Oxford Logic Guides 24

Introduction to Logic and to the Methodology of Deductive Sciences

Alfred Tarski

Fourth Edition
Edited by
Jan Tarski

Oxford Logic Guides

General Editors

Dov Gabbay
Angus Macintyre
Dana Scott

OXFORD LOGIC GUIDES

1. Jane Bridge: *Beginning model theory: the completeness theorem and some consequences*
2. Michael Dummett: *Elements of intuitionism*
3. A. S. Troelstra: *Choice sequences: a chapter of intuitionistic mathematics*
4. J. L. Bell: *Boolean-valued models and independence proofs in set theory* (1st edition)
5. Krister Segerberg: *Classical propositional operators: an exercise in the foundation of logic*
6. G. C. Smith: *The Boole-DeMorgan correspondence 1842-1864*
7. Alec Fisher: *Formal number theory and computability: a work book*
8. Anand Pillay: *An introduction to stability theory*
9. H. E. Rose: *Subrecursion: functions and hierarchies*
10. Michael Hallett: *Cantorian set theory and limitation of size*
11. R. Mansfield and G. Weitkamp: *Recursive aspects of descriptive set theory*
12. J. L. Bell: *Boolean-valued models and independence proofs in set theory* (2nd edition)
13. Melvin Fitting: *Computability theory: semantics and logic programming*
14. J. L. Bell: *Toposes and local set theories: an introduction*
15. Richard Kaye: *Models of Peano arithmetic*
16. Jonathan Chapman and Frederick Rowbottom: *Realtive category theory and geometric morphisms: a logical approach*
17. S. Shapiro: *Foundations without foundationalism*
18. J. P. Cleave: *A study of logics*
19. Raymond M. Smullyan: *Gödel's incompleteness theorems*
20. T. E. Forster: *Set theory with a universal set*
21. C. McLarty: *Elementary categories, elementary toposes*
22. Raymond M. Smullyan: *Recursion theory for metamathematics*
23. Peter Clote and Jan Krajíček: *Arithmetic, Proof Theory and Computational Complexity*
24. Alfred Tarski: *Introduction to Logic and to the Methodology of Deductive Sciences*, 4th edition edited by Jan Tarski

Introduction to Logic and to the Methodology of the Deductive Sciences

by Alfred Tarski
fourth edition edited by Jan Tarski

New York Oxford
OXFORD UNIVERSITY PRESS
1994

The original translation by Olaf Helmer was thoroughly revised for this edition.

Oxford University Press

Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Kuala Lumpur Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland Madrid

and associated companies in
Berlin Ibadan

New material in this Edition
Copyright © 1941, 1946, 1965, 1994 by Oxford University Press, Inc.
for the English Language

First American Edition, 1941
Second Edition, Revised, 1946
Third Edition, Revised, 1965
Fourth Edition, Revised, 1994

Published by Oxford University Press, Inc.,
200 Madison Avenue, New York, New York 10016

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press

Library of Congress Cataloging-in-Publication Data
Tarski, Alfred.
[O logice matematycznej i metodzie dedukcyjnej. English]
Introduction to logic and to the methodology
of the deductive sciences / by Alfred Tarski. — 4th ed. / edited by Jan Tarski.
p. cm. — (Oxford logic guides; 24)
Includes bibliographical references and index
ISBN 0-19-504472-X
1. Mathematics — Philosophy. 2. Arithmetic — Foundations.
I. Tarski, Jan. II. Title. III. Series
QA9.T28 1993
511.3 — dc20 93-6104

9 8 7 6 5 4 3 2 1
Printed in the United States of America
on acid-free paper

Contents

Second Part. Applications of Logic and Methodology in Constructing Mathematical Theories

From Author's Prefaces to Previous Editions

From the original edition (1936)

In the opinion of many laymen, mathematics is already today a dead science: after having reached an unusually high degree of development, it has become petrified in rigid perfection. This is an entirely erroneous view of the situation; there are but a few domains of scientific research which are passing through a phase of such intensive development at present as mathematics. Moreover, this development is extraordinarily manifold: mathematics is expanding its realm in all possible directions, it is growing in height, in width, and in depth. It is growing in height, since, on the soil of its old theories which look back upon hundreds if not thousands of years of existence, new problems appear again and again, and ever stronger and more refined results are being achieved. It is growing in width, since its methods permeate other branches of learning, its domain of investigation embraces ever wider ranges of phenomena, and new theories continue to enter into the large circle of mathematical disciplines. And finally, it is growing in depth, since its foundations become more and more firmly established, its methods perfected, and its principles stabilized.

It has been my intention in this book to give those readers who are interested in contemporary mathematics, but who are not actively concerned with it, at least a very general idea of that third line of mathematical development, i.e. of its growth in depth. My aim has been to acquaint the reader with the most important concepts of an independent discipline, which is known as mathematical logic, and which was created for the purpose of making the foundations of mathematics firmer and more profound; this discipline, in spite of its brief existence of barely a century, has already attained a high degree of perfection, and today its role in the totality of our knowledge transcends by far the originally intended boundaries.[†] It

[†]In this book, logic is treated primarily as a discipline which strengthens the foundations of mathematics. It would be natural, however, for a reader to ask whether this discipline may perhaps play a wider role, and certainly it was natural for the author to have been concerned with such matters. Cf. also the opening paragraph of the preface of 1941.—The author's remarks should appear in a clearer perspective when the reader takes note of a certain philosophical activity which was prominent, and with which the author was in close contact. In fact, the first publication of *Introduction to Logic*, in 1936, was about a decade after several philosophers and philosophically-minded scientists formed the Vienna Circle. Its members wrote extensively on various philosophical matters of broad significance, such as problems of language, knowledge, etc.; and as they strove for a maximum of clarity and precision, they employed freely the concepts of logic (and of the methodology of mathematics) in their discussions. Among its members were Carnap and Gödel, cf. footnote 8 on p. 130 and footnote 7

has been my intention to show that the concepts of logic permeate the whole of mathematics, that they comprehend all specifically mathematical concepts as special cases, and that logical laws are constantly applied—be it consciously or subconsciously—in mathematical reasonings. Finally, I have tried to present the basic principles employed in the construction of mathematical theories—principles which form the subject matter of still another discipline, the methodology of mathematics—and to show how one sets about using those principles in practice.

It has not been easy to carry this whole plan through within the framework of a relatively small book, especially without presupposing on the part of the reader any specialized mathematical knowledge, or any broader training in argumentation of an abstract character. Throughout the book a combination of the greatest possible intelligibility with the necessary conciseness had to be attempted, while taking a constant care for avoiding errors or cruder inexactitudes from the scientific standpoint. A language had to be used which deviates as little as possible from the language of everyday life. The employment of special logical symbolism had to be given up, although this symbolism is an invaluable tool which permits us to combine conciseness with precision, eliminates to a large extent the possibility of ambiguities and misunderstandings, and provides thereby an essential service in all subtler reasonings. The idea of a systematic treatment had to be abandoned from the beginning. Of the abundance of questions which present themselves, only a few could be discussed in detail, some could only be mentioned in passing, while still others had to be left out entirely, with the awareness that the choices which were made would inevitably exhibit a more or less arbitrary character. In those cases in which contemporary science has not yet taken any definite stand and offers a number of possible and equally correct solutions, it was out of the question to present objectively all known views. A decision in favor of a definite point of view had to be made. When making such a decision, I have taken care to choose a method of solution which would be simple insofar as possible, and which would lend itself to a popular mode of presentation, rather than one which would conform to my personal inclinations.—I do not have any illusion, that I have entirely succeeded in overcoming these and still other difficulties. [...]

---

on p. 128. Moreover, the interests and the premisses which were current among the members of the Vienna Circle were likewise adopted (to a greater or smaller extent) by a number of philosophers elsewhere, and in particular, by those who were associated with the Polish school; see *A Short Biographical Sketch of Alfred Tarski.*

From the first American edition (1941)

The present book is a partially modified and extended edition of my book *On Mathematical Logic and Deductive Method*, which appeared first in 1936 in Polish and then in 1937 in an exact German translation (under the title: *Einführung in die mathematische Logik und in die Methodologie der Mathematik*). In its original form it was intended as a popular scientific book; its aim was to present to the educated layman—in a manner which would combine scientific exactitude with the greatest possible intelligibility—a clear idea of that powerful trend of contemporary thought which is concentrated around modern logic. This trend arose originally from the somewhat limited task of stabilizing the foundations of mathematics. In its present phase, however, it has much wider aims. For it [aspires to relate to] the whole of human knowledge. In particular, one of its goals is to perfect and to sharpen the deductive method, which [not only has a central place in mathematics, but in addition, in just about] every domain of intellectual endeavor, serves as an indispensable tool for deriving conclusions from accepted assumptions.

The response accorded to the Polish and German editions, and especially some suggestions made by reviewers, gave rise to the idea of making the new edition not only a popular scientific book, but also a textbook upon which an elementary college course in logic and the methodology of deductive sciences could be based. The experiment seemed all the more desirable, in view of a certain lack of suitable elementary textbooks in this domain.

In order to carry out the experiment, it was necessary to make several changes in the book.

Some very fundamental questions and notions were entirely passed over or only touched upon in the previous editions, either because of their more technical character, or in order to avoid points of a controversial nature. As examples may be cited such topics as the difference between the use of certain terms in the context of logic and their use in the language of everyday life, the general method of verifying laws of the sentential calculus, the necessity of a sharp distinction between words and their names, the concepts of the universal class and the null class, the fundamental notions of the calculus [or, of the algebra] of relations, and finally the conception of methodology as a general science of sciences. In the present edition all of these topics are discussed (although not all in an equally thorough manner), since it seemed to me that to avoid them would constitute an essential gap in any textbook of modern logic. [...]

In contrast to the previous editions, in the present book I considered it necessary to familiarize the reader with the elements of logical symbolism. Nevertheless, the use of this symbolism in practice remains very restricted, and is limited mostly to exercises.

In the previous editions, the principal domain from which examples were drawn for illustrating general and abstract considerations was high-school mathematics; for it was my opinion, and still is, that elementary mathematics, and especially algebra, with its simple concepts and uniform methods of proof, is very appropriate for this purpose. Nevertheless, in the present edition, particularly in the newly added passages, I draw examples more frequently from other domains, especially from everyday life. [...]

The essential features of the book remain unchanged. The preface to the original edition, the major part of which [precedes the present preface], will give the reader an idea of the general character of the book. Perhaps, however, it is worthwhile to point out explicitly at this place what he or she should not expect to find in it.

First, the book contains no systematic and strictly deductive presentation of logic; such a presentation obviously would not lie within the framework of an elementary textbook. Originally it was my intention to include, in the present edition, an additional chapter entitled *Logic as a Deductive Science*, which—as an illustration of the general methodological discussion contained in Chapter VI—would outline a systematic development of some elementary parts of logic. For a number of reasons this intention could not be realized; but several exercises on this subject are included in Chapter VI, [and they should] to some extent compensate for the omission.

Secondly, apart from two rather short passages [and a few related exercises in Chapter VI], the book gives no information about the traditional Aristotelian logic, and contains no material drawn from it. But I believe that the space here devoted to traditional logic corresponds well enough to the small role, to which the place of this logic in modern science has been reduced; and I also believe that this opinion will be shared by most contemporary logicians.

And, finally, the book is not concerned with any problems belonging to the so-called logic and methodology of empirical sciences. I must say that I am inclined to doubt whether any special "logic of empirical sciences", as opposed to logic in general, or, to the "logic of deductive sciences", exists at all (at least so long as the word "*logic*" is used as in the present book—that is to say, as the name of a discipline which analyzes the meaning of the concepts common to all sciences, and establishes general laws governing these concepts). But this is a terminological rather than a factual problem. At any rate, the methodology of empirical sciences constitutes an important domain of scientific research. The knowledge of logic is of course valuable in the study of this methodology, as it is in the case of any other discipline. It must be admitted, however, that up to the present, logical concepts and methods have not found any specific or fertile applications in this domain. And it is certainly possible that this state of affairs is not simply a reflection of the present stage of methodological research. Perhaps it arises from the circumstance that, for the purpose of an adequate methodological treatment, an empirical science may have to be considered not simply as a scientific theory—that is, as a system of asserted statements arranged according to certain rules—, but rather as a complex, consisting partly of such statements and partly of human activities. It should be added that, in striking opposition to the high development of the empirical sciences themselves, the methodology of these sciences can hardly boast of comparably definitive achievements—despite the great efforts that have been made. [...] For these and other reasons, I see little rational justification for combining the discussion of logic and that of the methodology of empirical sciences in the same college course.

A few remarks concerning the arrangement of the book and its use as a college text.

The book is divided into two parts. The first gives a general introduction to logic and to the methodology of deductive sciences; the second shows, by means of a concrete example, the sort of applications which logic and methodology find in the construction of mathematical theories, and thus affords an opportunity to assimilate and to deepen the knowledge which is acquired in the first part. Each chapter is followed by appropriate exercises. Brief historical indications are contained in footnotes.

Passages, and even whole sections, which are set off by asterisks "*" both at the beginning and at the end, contain more difficult material or presuppose familiarity with other passages containing such material; they can be omitted without jeopardizing the intelligibility of subsequent parts of the book. An analogous remark applies to the exercises whose numbers are preceded by asterisks.

I feel that the book contains enough material for a full-year course. Its arrangement, however, makes it feasible to use it in half-year courses as well. If used as a text in half-year logic courses in a department of philosophy, I suggest a thorough study of its first part, including the more difficult portions, with omission of the entire second part. If the book is used in a half-year course in a mathematics department—for instance, in the foundations of mathematics—, I suggest the study of both parts of the book, with omission of the more difficult passages.

In any case, I should like to emphasize the importance of working out the exercises carefully and thoroughly; for they not only facilitate the assimilation of the concepts and the principles which are introduced, but also touch upon many problems for the discussion of which the text provides no opportunity.

I shall be very happy if this book contributes to a wider diffusion of logical knowledge. [...] For logic, by perfecting and by sharpening the tools of thought, makes men and women more critical—and thus makes less likely their being misled by all the pseudo-reasonings to which they are incessantly exposed in various parts of the world today.

Alfred Tarski

September 1940

The editor should like to mention those individuals to whom the author expressed (in the previous American editions) thanks and gratitude: Mr. K. J. Arrow, Dr. O. Helmer, Professor L. Henkin, Dr. A. Hofstadter, Mr. L. K. Krader, Professor Louise H. Lim (Chin), Dr. J. C. C. McKinsey, Professor E. Nagel, Miss Judith Ng, Professor W. B. Pitt, Professor W. V. Quine, Mr. M. G. White, Dr. P. P. Wiener.

Editor's Preface

Alfred Tarski's *Introduction to Logic and to the Methodology of Deductive Sciences* has become recognized as a classic. It was first published in Polish over 50 years ago, and was then the first popular-scientific book to be devoted to modern logic; its appreciation has grown over the years. It has been translated into eleven languages. It is still timely and continues to draw attention.

The English-language text has been out of print for a number of years. After the author's death in 1983 a discussion took place about publishing a new edition of the book. Two kinds of changes were then considered. On one hand, some modifications in the existing text seemed desirable, and on the other, a possible expansion of the book was envisaged.

Considerations of expansion led to plans for a forthcoming companion volume. These plans in turn affected some aspects of the present book.

For preparation of the new text, we naturally started with the latest American edition: the third, of 1965. (Its main text differs only slightly from the texts of the first two editions.) Our original intention was to keep the editorial changes to a minimum. Nonetheless, the changes turned out to be fairly numerous, and of several kinds.

(i) In Chapter V a change of convention was made, regarding the passage from a relation to a function. A function is now defined as a many-one relation, so that upon starting with "xRy", we are led to the functional formula: "$f(x) = y$".—This convention follows the original Polish text. In the previous American editions, instead, a function was defined as a one-many relation, and one was led to "$x = f(y)$". The difference is trivial, but it affects several sections of the chapter, and it required modifying not only formulas but also the discussion.

(ii) Some of the exercises in Chapter VI were intended as an introduction to a deductive development of logic. Their content will now be utilized in the companion volume, and the exercises were therefore removed. New exercises, which are of the same general kind but are somewhat more elementary, are provided in their place.

(iii) The section *Suggested Readings* was removed. An edited version of this section will be included in the companion volume.

(iv) A new section, *A Short Biographical Sketch of Alfred Tarski*, is included.

(v) A number of editor's footnotes were added. They are marked with daggers: "†", "‡".

(vi) Notation was changed so as to conform to the prevalent usage of today. This applies especially to the notation for quantifiers and for classes.

The terminology was likewise largely modernized. One important change was the replacement of "*biunique function*" by "*bijective function*". Furthermore, a few historical and other remarks were no longer appropriate, and they were brought up to date.

(vii) A few excerpts were set in italics for emphasis, even though they are not displayed.

(viii) There are approximately 25 sentences whose content was altered (including some in exercises and in the author's preface of 1941).

(ix) Extensive stylistic changes were made throughout the book.

(x) Three asterisks are now used to signal a change of topic within a section. The gaps which were employed for this purpose in the previous editions had several drawbacks.

(xi) A few simple adjustments were made in section titles. These adjustments relate to the modernization of terminology and to stylistic changes (points (vi) and (ix)).

(xii) Finally, the index was revised so as to take into account the various changes, and was modified somewhat otherwise.

We should like to comment on points (viii) and (ix). It is always difficult to eliminate completely the rough spots in a book, and in this case there were particular adverse factors. The English text was produced as a "translation of a translation", with some additions. This work was carried out under the pressure of time, and under otherwise difficult circumstances. Finally, the subsequent editions were photographic reprints, in which one could accommodate only minimal changes of the main text.

For an example of a change of content we refer to Section 19, where we modified the author's critical remarks concerning the traditional usage of the term "*equality*" in geometry.

With regard to matters of style, the previous text clearly lost some quality by the double-translation effect. As example, the word "*thing*" appeared throughout the book, with reference to numbers, classes, etc. The original Polish edition had "*przedmiot*", meaning "*object*", this was translated into German as "*Ding*", and then into English as "*thing*". This last word has now been replaced throughout by "*object*".

While going over the previous English text we checked numerous passages against the original Polish, and in a number of cases we attempted to provide an improved translation.

For the editorial revision of the main text, footnotes, and exercises, the following rules of procedure were adopted and strictly followed: Aside from the revisions described in (i) and (ii), there are no changes other than those involving individual sentences. No sentences were deleted, no sentences were rearranged, and no sentences were added except as editor's footnotes. Moreover, the structure of each paragraph was left intact. (Two exceptions were made: In Section 57 the division into paragraphs was changed, so as to conform to the original edition, and in Section 58 the former statement

(III) now has the place (V). In addition, the foregoing rules were modified slightly for the prefaces; stylistic changes were made there without comment, as elsewhere, but other alterations are indicated.) We hoped that by proceeding in this way, the editorial revision would not distort the over-all plan of the author nor the portrayal of his pattern of thought.

Let us return to the possible expansion. The author made a step in that direction when he adapted his article *Truth and Proof* for the fifth German edition of the book. He mentioned in the preface of 1941, moreover, certain thoughts about an additional chapter: *Logic as a Deductive Science.*

A project was set up, accordingly, by Dana Scott and the present editor sometime after the discussion of 1983 to arrange an appropriate text for a new edition of *Introduction to Logic.* This text was to incorporate needed alterations to the existing version, and to include several of the author's articles as well as some new material. As the project progressed, it became apparent that the combination of the articles and the new material would be better suited for an independent volume, and the plan for a single enlarged book was dropped.

At the present time we are optimistic that the new book will appear fairly soon, and that it will be welcome as a continuation of *Introduction to Logic* and as its companion volume.

Unfortunately Scott had to limit his participation in the project rather substantially after the initial stages. In particular, the editorial work on *Introduction to Logic* was done by the present editor.

* * *

My thanks go, in the first place, to Professor Dana Scott, who took the initiative regarding the expansion and the revision of the original text. While his subsequent participation was greatly reduced (as mentioned), his advice was always very valuable. He, moreover, arranged for extensive reworkings of the book on a computer at Carnegie Mellon University. I also thank Professor Leon Henkin, who provided warm encouragement to undertake the project, and Professor Steven Givant, who helped with many clarifying discussions. Furthermore, Professor John Corcoran read critically a version of the revised text and, having had long experience in teaching elementary logic, suggested many improvements and corrections; his help is deeply appreciated. Finally, the book was composed in the system TEX; several individuals contributed to carrying out this task. I am very grateful to them, and to Oxford University Press, for their care in producing this book.

J. T.

Berkeley, California, March 1993

Photograph of Alfred Tarski — June 1953

Alfred Tarski

A Short Biographical Sketch of Alfred Tarski

Alfred Tarski was born in Warsaw, Poland, on January 14, 1901. He entered the University of Warsaw in the fall of 1918, and completed his doctorate in the spring of 1924 under the direction of the logician S. Leśniewski. He remained at the University of Warsaw until 1939.

In the 1920's two centers of learning in Poland gained a particular recognition, for logic and philosophy on one hand, and for mathematics on the other. They were at the two cities, Warsaw and Lwów (now: Lviv), and are referred to jointly as the Polish school (of logic, etc.). Tarski took advantage of the opportunities that these centers offered, and, with his extraordinary talents and abilities, he quickly became one of the leading members.

The period 1923–1939 was very fruitful for Tarski, and a number of important works by him appeared at that time. Several of these works were devoted to sentential calculus, to set theory, and to topics of methodology (such as the foundations of metamathematics). One article, which involves set theory as well as mathematics proper, established the well-known Banach–Tarski paradox.[†] Furthermore, Tarski started to pursue other problems, and came upon many of the ideas which he developed later in life.

Among Tarski's works of that period, the monograph on truth has a central place; he dealt there with a problem of classical philosophy (dating back to Aristotle). This investigation depended on the techniques of methodology, but otherwise was outside the main currents of mathematics. It had its background in the extensive lectures and discussions of Tarski's teachers: Kotarbiński, Leśniewski, and Łukasiewicz; the first-mentioned of these was a philosopher rather than a logician.

In this monograph Tarski showed, for the case of a language whose sentences are suitably limited, how one can give a precise definition of the phrase "*is a true sentence*". This work brought to a sharp focus several concepts and ingredients which are essential for an analysis of the notion of truth. These concepts and ingredients are of value for other kinds of problems as well.— Tarski had various contacts with philosohy from the time of his university studies, and these contacts became a great deal broader after the appearance of the monograph. For instance, he participated in a number of philosophical congresses. He moreover contributed other articles

[†]This paradox (which Tarski discovered jointly with the Polish mathematician S. Banach) describes the theoretical possibility of transforming a sphere into two spheres, each the same size as the original. This possibility, which depends on peculiarities of infinite sets, was noted on various occasions in popular writing.

on semantics; some of these are technical, while some were intended for a wider audience. Basically, however, he did not have the disposition on a philosopher, but rather, had that of a mathematician.

During the summer of 1939 Tarski attended a scientific congress in the USA. When the Second World War broke out, he had to forego returning to Poland. In 1942 he joined the faculty of the University of California in Berkeley, where he remained until his death on October 27, 1983.

In Berkeley, Tarski continued his research, which then dealt with a rich variety of topics in the foundations of mathematics. In particular, he extended his earlier work on completeness and decision problems,[††] and also on the axiomatization of geometry. He, moreover, carried out extensive investigations which lie on the borderline between algebra and logic. His work on the theory of models, in his later years, falls into this category and is regarded as one of his greatest achievements.

Especially noteworthy among Tarski's publications is the book of 1956: *Logic, Semantics, Metamathematics*, which contains a selection of his articles of more general interest, from the period 1923–1938 and in English translation, and which summarizes a significant part of his career.—We may note in this connection, that in science one seldom reads the original article after a lapse of a few years. By then one can usually find the description of a consequential discovery in a more attractive form, in a review or in a book. A collection of original articles may nonetheless be valuable, for historical or other purposes; a second edition of such a collection, however, would be a great rarity. In the case of Tarski's book, the second edition (which appeared in 1983) provides strong evidence for the exceptional value of the articles.[‡]

Tarski's credits extend far beyond his publications. He was highly influential among his colleagues, and was very successful as a teacher. He had over 20 doctoral students, of whom unusually many became leading logicians. He moreover built a program of logic in Berkeley, a program

[††]Tarski's construction of decision procedures for elementary theories of arithmetic of real numbers and geometry, i.e. the construction of algorithms by which one can prove or disprove any given sentence of these disciplines, remains a landmark among investigations of decision problems. (Cf. the footnotes on pp. 126–127.) It should be said, however, that the procedures in question are quite complicated, and are hardly suitable for extensive applications. In particular, the procedure for geometry would not simplify at all the proofs of the standard theorems of high-school geometry.—On the other hand, the interest in algorithms has grown with the increased use of computers, and this applies to algorithms of all kinds, whether they are exploited in practice or not. The decision procedures in question can therefore claim a place within *theoretical computer science* as well as within methodological studies.

[‡]The preparation of the two editions required a considerable effort on the part of several of Tarski's colleagues. In particular, the British scientist J. H. Woodger took the initiative in this direction and translated the articles. The second edition appeared under the care of the American philosopher J. Corcoran.

which attracted students and scholars from all over the world. Now, nearly a decade after his death, Berkeley remains an important center for the foundations of mathematics.

In 1971 an international symposium was held in his honor in Berkeley. In recognition of his achievements he was elected to the National Academy of Sciences, to the British Academy, and to the Dutch Academy. He received three honorary doctorates.

It would be natural to ask: What kind of a person was Tarski? How did his friends and colleagues find him? What did he think? Such questions are never easy to answer. We hope, however, that the few details which follow, and which are largely unrelated, will be of some interest.

Tarski used to emphasize the importance of quality in mathematical writing, and in particular, the need for bringing out intuitive aspects. Accordingly, in reviewing doctoral dissertations which were submitted to him, he always looked critically at the form of the exposition. It was not enough just to have statements of the results, with proofs. For instance, the significance of the problem had to be explained clearly. The definitions not only had to be stated with precision, but insofar as possible, had to be intuitively understandable.

The emphasis which has just been described served various pedagogical purposes, and reflected Tarski's linguistic sensitivity (and of course, his over-all high standards). In addition, it reflected his profound attitude, that research should have an esthetic aspect, that it should play a role in fulfilling esthetic needs.—And he appreciated, naturally, also other sources of esthetic expression, in particular poetry and visual arts. He enjoyed, for instance, reciting from his favorite poets (in several languages).

A different matter: Tarski had broad interests, as the reader may have guessed from the foregoing; he moreover liked to talk at length on a variety of subjects. Furthermore, like many people from Poland, he showed a deep emotional involvement about politics; this was apparent especially during his years in Berkeley. He was distressed at the sad political situation that was spread over much of the globe. On occasions he would repeat, like a prophet: One day we will wake up and will see that all these beastly regimes are gone, that they have faded like a horrible nightmare.

As to his nature, he had a strong will and immense self-discipline, and could concentrate on his work to the exclusion of (just about) everything else. This was a factor which certainly helped him in achieving the numerous successes, but which also made him a difficult person to deal with.

As to his diversions, he particularly enjoyed gardening. This perhaps was to him a reminder of his first intellectual passion, which was for biology. In fact, it was only in the course of his university studies that he decided to switch to logic.

Let us return to Tarski's early life. He was the elder of two sons in a

Jewish family. His parents did not have distinguished careers. His mother, however, was described as a woman with a superb memory. No doubt Tarski inherited much of his talent from her.

His secondary education (at *gymnasium*) had standards of excellence which could not often be found. For instance, his curriculum included a total of five languages, modern and classical. (Moreover, because of the political situation, he could study Polish only by a special arrangement, and he had lessons of Hebrew as well.) Several of the teachers at that school, in fact, were very able individuals; one of his teachers was an enterprising man who succeeded in obtaining a doctorate in mathematics by an independent effort.—At the university Tarski studied under professors who became prominent in years to come. We already mentioned those in logic and philosophy. Among his mathematics teachers were Mazurkiewicz and Sierpiński, both of whom made significant contributions to the theory of sets.

In the first part of his career he encountered two major hardships. His university position as docent[‡‡] provided only a meager income, which he had to supplement by teaching at a *gymnasium*. In that period he had one hope for a change, when a new professorship in mathematical logic was created (in 1928) at Lwów. He however lost the competition for this professorship, and in the succeeding years no other positions were available which would be appropriate for him. He therefore had to reconcile himself with such uneasy arrangements for another decade.—And then, in September 1939, having come to the USA, he found himself separated from his family and close friends in Poland. His wife and two children were unable to join him until the first days of 1946.

When he came to Berkeley, he was still a man of the Old World, and his habits and ways of thinking were deeply engrained. In some respects he adapted quickly to the new surroundings, in some others much less so. He continued to have a special feeling for his days in Warsaw, where he spent his childhood and later enjoyed the period of his greatest creativity. There he found many difficulties, but also a memorable intellectual atmosphere: very stimulating, and having a kind of a magical quality about it as well.

In Berkeley he discovered such attractions as the scenery of the American West within easy reach and an agreeable climate. He came to appreciate the prevailing informality. Indeed, he acquired a great liking for Berkeley and its vicinity, and except for one sabbatical year, he never left the city for long.

[‡‡] A position of docent was approximately at the level of associate professor, but without a regular salary (and without tenure). The docent was paid primarily for his lectures. Such positions were common in former years, and can also be found today.

Introduction to Logic

This page intentionally left blank

First Part

Elements of Logic
Deductive Method

This page intentionally left blank

I

On the Use of Variables

1 Constants and variables

Every scientific theory is a system of sentences which are accepted as true and which may be called LAWS or ASSERTED STATEMENTS (sometimes one says, for short, simply STATEMENTS). In mathematics these statements follow one another in a definite order, and in accordance with certain principles which will be discussed in detail in Chapter VI; in view of these principles, the statements are generally accompanied by arguments whose purpose is to demonstrate their truth. Arguments of this kind are referred to as PROOFS, and the statements established by them are called THEOREMS.

Among the terms and symbols occurring in mathematical theorems and proofs, we distinguish CONSTANTS and VARIABLES.

In arithmetic, for instance, we encounter such constants as "*number*", "*zero*" ("0"), "*one*" ("1"), "*sum*" ("+"), and many others.[1] Each of these terms has a well-determined meaning, which remains unchanged throughout the course of the considerations.

As variables we employ, as a rule, selected letters; e.g. in arithmetic we generally choose a few lower-case letters of the Latin alphabet: "*a*", "*b*",

[1] By "*arithmetic*" we shall here understand that part of mathematics which is concerned with the investigation of general properties of numbers, relations among numbers, and operations on numbers. In place of the word "*arithmetic*" the term "*algebra*" is frequently used, particularly in high-school mathematics. We have preferred the term "*arithmetic*" because, in higher mathematics, the term "*algebra*" is reserved for the much more special theory of algebraic equations. (Some time ago the term "*algebra*" acquired a wider meaning, which, however, still is different from that of "*arithmetic*".)—In this book the term "*number*" will always be used with the meaning which is normally attached to the term "*real number*" in mathematics; that is to say, it will cover integers and fractions, rational and irrational numbers, positive and negative numbers, but not imaginary or complex ones.

"c", ..., "x", "y", "z". As opposed to the constants, the variables do not possess any meaning by themselves. Thus, the question:

does zero have such and such a property?

e.g.:

is zero an integer?

can be answered in the affirmative or in the negative; the answer may be true or false, but at any rate it is meaningful. A question about x, on the other hand, for example the question:

is x an integer?

cannot be answered meaningfully.

* * *

In some textbooks of elementary mathematics, particularly in the less recent ones, one does occasionally come across formulations which convey the impression that it is possible to attribute an independent meaning to variables. One might find an explanation that the symbols "x", "y", ... also denote certain numbers or quantities, not "constant numbers" however (which are denoted by constants such as "0", "1", ...), but so-called "variable numbers" or "variable quantities". Statements of this kind stem out of a gross misunderstanding. The "variable number" x which one tries to envisage could not possibly have any specified property, for instance, it could be neither positive nor negative nor equal to zero; or rather, the properties of such a number would have to change from case to case, that is to say, the number would sometimes be positive, sometimes negative, and sometimes equal to zero. But entities of such a kind are not to be found in our world at all; their existence would contradict the fundamental laws of thought. The classification of the symbols into constants and variables, therefore, does not correspond to any of the familiar classifications of the numbers.

2 Expressions containing variables—sentential and designatory functions

In view of the fact that variables do not have a meaning by themselves, such phrases as:

x is an integer

are not sentences, although they have the grammatical form of sentences; they do not express definite assertions and can be neither confirmed nor refuted. We can say, however, that from the expression:

x is an integer

we obtain a sentence when we replace "x" in it by a constant denoting a definite number; for instance, if "x" is replaced by the symbol "1", the result is a true sentence, while a false sentence arises on replacing "x" by "$\frac{1}{2}$". An expression of this kind, which contains variables and becomes a sentence on replacement of these variables by constants, is called a SENTENTIAL FUNCTION. But mathematicians, by the way, are somewhat reluctant to employ this expression, because they use the term "*function*" in a different sense. More often the word CONDITION is employed instead; and sentential functions and sentences which are composed entirely of mathematical symbols (and not of words of everyday language), such as:

$$x + y = 5,$$

are usually referred to by mathematicians as FORMULAS. In place of "*sentential function*" we shall sometimes simply say "*sentence*"—but only in cases where there is no danger of any misunderstanding.

The role of the variables in a sentential function at times was compared, very appropriately, with that of the blanks which are in a questionnaire; just as the questionnaire acquires a definite content after the blanks have been filled in, so a sentential function becomes a sentence after constants have been inserted in place of the variables. The replacement of the variables in a sentential function by constants—equal constants taking the place of equal variables—may lead to a true sentence; in such a case, the objects denoted by those constants are said to SATISFY the given sentential function. For example, the numbers 1, 2, and $2\frac{1}{2}$ satisfy the sentential function:

$$x < 3,$$

but the numbers 3, 4, and $4\frac{1}{2}$ do not.†

* * *

Besides the sentential functions, there are certain other expressions which contain variables and merit our attention, namely, the so-called DESIGNATORY or DESCRIPTIVE FUNCTIONS. They are expressions which, upon replacement of the variables by constants, turn into designations (or, descriptions) of objects. For example, the expression:

$$2x + 1$$

is a designatory function, because we obtain the designation of a certain number (e.g., the number 5), if in it we replace the variable "x" by an

---

† In order to illustrate the provision, that of equal constants taking the place of equal variables, one may consider the inequality:

$$x^2 + x + y^2 > 7.$$

This inequality is satisfied, for instance, when "x" is replaced (in both places) by "3" and "y" by "1". It is not satisfied when both "x" and "y" are replaced by "1".

arbitrary numerical constant, that is, by a constant denoting a number (e.g. "2"; we assume the usual convention, that "$2x$" means "$2 \cdot x$").

Among the designatory functions occurring in arithmetic we have, in particular, all so-called algebraic expressions, which are composed of variables, numerical constants, and symbols of the four fundamental arithmetical operations; these expressions are such as the following††:

$$x - y, \quad \frac{x+1}{y+2}, \quad 2 \cdot (x + y - z).$$

On the other hand, algebraic equations, which are formulas consisting of two algebraic expressions connected by the symbol "=", are sentential functions. For discussing equations, a special terminology has been adopted in mathematics; in particular, the variables occurring in an equation are referred to as the unknowns, and the numbers satisfying the equation are called the roots of the equation. E.g., in the equation:

$$x^2 + 6 = 5x$$

the variable "x" is the unknown, while the numbers 2 and 3 are the roots of the equation.

* * *

Of the variables "x", "y", ... employed in arithmetic, one says that they STAND FOR DESIGNATIONS OF NUMBERS or that numbers are VALUES of these variables. These phrases mean approximately the following: a sentential function containing the symbols "x", "y", ... becomes a sentence, if these symbols are replaced by constants which designate numbers (and not by expressions designating operations on numbers or relations among numbers, nor by expressions designating objects outside the field of arithmetic like geometrical configurations, animals, plants, etc.). The variables occurring in geometry, likewise, stand for designations of points and of geometrical figures. Moreover, the concept of designation applies to compound expressions, and so the designatory functions which we meet in arithmetic, like the variables themselves, stand for designations of numbers.—Sometimes one says briefly that the symbols "x", "y", ... themselves, as well as the designatory functions constructed with their help, denote numbers or are designations of numbers, but this is simply an abbreviative terminology.

††Sometimes the word "*formula*" is used also for a designatory function which is expressed in symbols only. For instance, one might say that "ab" is the formula for the area of a rectangle, where a and b are the lengths of two adjacent sides. We see therefore two uses for this particular word, involving two different kinds of expressions. This double usage should not cause any difficulties, but it points to the need for keeping the distinction between the two kinds of expressions firmly in mind.

3 Construction of sentences in which variables occur—universal and existential sentences

Apart from the replacement of variables by constants, there is also another way in which sentences can be obtained from sentential functions. Let us consider the formula:

$$x + y = y + x.$$

It is a sentential function which contains the two variables "x" and "y" and is satisfied by an arbitrary pair of numbers; no matter what numerical constants we put in place of "x" and "y", we always obtain a true formula. We express this fact briefly in the following manner:

for any numbers x and y, $x + y = y + x$.

The expression which has just been given is a genuine sentence and, indeed, a true sentence; we recognize it as one of the fundamental laws of arithmetic, namely the commutative law for addition. Many important theorems of mathematics are formulated in a similar way; that is, they are instances of so-called UNIVERSAL SENTENCES, or, of SENTENCES OF UNIVERSAL CHARACTER, which assert that arbitrary objects of a certain category (e.g., in the case of arithmetic, arbitrary numbers) have such and such a property.—One should be aware that in the formulation of universal sentences, the phrase "*for any objects* (or *numbers*) x, y, ..." is often omitted and has to be inserted mentally; for instance, the commutative law for addition may simply be given in the following form:

$$x + y = y + x.$$

This has become a well-accepted usage, to which we shall generally adhere in our future discussions.

* * *

Let us now consider the sentential function:

$$x > y + 1.$$

This formula fails to be satisfied by every pair of numbers; if, for instance, "3" is put in place of "x" and "4" in place of "y", the false sentence:

$$3 > 4 + 1$$

is obtained. Therefore, if one says:

for any numbers x and y, $x > y + 1$,

one does undoubtedly state a meaningful sentence, but one which is obviously false. On the other hand, there are pairs of numbers which satisfy the sentential function under consideration; for example, if "x" and "y" are replaced by "4" and "2", respectively, the result is the true formula:

$$4 > 2 + 1.$$

The fact that such numbers exist is expressed briefly by the following phrase:

for some numbers x and y, $x > y + 1$,

or, using a form which is more frequently employed:

there are numbers x and y such that $x > y + 1$.

The expressions just given are true sentences; they are examples of EXISTENTIAL SENTENCES, or, of SENTENCES OF EXISTENTIAL CHARACTER, stating the existence of objects (e.g. numbers) with a certain property.

With the help of the methods just described, we can obtain sentences from any given sentential function; but it depends on the content of the sentential function whether we arrive at a true or at a false sentence. The following example may serve as a further illustration. The formula:

$$x = x + 1$$

is satisfied by no number; hence, no matter whether the words "*for any number x*" or "*there is a number x such that*" are prefixed, the resulting sentence will be false.

* * *

In contradistinction to sentences of universal or existential character, we may refer to sentences not containing any variables, such as:

$$3 + 2 = 2 + 3,$$

as SINGULAR SENTENCES. This classification is not at all exhaustive, since there are many sentences which cannot be placed into any of the three categories mentioned above. We include as example the following statement:

for any numbers x and y there is a number z such that
$x = y + z$.

Sentences of this type are called UNIVERSAL-EXISTENTIAL SENTENCES (as opposed to the existential sentences considered before, which may also be called ABSOLUTELY EXISTENTIAL SENTENCES); they assert the existence of numbers having a certain property, but on condition that certain other numbers (which can be arbitrary) have been chosen first.

4 Universal and existential quantifiers; free and bound variables

Phrases like:

for any x, y, ...

and

there are x, $y, \ldots$ such that

are called QUANTIFIERS; the former is said to be a UNIVERSAL, the latter an EXISTENTIAL QUANTIFIER. One also thinks of quantifiers as OPERATORS; however, there are expressions which are likewise considered to be operators but which are not quantifiers.—In the preceding section we tried to explain the meaning of both kinds of quantifiers. Their significance depends on the fact that an expression containing variables can be a sentence, that is, the statement of a well-determined assertion, only if quantifiers or other operators are employed (in an explicit or in an implicit manner). Without the help of operators, the usage of variables in the formulation of mathematical theorems would be ruled out.

* * *

In everyday language it is not customary to use variables (though in principle one could), and for this reason quantifiers are also not in use. There are, however, certain words in general usage which exhibit a very close connection with quantifiers, namely, such words as "*every*", "*all*", "*a certain*", "*some*". The connection becomes obvious when we observe that expressions like:

all men are mortal

or

some men are wise

have about the same meaning as sentences which are formulated with the help of quantifiers, according to the following schemes:

for any x, if x is a man, then x is mortal

and

there is an x, such that x is both a man and wise,

respectively.

* * *

For the sake of convenience, quantifiers are often replaced by symbolic expressions. Accordingly, let us agree to write in place of:

for any objects (or *numbers*) x, $y, \ldots$

and

there exist objects (or *numbers*) x, $y, \ldots$ *such that*

the following symbolic expressions:

$$\forall_{x,y,\ldots} \quad \text{and} \quad \exists_{x,y,\ldots}$$

respectively (with the understanding that the sentential functions which follow the quantifiers are put inside parentheses). Then, for instance, the statement which was given at the end of the preceding section as an example of a universal-existential sentence assumes the following form:

(I) $$\forall_{x,y}, \exists_z (x = y + z).$$

* * *

A sentential function in which the variables "x", "y", "z", ... occur automatically becomes a sentence as soon as one prefixes to it one or several operators containing all those variables. However, if some of the variables do not occur in the operators, then the expression in question remains a sentential function, without becoming a sentence. For example, the formula:

$$x = y + z$$

changes into a sentence if preceded by one of the phrases:

for any numbers x, y, and z;

there are numbers x, y, and z such that;

for any numbers x and y, there is a number z such that;

and so on. But if we prefix only the quantifier:

there is a number z such that or $\exists_z$

we do not yet arrive at a sentence; however, the resulting expression, that is:

(II) $$\exists_z (x = y + z)$$

is undoubtedly a sentential function, for it immediately becomes a sentence when we substitute appropriate constants in place of "x" and "y" and leave "z" unaltered, or else, when we prefix another suitable quantifier, e.g.:

for any numbers x and y or $\forall_{x,y}$.

We therefore see that, among the variables which may occur in a sentential function, two different kinds can be distinguished. The presence of variables of the first kind—they will be called FREE OR REAL VARIABLES—indicates that the expression under consideration is a sentential function and not a sentence; in order to convert the given sentential function to a sentence, it is necessary to replace these variables by constants, or else to put operators that contain these variables in front of the sentential function (or to combine the two kinds of modification). The remaining, so-called BOUND OR APPARENT VARIABLES, however, should not be changed in such a transformation. In the above sentential function of (II), for instance, "x" and "y" are free variables, while the symbol "z" occurs twice as a bound variable; on the other hand, expression (I) is a sentence, and thus contains bound variables only.

* * *

*It depends entirely upon the structure of the sentential function, that is, upon the presence and the position of the operators, whether any particular variable which occurs in it is free or bound. This may best be seen by means of a concrete example. Let us consider, for instance, the following sentential function:

(III) *for any number* x, *if* $x = 0$ *or* $y \neq 0$, *then there exists a number* z *such that* $x = y \cdot z$.

This function begins with a universal quantifier containing the variable "x", and therefore the latter, which occurs three times in this function, occurs at all these places as a bound variable; at the first place it constitutes a part of the quantifier, while at the other two places it is, as we say, BOUND BY THE QUANTIFIER. The situation is similar with respect to the variable "z". For, although the initial quantifier of (III) does not contain this variable, we can nevertheless recognize that a part of (III) opens with an existential quantifier containing the variable "z"; this is the sentential function:

(IV) *there exists a number* z *such that* $x = y \cdot z$.

Both places where the variable "z" occurs in (III) belong to the partial function which is stated in (IV). We therefore say that "z" occurs everywhere in (III) as a bound variable; at the first place it is a part of the existential quantifier, and at the second place it is bound by that quantifier. As to the variable "y" also occurring in (III), we see that there is no quantifier in (III) which contains this variable, and therefore it occurs in (III) twice as a free variable.

The fact that quantifiers bind variables—that is, that they change free into bound variables in the sentential functions which follow them—constitutes a very important property of the quantifiers. Several other expressions are known which have an analogous property; with some of them we shall become acquainted later (in Sections 20 and 22), while some others—such as, for instance, the integral sign—are central in calculus and in various other branches of higher mathematics. The term "*operator*" is the general term which is used to denote all expressions having this property.†*

†This may be an opportune place to comment on two of the terms which have been introduced. The word "*operator*" has various connotations; it is used in a number of contexts, in particular in logic, in mathematics, in everyday life. To avoid ambiguities, one can refer to expressions under consideration as *variable-binding operators*.—With regard to "*sentential functions*" (or, "*conditions*"), other terms have also been used with the same meaning. For instance, one finds the analogous phrase "*propositional functions*" in *Principia Mathematica* (cf. footnote 1 on p. 18). Moreover, such phrases have recently been called OPEN SENTENCES.

5 The importance of variables in mathematics

As we saw in Section 3, variables play a leading role in the formulation of mathematical laws. The foregoing remarks do not imply, however, that it would be impossible in principle to formulate such statements without the use of variables. But in practice it would scarcely be feasible to do without them, since even comparatively simple sentences would assume a complicated and obscure form. As an illustration, let us consider the following theorem of arithmetic:

$$\textit{for any numbers } x \textit{ and } y,\ \ x^3 - y^3 = (x - y) \cdot (x^2 + xy + y^2).$$

Without the use of variables, this theorem would look as follows:

> *the difference of the third powers of any two numbers is equal to the product of the difference of these numbers and a sum of three terms, the first of which is the square of the first number, the second the product of the two numbers, and the third the square of the second number.*

Variables take on an even greater importance, from the standpoint of the economy of thought, in the construction of mathematical proofs. This will be readily observed by a reader who attempts to eliminate the variables in any of the proofs which will be met in the course of our future considerations. And we should be aware that these proofs are much simpler than the average arguments which are found in the diverse fields of higher mathematics; attempts at carrying such arguments through without the help of variables would meet with the greatest of difficulties. It may be added, that the introduction of variables has allowed the development of so fertile a method for the solution of mathematical problems as the method of equations. Without exaggeration it can be said, that *the invention of variables constitutes a turning point in the history of mathematics; with these symbols man acquired a tool that prepared the way for the tremendous development of mathematical science and for the solidification of its logical foundations.*[2]

[2]Variables were used already in ancient times by Greek mathematicians and logicians,—though only in special circumstances and in rare cases. At the beginning of the 17th century, mainly under the influence of the work of the French mathematician F. VIETA (1540–1603), people began to work systematically with variables and to employ them consistently in mathematical considerations. Only at the end of the 19th century, however, when the notion of a quantifier had become firmly established, was the role of variables in scientific language and especially in the formulation of mathematical theorems fully recognized; this was largely the merit of the outstanding American logician and philosopher C. S. PEIRCE (1839–1914).

Exercises

1. Which of the following expressions are sentential functions, and which are designatory functions?

(a) *x is divisible by* 3,
(b) *the sum of the numbers x and* 2,
(c) $y^2 - z^2$,
(d) $y^2 = z^2$,
(e) $x + 2 < y + 3$,
(f) $(x + 3) - (y + 5)$,
(g) *the mother of x and z*,
(h) *x is the mother of z.*

2. Give examples of sentential and of designatory functions from the field of geometry.

3. The sentential functions which are encountered in arithmetic and which contain only one variable (which, however, may occur at several different places in the given sentential function) can be divided into three categories: (i) functions satisfied by every number; (ii) functions not satisfied by any number; (iii) functions satisfied by some numbers, and not satisfied by others.

To which of these categories do the following sentential functions belong?

(a) $x + 2 = 5 + x$,
(b) $x^2 = 49$,
(c) $(y + 2) \cdot (y - 2) < y^2$,
(d) $y + 24 > 36$,
(e) $z = 0$ *or* $z < 0$ *or* $z > 0$,
(f) $z + 24 > z + 36$.

4. Give examples of universal, of absolutely existential, and of universal-existential theorems from the fields of arithmetic and geometry.

5. By writing quantifiers containing the variables "x" and "y" in front of the sentential function:

$$x > y$$

it is possible to obtain various sentences from it, for instance:

for any numbers x and y, $x > y$;
for any number x, there exists a number y such that $x > y$;
there is a number y such that, for any number x, $x > y$.

Formulate them all (there are six altogether) and determine which of them are true.

6. Do the same as in exercise 5 for the following sentential functions:

$$x + y^2 > 1$$

and

x is the father of y

(in the latter, the variables "x" and "y" are assumed to stand for names of human beings).

7. State a sentence of everyday language that has the same meaning as:

for every x, if x is a dog, then x has a good sense of smell

and that contains no quantifiers or variables.

8. Replace the sentence:

some snakes are poisonous

by one which has the same meaning but is formulated with the help of quantifiers and variables.

9. Distinguish in the following expressions between free and bound variables:

(a) *x is divisible by y;*
(b) *for any x, $x - y = x + (-y)$;*
(c) *if $x < z$, then there is a number y such that $x < y$ and $y < z$;*
(d) *for any number y, if $y > 0$, then there is a number z such that $x = y \cdot z$;*
(e) *if $x = y^2$ and $y > 0$, then, for any number z, $x > -z^2$;*
(f) *if there exists a number y such that $x > y^2$, then, for any number z, $x > -z^2$.*

Reformulate the above expressions by replacing the quantifiers by the symbols which were introduced in Section 4.

***10.** If, in the sentential function in (e) of the preceding exercise, we replace the variable "z" in both places by "y", we obtain an expression in which "y" occurs in some places as a free and in others as a bound variable; in what places and why?†

(In view of certain difficulties in operating with expressions in which the same variable occurs both bound and free, some logicians prefer to avoid such expressions altogether and not to treat them as sentential functions.)

***11.** Try to state quite generally under which conditions a variable occurs at a certain place of a given sentential function as a free variable, and under which conditions, as a bound variable.

† In view of such situations as in this exercise, it is preferable to speak of free and of bound occurrences of variables, and not simply of free and of bound variables in a sentential function. This additional precision, however, sometimes is awkward to put into practice.

12. Determine which numbers satisfy the sentential function:

there is a number y *such that* $x = y^2$,

and which satisfy:

there is a number y *such that* $x \cdot y = 1$.

***13.** In addition to the symbols for quantifiers introduced in Section 4, we shall introduce in Chapter II the symbols "$\sim$", "$\wedge$", "$\vee$", "$\rightarrow$", and "$\leftrightarrow$" to replace the respective expressions of ordinary language "*not*" ("*it is not the case that*"), "*and*", "*or*", "*if ..., then ...*", and "*if, and only if*". Translate the following formulas into ordinary language:

(a) $\forall_x[(x = y \vee y < x) \leftrightarrow \sim (x < y)]$,
(b) $\exists_x[(0 < x \wedge x < y) \wedge \sim (x + 1 = y)]$,
(c) $\forall_{x,y}[x + y = 4 \rightarrow \exists_z(x < z \wedge z < y)]$.

Express the following sentential functions in logical symbols ††:

(d) *For any number* x, *there are numbers* y *and* z *such that* $y < 2$ *and* $x < z$.
(e) *For every* x, $x^2 + 6 = 5y$ *if, and only if,* $x = 2$ *or* $x = 3$.
(f) *There are* x *and* y *such that:* $x < y$ *and it is not the case that, for every* z, $x + z < y + z$.

For each of the sentential functions in (a)–(f) point out which variables occur free and which occur bound. If a certain variable occurs free, give, if possible, examples of numbers which satisfy this sentential function and of numbers which do not satisfy it. If a sentential function is a sentence, determine whether it is true or false.

††Some of the following expressions may be sentences rather than sentential functions. However, it is appropriate to regard a sentence as a special case of a sentential function, namely the case where the number of free variables reduces to zero. (In the same way, a constant like "*two*" can be regarded as a special case of a designatory function.)—These remarks may be contrasted with those of Section 2, according to which a sentential function is sometimes referred to as a sentence. This latter usage has a different sense: In such cases one tolerates a mild abuse of language for the sake of expediency.

II

On the Sentential Calculus

6 Logical constants; the old logic and the new logic

The constants with which we have to deal in any scientific theory may be divided into two large groups. The first group consists of terms which are specific for a given theory. In the case of arithmetic, for instance, they are the terms denoting either individual numbers or whole classes of numbers, relations among numbers, operations on numbers, etc.; the constants which we used in Section 1 as examples belong here, among others. On the other hand, in most statements of arithmetic there are also terms of a much more general character—terms which are encountered constantly in considerations of everyday life and in every possible field of science, and which are indispensable for conveying human thoughts and for carrying out arguments in any field whatsoever; such words as "*not*", "*and*", "*or*", "*is*", "*every*", "*some*", and many others belong here. There is a special discipline, called LOGIC, which is considered to be the basis for all other sciences, and where one aims to establish the precise meaning of such terms and to determine the most general laws which govern them.

Logic evolved into an independent science long ago, earlier even than arithmetic and geometry. And yet it was relatively recently, in the latter part of the nineteenth century—after a long period of almost complete stagnation—that this discipline began an intensive development, in the course of which it underwent a complete transformation and acquired a character similar to that of mathematical disciplines; in this new form it is known as MATHEMATICAL or SYMBOLIC LOGIC, and it has also been called LOGISTIC. The new logic surpasses the old in many respects,—not only because of the solidity of its foundations and the perfection of the methods which are employed in its development, but mainly on account of the wealth of concepts which have been investigated and the wealth of laws which have been discovered. Fundamentally, the old traditional logic forms only a fragment of the new, moreover a fragment which is entirely insignificant from the standpoint of requirements of other sciences, and of mathematics in particular. In view of these circumstances, and of the aim which we have here, there will be but very little opportunity in this book for drawing material from traditional logic for our considerations.[1]

7 Sentential calculus; the negation of a sentence, the conjunction and the disjunction of sentences

Among logical terms there is a small distinguished group, consisting of such words as "*not*", "*and*", "*or*", "*if..., then...*". All these words are well known to us from everyday language, and serve to construct compound sentences from simpler ones. In grammar they are classified as the so-called sentential conjunctions. We see here one reason, why the presence of these terms should not be regarded as a specific feature of any particular science. To establish the precise meaning and the usage of these terms, which are also known as SENTENTIAL CONNECTIVES, is the task of that part of logic which is the most elementary and the most fundamental; this part is called SENTENTIAL CALCULUS, or sometimes, PROPOSITIONAL CALCULUS, or PROPOSITIONAL or SENTENTIAL LOGIC.[2 †]

We shall now discuss the meaning of the most important terms of sentential calculus.

* * *

[1]Logic was created by ARISTOTLE, the illustrious Greek thinker of the 4th century B.C. (384–322); his logical writings are collected in the work *Organon*. As the creator of mathematical logic we have to look upon the great German philosopher and mathematician of the 17th century G. W. LEIBNIZ (1646–1716). However, LEIBNIZ's works on logic failed to have much influence upon the subsequent development of logical investigations; there was even a period during which they sank into oblivion. A continuous development of mathematical logic began only near the middle of the 19th century, at about the time when the logical system of the English mathematician G. BOOLE was published (1815–1864; principal work: *An Investigation of the Laws of Thought*, London 1854). Later the new logic found its full expression in the epochal work of the great English logicians A. N. WHITEHEAD (1861–1947) and B. RUSSELL (1872–1970): *Principia Mathematica* (Cambridge 1910–1913).

[2]Historically, the first systematic exposition of sentential calculus is contained in the work *Begriffsschrift* (Halle 1879) of the German logician G. FREGE (1848–1925) who, without doubt, was the greatest logician of the 19th century. The eminent Polish logician and historian of logic J. ŁUKASIEWICZ (1878–1956) succeeded in giving the sentential calculus a particularly simple and precise form, and prompted extensive investigations concerning this calculus.

[†]For a discussion of the two options, "*calculus*" and "*logic*", we refer the reader to footnote on p. 59.

With the help of the word "*not*" one forms the NEGATION of any sentence; two sentences, one of which is the negation of the other, are called CONTRADICTORY SENTENCES. In sentential calculus, the word "*not*" is put in front of the whole sentence, while in everyday language it is customary to place it with the verb; or should it be desirable to have it at the beginning of the sentence, it has to be replaced by the phrase "*it is not the case that*". For example, the negation of the sentence:

1 is a positive integer

reads as follows:

1 is not a positive integer,

or else:

it is not the case that 1 is a positive integer.

Whenever we utter the negation of a sentence, we intend to express the idea that the sentence is false; if the sentence is actually false, its negation is true, while otherwise its negation is false.

* * *

When two (or more) sentences are joined by the word "*and*", the result is their so-called CONJUNCTION, or, LOGICAL PRODUCT; the sentences which are joined in this manner are called the MEMBERS OF THE CONJUNCTION or the FACTORS OF THE LOGICAL PRODUCT. If, for instance, the sentences:

2 is a positive integer

and

$$2 < 3$$

are joined in this way, we obtain the conjunction:

2 is a positive integer and $2 < 3$.

When we assert the conjunction of two sentences, this is tantamount to stating that both sentences of which the conjunction is formed are true. If this is actually the case, then the conjunction is true, but if at least one of its members is false, then the whole conjunction is false.

* * *

When joining sentences by means of the word "*or*", one obtains the DISJUNCTION of those sentences, which is also called the LOGICAL SUM; the sentences forming the disjunction are called the MEMBERS OF THE DISJUNCTION or the SUMMANDS OF THE LOGICAL SUM. Now, in everyday language, the word "*or*" has at least two different meanings. Taken in the so-called NON-EXCLUSIVE MEANING, the disjunction of two sentences expresses only that at least one of these sentences is true, without saying anything as to whether or not both sentences may be true; taken in another meaning, known as the EXCLUSIVE one, the disjunction of two sentences asserts that one of the sentences is true but that the other is false. To illustrate, let us suppose we see the following notice put up in a bookstore:

Customers who are teachers or college students are entitled to a special reduction.

Here the word "*or*" is undoubtedly used in the first meaning, since one would not refuse the reduction to a teacher who is at the same time a college student. On the other hand, if a child has asked to be taken on a hike in the morning and to a theater in the afternoon, and we reply:

no, we shall go on a hike or we shall go to the theater,

then our usage of the word "*or*" is obviously of the second kind, since we intend to comply with only one of the two requests. In logic and in mathematics the word "*or*" is used always in the first, non-exclusive meaning; the disjunction of two sentences is considered true if at least one of its members is true, and otherwise false. For instance, we may assert:

every number is positive or less than 3,

although there are numbers which are both positive and less than 3. In order to avoid misunderstandings it would be expedient, in everyday as well as in scientific language, to use the word "*or*" by itself only in the first meaning, and to replace it by the compound expression "*either... or...*" whenever the second meaning is intended.

* * *

*Even if we confine ourselves to those cases in which the word "*or*" occurs in its first meaning, we find quite noticeable differences between its usage in everyday language and that in logic. In common language two sentences are joined by the word "*or*" only when they are in some way connected in form and content. (The same applies, though perhaps to a lesser degree, to the usage of the word "*and*".) It is not altogether clear what kinds of connections would be appropriate here, and any attempt at their detailed analysis and description would lead to considerable difficulties. As we shall see, such connections are disregarded in contemporary logic, where consequently one has to allow some strange examples; and indeed, anybody unfamiliar with its language would presumably be little inclined to consider a phrase such as:

$$2 \cdot 2 = 5 \text{ or } \textit{New York is a large city}$$

as a meaningful expression, and even less so to accept it as a true sentence. Moreover, the usage of the word "*or*" in everyday speech is influenced by certain psychological factors. Usually we state a disjunction of two sentences only if we believe that one of them is true but wonder which one. For example, if we look upon a lawn in normal light, it will not enter our mind to say that the lawn is green or blue, since we are able to affirm something simpler, and at the same time stronger, namely that the lawn is green. Sometimes we even take the utterance of a disjunction as an admission by the speaker that he or she does not know which member of the disjunction is true, and which is false. And if we later arrive at the conviction that the speaker knew at the time that one—and, specifically, which—of the members was false, we are inclined to look upon the whole disjunction as a false sentence, even though the other member should be undoubtedly true. For instance, let us imagine that a friend of ours, upon being asked when he will leave town, answers that he is going to do so today, tomorrow, or the day after. Should we later ascertain that, at that time, he had already decided to leave the same day, we shall probably get the impression that we were deliberately misled and that he told us a lie.

When creators of contemporary logic were introducing the word "*or*" into their considerations, they desired, perhaps subconsciously, to simplify its meaning; in particular, they endeavored to render this meaning clearer and independent of all psychological factors, especially of the presence or absence of knowledge. Consequently, they decided to extend the usage of the word "*or*", and to consider the disjunction of any two sentences as a meaningful whole, even when no connection between their contents or forms should exist; and they also decided to make the truth of a disjunction—like that of a negation or a conjunction—dependent only on the truth of its members. Therefore, a person using the word "*or*" according to contemporary logic will consider the expression given above:

$$2 \cdot 2 = 5 \;\; \textit{or New York is a large city}$$

as a meaningful and in fact a true sentence, since its second part is surely true. Similarly, if we assume that our friend, who was asked about the date of his departure, used the word "*or*" in its strict logical meaning, we shall be compelled to regard his answer as true, independently of our opinion as to his intentions.*

8 Implications or conditional sentences; implications in the material meaning

If we combine two sentences by the words "*if..., then...*", we obtain a compound sentence which is designated as an IMPLICATION or a CONDITIONAL SENTENCE. The subordinate clause to which the word "*if*" is prefixed is called ANTECEDENT, and the principal clause introduced by the word "*then*" is called CONSEQUENT. By asserting an implication one claims that the antecedent cannot be true when the consequent is false. An implication is thus true in each of the following three cases: (i) both the antecedent and the consequent are true, (ii) the antecedent is false and the consequent is true, (iii) both the antecedent and the consequent are false; and only in the fourth possible case, when the antecedent is true and the consequent is false, is the whole implication false. It follows that he or she who accepts an implication as true, and at the same time accepts its antecedent as true, cannot but accept its consequent; and whoever accepts an implication as true and rejects its consequent as false, must also reject its antecedent.

* * *

*As in the case of disjunctions, considerable differences between the usage of implications in logic and in everyday language manifest themselves. Again, in ordinary language, we tend to join two sentences by the words "*if..., then...*" only when there is some connection between their forms and contents. Connections which would be suitable for this purpose are hard to characterize in a general way, and only at times is their nature relatively clear. In particular, we often expect a given implication to be such that the consequent follows necessarily from the antecedent, that is to say, that if the antecedent is true under certain circumstances, then the consequent must also be true under the same circumstances (and that possibly we can even deduce the consequent from the antecedent on the basis of some general laws which we might not always be able to quote explicitly). And here, as before, an additional psychological factor manifests itself; usually we formulate and assert an implication only if we have no exact knowledge as to whether or not the antecedent and the consequent are true. Otherwise the use of an implication seems unnatural, and doubts may arise with regard to its meaningfulness and its truth.

The following example may serve as an illustration. Let us consider the law of physics:

every metal is malleable,

and let us put it in the form of an implication containing variables:

if x is a metal, then x is malleable.

If we believe in the truth of this universal law, we must also believe in the truth of all of its particular cases, that is, of all implications obtainable by replacing "x" by names of arbitrary materials such as iron, clay, or wood. And, indeed, it turns out that all sentences which are obtained in this way satisfy the conditions given above for a true implication; it never happens that the antecedent is true while the consequent is false. In every implication of this kind, moreover, there is a close connection between the antecedent and the consequent, which depends on their having the same subject. We notice, finally, that by assuming the antecedent of any of such implications to be true, we can deduce the consequent (for instance, from "*iron is a metal*" we can deduce "*iron is malleable*") by referring to the general law that every metal is malleable.

Nevertheless, some of the sentences which we have been led to consider seem artificial and doubtful from the point of view of common language. One would not hesitate to use the universal law given above, nor to use any of the particular cases which are obtained by replacing "x" by the name of a material, as long as one does not know whether it is a metal or whether it is malleable. But if we replace "x" by "*iron*", we are confronted with a case in which the antecedent and the consequent are certainly true; then the implication seems indeed artificial, and we prefer to use, instead, an expression such as:

since iron is a metal, it is malleable.

Similarly, if for "x" we substitute "*clay*", we obtain an implication with a false antecedent and a true consequent, and we are inclined to replace it by the expression:

although clay is not a metal, it is malleable.

And finally, replacing "x" by "*wood*" results in an implication with a false antecedent and a false consequent; in this case, if we want to retain the structure of an implication, we should alter the grammatical form of the verbs:

if wood were a metal, then it would be malleable.

* * *

The logicians, with due regard for the needs of scientific languages, adopted the same procedure with respect to the phrase "*if..., then ...*" as they had done in the case of the word "*or*". They decided to simplify and to clarify the meaning of this phrase, and to free it from extraneous factors. For this purpose they extended the usage of this phrase, by considering an implication as a meaningful sentence even if no connection whatsoever exists between its two members, and they made the truth or falsity of an implication dependent exclusively upon the truth or falsity of the antecedent and of the consequent. To characterize this situation briefly, we say that contemporary logic uses IMPLICATIONS IN THE MATERIAL MEANING, or simply, MATERIAL IMPLICATIONS; this is opposed to the usage of the concept of IMPLICATION IN THE FORMAL MEANING, or, of FORMAL IMPLICATION, in which case the presence of a certain formal connection between the antecedent and the consequent is an indispensable condition for the meaningfulness and the truth of the implication. The concept of formal implication perhaps has not been made completely clear, but at any rate, it is narrower than that of material implication; every meaningful and true formal implication is at the same time a meaningful and true material implication, but not vice versa.

In order to illustrate the foregoing remarks, let us consider the following four sentences:

if $2 \cdot 2 = 4$, *then New York is a large city;*

if $2 \cdot 2 = 5$, *then New York is a large city;*

if $2 \cdot 2 = 4$, *then New York is a small city;*

if $2 \cdot 2 = 5$, *then New York is a small city.*

In everyday language, these sentences would hardly be regarded as meaningful, much less as true. From the point of view of mathematical logic, on the other hand, they are all meaningful, the third sentence being false, while the remaining three are true. Of course, we do not thereby suggest that sentences like these are particularly relevant from any viewpoint whatever, or that we want to apply them as premisses in our arguments.

* * *

It would be a mistake to think that the difference between everyday language and the language of logic, which has been brought to light here, is fixed and definite, and that the customs outlined above, regarding the usage of the words "*if..., then...*" in common language, do not admit any exceptions. Actually, the usage of these words fluctuates to some extent, and if we look around, we can find cases which trespass the sense of formal implication. Let us imagine that a friend of ours is confronted with a very diflficult problem and that we do not believe that he will ever solve it. We can then express our disbelief in a jocular form by saying:

if you solve this problem, I shall eat my hat.

The sense of this utterance is quite clear. We have here an implication whose consequent is undoubtedly false; and since we affirm the truth of the whole implication, we therefore, at the same time, reject the truth of the antecedent; that is to say, we express our conviction that our friend will fail to solve the problem in which he is interested. But it is also quite clear that the antecedent and the consequent of our implication are not connected in any way, so that we have a typical case of a material and not of a formal implication.

* * *

The divergency between the usage of the phrase "*if..., then...*" in ordinary language and its usage in mathematical logic has been at the root of lengthy and even passionate discussions,—in which, by the way, professional logicians took only a minor part.[3] (It is perhaps surprising, that considerably less attention was paid to the analogous divergency in the

[3]It is interesting to notice thnt the beginning of these discussions dates back to antiquity. It was the Greek philosopher PHILO OF MEGARA (in the 4th century B.C.)

case of the word "*or*".) It has been objected that logicians, on account of their adoption of the concept of material implication, arrived at paradoxes and even at plain nonsense. This has resulted in an outcry for a reform of logic, and in particular, for bringing about a far-reaching rapprochement between logic and ordinary language with regard to the use of implication.

It would be hard to grant that these criticisms are well founded. There is no phrase in ordinary language which has a precisely determined meaning. It would scarcely be possible to find two people who would use every word with exactly the same meaning, and even in the language of a single person the meaning of a given word may vary from one period of the person's life to another. Moreover, the meaning of words of everyday language is usually very complicated; it depends not only on the external form of the word, but also on the circumstances in which it is uttered, and sometimes even on subjective psychological factors. If a scientist wants to transfer a concept from everyday life into a science and to establish general laws concerning this concept, he (or she) must always make its content clearer, more precise, and simpler, and free it from inessential attributes; it does not matter here whether he is a logician who is concerned with the phrase "*if..., then...*", or, for instance, a physicist wanting to establish the exact meaning of the word "*metal*". In whatever way the scientist realizes his task, the resulting usage of the term will deviate more or less from the practice of everyday language. If, however, he states explicitly in what sense he decides to use the term, and if afterwards he acts always in accordance with this decision, then nobody will be in a position to object, or to argue that his procedure leads to nonsensical results.

Nevertheless, in connection with the discussions that have taken place, some logicians attempted to reform the theory of implication. Generally they do not deny a place in logic to material implication, but they are anxious to find also a place for another concept of implication, for instance, of such a kind that one must be able to deduce the consequent from the antecedent in order to have a true implication; they even desire, so it seems, to place the new concept in the foreground. These attempts are of a relatively recent date, and it is too early to pass a final judgment as to their value.[4] But today *it appears certain that the theory of material implication will surpass all other theories in simplicity; and, in any case,*

who presumably was the first in the history of logic to advocate the usage of material implication; this was in opposition to the views of his master, DIODORUS CRONUS, who proposed the use of implication in a narrower sense, rather related to what is called here the formal meaning. Somewhat later (in the 3d century B.C.)—and probably under the influence of PHILO—various possible conceptions of implication were considered by the Greek philosophers and logicians of the Stoic school (whose writings included the first discussions of sentential calculus).

[4] The first attempt to develop a systematic theory of implication of this kind was made by the American philospher and logician C. I. LEWIS (1883–1964).

*it must not be forgotten that logic, which has been founded upon this simple concept, turned out to be a satisfactory basis for the most complicated and most subtle of mathematical reasonings.**

9 The use of implications in mathematics

The phrase "*if..., then...*" is among those expressions of logic which are the most widely used in other sciences and, especially, in mathematics. Mathematical theorems, particularly those of universal character, tend to have the form of implications; the antecedent is called in mathematics the HYPOTHESIS, and the consequent is called the CONCLUSION.

As a simple example of a theorem of arithmetic having the form of an implication, we may quote the following sentence†:

if x is a positive number, then $2x$ is a positive number

in which "*x is a positive number*" is the hypothesis, while "*$2x$ is a positive number*" is the conclusion.

Apart from this classical form of mathematical theorems (so to speak), there are various alternative formulations, in which the hypothesis and the conclusion are connected through some other means than by the phrase "*if..., then...*". The theorem just mentioned, for instance, can be paraphrased in any of the following ways:

from: x is a positive number, it follows: $2x$ is a positive number;

the hypothesis: x is a positive number, implies (or, *has as a consequence) the conclusion: $2x$ is a positive number;*

the condition: x is a positive number, is sufficient for $2x$ to be a positive number;

for $2x$ to be a positive number it is sufficient that x be a positive number;

the condition: $2x$ is a positive number, is necessary for x to be a positive number;

for x to be a positive number it is necessary that $2x$ be a positive number.

†Strictly speaking, the expression which follows is a sentential function rather than a sentence. We recall that the term "*theorem*" was first introduced for proved sentences, but it is convenient and natural to extend the scope of this term also to proved sentential functions. (Note that, if we were to adjoin the universal quantifier to this sentential function, or to any other, then we would not have an implication in the strict sense.)

Therefore, instead of asserting a conditional sentence, it is usually just as appropriate to say that the hypothesis IMPLIES the conclusion or HAS it AS A CONSEQUENCE, or that it is a SUFFICIENT CONDITION for the conclusion; or one can express the same idea by saying that the conclusion FOLLOWS from the hypothesis, or that it is a NECESSARY CONDITION for the latter. A logician may raise certain objections against some of the formulations given above, but they are in general use in mathematics.

* * *

*The objections which might be raised here concern those of the above statements in which any of the words "*hypothesis*", "*conclusion*", "*consequence*", "*follows*", "*implies*" occur.

In order to understand the essential points in these objections, we observe first that, strictly speaking, those statements differ in content from the one originally given. While in the "classical form" of the theorem we talk about numbers, properties of numbers, operations upon numbers, and so on—in short, about objects with which mathematics is concerned—, in the formulations now under discussion we talk about hypotheses, conclusions, conditions, that is, about sentences or sentential functions occurring in mathematics. It might be noted on this occasion that, in general, people do not distinguish clearly enough the terms which denote the objects dealt with in a given science, from those terms which denote various kinds of expressions occurring within it. This can be observed, in particular, in the domain of mathematics, especially on the elementary level. It appears that only relatively few individuals are aware of the fact that such terms as "*equation*", "*inequality*", "*polynomial*", or "*algebraic fraction*", which are met at every turn in textbooks of elementary algebra, do not belong to the domain of mathematics or of logic (in the strict sense of the words), since they do not denote objects which are considered in these domains; equations and inequalities are certain special sentential functions, while polynomials and algebraic fractions—especially as they are treated in elementary textbooks—are particular instances of designatory functions (cf. Section 2). The confusion on this point is reinforced by the fact that terms of this kind are frequently used in the statements of mathematical theorems. This has become a very common usage, and perhaps it is not worth our while to put up a stand against it, since it does not present any particular danger; but it might be useful to recognize that, for many theorems formulated with the help of such terms, there are other formulations in which those terms do not occur at all, and which are therefore logically more satisfactory. For instance, the theorem:

the equation: $x^2 + ax + b = 0$ *has at most two roots*

can be expressed in a more appropriate manner as follows:

there are at most two numbers x *such that* $x^2 + ax + b = 0$.

* * *

Let us return to the controversial formulations of an implication in order to emphasize one further point, which is particularly important. In these formulations we assert that one sentence, namely the antecedent of the implication, has another—the consequent of the implication—as a consequence, that is, that the second follows from the first. Ordinarily when we express ourselves in this way, we have in mind that the assumption of the first sentence being true leads us, so to speak, necessarily to the same assumption concerning the second sentence (and that perhaps we are even able to derive the second sentence from the first). As we know already from Section 8, however, in contemporary logic the meaningfulness of an implication does not depend on whether its consequent has any such connection with its antecedent; and it is interesting to look again at the examples which were given there. Anyone who was shocked by the fact that the expression:

if $2 \cdot 2 = 4$, *then New York is a large city*

is considered in logic as a meaningful and even true sentence, will find it still harder to reconcile himself with a transformed phrase such as:

the hypothesis that $2 \cdot 2 = 4$ *has as a consequence that New York is a large city.*

We see that the alternate ways of formulating or transforming a conditional sentence lead to paradoxical-sounding utterances, and so they intensify the discrepancies between common language and mathematical logic. It is therefore not surprising that such reformulations repeatedly brought about various misunderstandings, and have been one of the causes of those passionate and frequently sterile discussions which we mentioned above.

* * *

From the purely logical point of view, we can obviously avoid all objections which are raised here by stating explicitly once and for all, that we shall disregard the literal meaning of the formulations in question, and that we shall attribute to them exactly the same contents as to ordinary conditional sentences. But this would be inconvenient in another respect; for there are situations—though not in logic itself, but in a field closely related to it, namely, in the methodology of deductive sciences (cf. Chapter VI)—where we talk about sentences and about the relation of consequence among them, and where we use such terms as "*implies*" and "*follows*" in a meaning which is different, and which is more closely akin to the ordinary one. It would therefore be better to avoid such formulations altogether, especially since we have several others at our disposal which are not open to any of these objections.*

10 Equivalence of sentences

We shall now consider a still another sentential connective. It is one which is rather rarely met in everyday language, namely, the phrase "*if, and only if*". If any two sentences are joined by this phrase, the result is a compound sentence called an EQUIVALENCE. The two sentences which are connected in this way are referred to as the LEFT and the RIGHT SIDE OF THE EQUIVALENCE. By asserting the equivalence of two sentences, we exclude the possibility that one is true and the other false; an equivalence, therefore, is true if its left and right sides are either both true or both false, and otherwise the equivalence is false.

The sense of an equivalence can also be characterized in a different way. If, in a conditional sentence, we interchange the antecedent and the consequent, we obtain a new sentence which is called the CONVERSE OF THE GIVEN SENTENCE, or simply, the CONVERSE SENTENCE. Let us take, for instance, the following implication as the given sentence:

(I) *if x is a positive number, then $2x$ is a positive number;*

the converse of this sentence will then be:

(II) *if $2x$ is a positive number, then x is a positive number.*

As this example shows, the converse of a true sentence sometimes is true. In order to see that this is not a general rule, it suffices to replace "$2x$" by "x^2" in (I) and in (II); sentence (I) then remains true, while sentence (II) will become false. If, now, it happens that a given conditional sentence and its converse are both true, then their simultaneous truth can also be expressed by joining the antecedent and the consequent of either of the two sentences by the words "*if, and only if*". Accordingly, the above two implications—the original sentence in (I) and its converse in (II)—may be replaced by a single sentence:

x is a positive number if, and only if, $2x$ is a positive number

(in which the two sides of the equivalence may also be interchanged).

There are, incidentally, several other formulations which can be used to express the same idea, e.g.:

from: x is a positive number, it follows: $2x$ is a positive number, and conversely;

the conditions that x is a positive number and that $2x$ is a positive number are equivalent with (or, *equivalent to) each other;*

the condition that x is a positive number is both necessary and sufficient for $2x$ to be a positive number;

for x to be a positive number it is necessary and sufficient that $2x$ be a positive number.

Instead of joining two sentences by the phrase "*if, and only if*", one can therefore also say that the RELATION OF CONSEQUENCE HOLDS between these two sentences IN BOTH DIRECTIONS, or, that the two sentences are EQUIVALENT, or finally, that each of the two sentences represents a NECESSARY AND SUFFICIENT CONDITION for the other.

11 The formulation of definitions and its rules

The phrase "*if, and only if*" is used very frequently in presenting DEFINITIONS; these should be regarded as conventions stipulating what meanings are to be attributed to expressions which previously did not appear in a certain discipline, and which may therefore not be comprehensible. Imagine, for instance, that in arithmetic the symbol "$\leq$" has not as yet been employed, but that we want to introduce it now into the considerations (looking upon it, in the usual way, as an abbreviation of the expression "*is less than or equal to*" or of "*is not greater than*"). For this purpose it is necessary to define this symbol first, that is, to explain exactly its meaning in terms of expressions which are already known and whose meanings do not give rise to any doubt. To achieve this, we introduce the following definition,—assuming that "$>$" belongs to the symbols already known:

we say that $x \leq y$ *if, and only if, it is not the case that* $x > y$.

The definition which we have just formulated stipulates the equivalence of the two sentential functions:

$$x \leq y$$

and

it is not the case that $x > y$;

this equivalence has the important feature of allowing the transformation of the formula "$x \leq y$" into an equivalent expression, which no longer contains the symbol "$\leq$" but is formulated by using only those terms which have already been comprehensible to us. The same holds for any formula which is obtained from "$x \leq y$" by replacing "x" and "y" by arbitrary symbols or expressions designating numbers. The formula:

$$3 + 2 \leq 5,$$

for instance, is equivalent with the sentence:

it is not the case that $3 + 2 > 5$;

since the latter is a true assertion, so is the former. Similarly, the formula:

$$4 \leq 2 + 1$$

is equivalent to the following:

$$\textit{it is not the case that } 4 > 2 + 1,$$

both being false assertions. The foregoing remark about allowing a transformation applies also to more complicated sentences (and sentential functions); for instance, by transforming the sentence:

$$\textit{if: } x \leq y \textit{ and } y \leq z, \textit{ then: } x \leq z,$$

we obtain:

$$\textit{if: it is not the case that } x > y \textit{ and it is not the case that } y > z, \textit{ then: it is not the case that } x > z.$$

In short, by virtue of the definition given above, we are in a position to transform any simple or compound sentence containing the symbol "$\leq$" into an equivalent one which no longer contains it[†]; to say in other words, we can translate it into a language in which the symbol "$\leq$" does not occur. And the role of definitions within the mathematical disciplines depends on just this fact.

* * *

For a definition to fulfill its assigned task well, certain precautionary measures have to be observed in its formulation. Accordingly, special rules have been established, the so-called RULES OF DEFINITION, which specify how to construct definitions correctly. We shall not give here a precise statement of these rules, but shall only remark that, on their basis, every definition may assume the form of an equivalence; the left side of this equivalence, the DEFINIENDUM, should be a short, grammatically simple sentential function containing the constant to be defined; the right side, the DEFINIENS, may be a sentential function of an arbitrary structure, but containing only those constants whose meaning is understood; that is, the meaning of each of those constants either is immediately obvious or has already been explained. In particular, the constant to be defined, or any expression that was previously defined with its help, must not occur in the definiens; otherwise the definition is incorrect, since it contains an error known as a VICIOUS CIRCLE IN THE DEFINITION (just as one speaks of a VICIOUS CIRCLE IN THE PROOF, when the argument which is intended to establish a certain sentence is based upon the sentence itself, or upon some other sentence previously deduced with its help). Finally, in order to emphasize the character of a definition as a convention, and to distinguish it from other statements which have the form of an equivalence, it is expedient to prefix to it words such as "*we say that*".—It is now easy to verify that

---

[†] Of course, the equivalence also allows one to proceed in the opposite direction: to transform certain expressions into shorter ones, by utilizing the symbol "$\leq$".

the above definition of the symbol "$\leq$" satisfies all of these conditions; it has the definiendum:

$$x \leq y,$$

while the definiens reads:

it is not the case that $x > y$.

It is also worth noticing that mathematicians, in stating definitions, prefer the words "*if*" or "*in case that*" to the phrase "*if, and only if*". To the definition of the symbol "$\leq$" they would, presumably, give the following form:

we say that $x \leq y$, *if it is not the case that* $x > y$.

It looks as if such a definition stated only that the definiendum follows from the definiens, without indicating that the relation of consequence also holds in the opposite direction, and as if it therefore failed to express the equivalence of definiendum and definiens. But what we actually have here, is a tacit convention to the effect that "*if*" or "*in case that*", if used to join definiendum and definiens, are to mean the same as the phrase "*if, and only if*" ordinarily does.—It may be added, that the form of an equivalence is not the only form in which definitions may be presented.

12 Laws of sentential calculus

Having completed our discussion of the most important sentential connectives, we are now in position to clarify the character of the laws of sentential calculus.

Let us consider the following sentence:

if 1 *is a positive number and* $1 < 2$, *then* 1 *is a positive number.*

This sentence is obviously true, it contains exclusively constants belonging to the domain of logic and arithmetic, and yet the idea of listing this sentence as a special theorem in a textbook of mathematics would not occur to anybody. If one reflects why this is so, one comes to the conclusion that this sentence is completely uninteresting from the standpoint of arithmetic; it fails to enrich our knowledge about numbers in any way, since its truth does not depend at all upon the content of the arithmetical terms occurring within it, but only upon the meaning of the words "*and*", "*if*", "*then*". In order to convince ourselves of this, let us replace in the sentence under consideration the two components:

1 *is a positive number*

and

$$1 < 2$$

by any other sentences from an arbitrary domain; we arrive at a series of sentences, each of which is true, like the original; for example:

if the given figure is a rhombus and the same figure is a rectangle, then the given figure is a rhombus;

if today is Sunday and the sun is shining, then today is Sunday.

In order to express the truth of such sentences in a more general form, we shall introduce the variables "p" and "q", stipulating that these symbols are not designations of numbers or of any other objects, but that they stand for whole sentences; variables of this kind are known as SENTENTIAL VARIABLES. In particular, in the sentence under consideration we shall replace the phrase:

1 *is a positive number*

by "p", and the formula:

$$1 < 2$$

by "q"; in this way we arrive at the sentential function:

if p *and* q, *then* p.

This sentential function has the property that one necessarily obtains true sentences when arbitrary sentences are substituted for "p" and "q". This observation may be put into the form of a universal statement:

for any p *and* q, *if* p *and* q, *then* p.

We have obtained here our first example of a law of sentential calculus, which will be referred to as the LAW OF SIMPLIFICATION FOR LOGICAL MULTIPLICATION. The sentence which we considered at first is simply a special instance of this universal law—just as, for example, the formula:

$$2 \cdot 3 = 3 \cdot 2$$

is a special instance of the universal arithmetical law:

for arbitrary numbers x *and* y, $x \cdot y = y \cdot x$.

In a similar way, other laws of sentential calculus can be found. We give here a few examples of such laws; in their formulation we omit the universal quantifier "*for any p, q, ...*"—in accordance with the usage mentioned in Section 3, which has become almost a rule throughout sentential calculus.

If p, *then* p.

If p, *then* q *or* p.

If p *implies* q *and* q *implies* p, *then* p *if, and only if,* q.

If p *implies* q *and* q *implies* r, *then* p *implies* r.

The first of these four statements is known as the LAW OF IDENTITY, the second is called the LAW OF LOGICAL ADDITION, and the fourth, the LAW OF THE HYPOTHETICAL SYLLOGISM.† ††

Just as arithmetical theorems of universal character express properties of arbitrary numbers, so laws of sentential calculus assert, so one may say, properties of arbitrary sentences. Variables of only one kind occur in these laws, namely those variables which stand for arbitrary sentences; this circumstance is characteristic for sentential calculus, and is decisive for its great generality and for the wide range of its applications.

13 The symbolism of sentential calculus; compound sentential functions and truth tables

There is a simple and general method, called the METHOD OF TRUTH TABLES or OF MATRICES, which enables us to determine whether any given sentence from the domain of sentential calculus is true, and therefore, whether it can be counted as a law of this calculus.[5]

For describing this method it is convenient to introduce a special symbolism. We shall replace the sentential connectives:

not; and; or; if..., then...; if, and only if,

by the symbols:

$\sim$; $\wedge$; $\vee$; $\rightarrow$; $\leftrightarrow$,

respectively. The first of these symbols is to be placed in front of the expression whose negation one wants to obtain; the remaining symbols are always put between two expressions ("$\rightarrow$" therefore stands in the place of the word "*then*", while the word "*if*" is simply omitted). From one or two simpler expressions we are led, in this way, to a more complicated expression; and if we want to use the latter for the construction of another expression which is still more complicated, we enclose it in parentheses.

---

†The reader should note the uneven spaces in the above "verbal formulas". Here the large spaces play a role which is similar to that of parentheses in equations.

††The law of logical addition is also called more briefly the LAW OF ADDITION. Still another name for this law is: LAW OF SIMPLIFICATION FOR LOGICAL ADDITION. This latter name originates in a certain analogy between this law and the law of simplification for logical multiplication, given above. This analogy arises in more advanced considerations (see e.g. *Supplementary exercises* of Chapter VI), but for now, the name "*law of simplification*" appears to be a misnomer: If we assume that p and conclude that q *or* p, we seem to be complicating things, not simplifying.

[5]This method originates with PEIRCE (who was cited already at an earlier occasion; cf. footnote 2 on p. 12).

With the help of variables, parentheses, and the constants listed above (and sometimes also with the help of additional constants of a similar character, which will not be discussed here), we are able to write down all sentential functions and all sentences belonging to the domain of sentential calculus. Apart from various sentential variables, the simplest sentential functions are the following:

$$\sim p, \qquad p \wedge q, \qquad p \vee q, \qquad p \rightarrow q, \qquad p \leftrightarrow q$$

(and other similar expressions which differ from these only in the shape of the variables used). As an example of a compound sentential function which is more involved, let us note the expression:

$$(p \vee q) \rightarrow (p \wedge r),$$

which we read, translating symbols into common language:

if p or q, then p and r.

A still more complicated expression is the law of the hypothetical syllogism which was given before, and which now assumes the form:

$$[(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r).$$

* * *

We can easily convince ourselves that every sentential function which is constructed in such a way is a so-called TRUTH FUNCTION. This means the following: if a sentence is obtained from that function by substituting whole sentences for the variables, its truth or falsity depends exclusively on the truth or falsity of the sentences which have been substituted.† As to the simplest sentential functions: "$\sim p$", "$p \wedge q$", and so on, this follows immediately from the remarks made in Sections 7, 8, and 10 regarding the meaning which is attributed in logic to the words "*not*", "*and*", and so on. But the same applies, likewise, to longer compound functions. Let us consider, for instance, the function "$(p \vee q) \rightarrow (p \wedge r)$" mentioned above. A sentence which is obtained from it by substitution is an implication, and its truth therefore depends only on the truth or falsity of its antecedent and of its consequent—and the two possibilities, truth and falsity, are referred to jointly as TRUTH VALUES; next, the truth value of the antecedent, which is a disjunction obtained from "$p \vee q$", depends only on the truth values of the sentences which were substituted for "p" and "q", and similarly the truth

† A truth function is therefore analogous to what may be called a NUMERICAL FUNCTION, i.e., to a designatory function whose designation (a number) depends exclusively on the numerical constants which have been substituted for the variables.—However, the term "*truth function*" is not widely used, and in practice one would refer to a formula like "$p \rightarrow (p \wedge q)$" as a *compound sentenial function*, or as an *expression of sentential calculus.*

value of the consequent depends only on the truth values of the sentences substituted for "p" and "r". So, finally, the truth value of the whole sentence obtained from the sentential function under consideration depends exclusively on the truth values of the sentences which were substituted for "p", "q", and "r". ††

We now consider a sentence which is obtained by substitution from a given sentential function; in order to see in detail how its truth value depends on the truth values of the sentences which are substituted for the variables, one constructs what is called the TRUTH TABLE or the MATRIX for this function. We shall now describe such constructions, and we begin by giving the table for the function "$\sim p$":

p	$\sim p$
T	F
F	T

And here is the joint truth table for the other elementary sentential functions, that is, for "$p \wedge q$", "$p \vee q$", and so on:

p	q	$p \wedge q$	$p \vee q$	$p \rightarrow q$	$p \leftrightarrow q$
T	T	T	T	T	T
F	T	F	T	T	F
T	F	F	T	F	F
F	F	F	F	T	T

The meaning of these tables becomes clear at once if we note that the letters "T" and "F" are abbreviations of "*true sentence*" and "*false sentence*", respectively. To give an example, we find in the second table, in the second line below the headings "p", "q", and "$p \rightarrow q$", the letters "F", "T", and "T", respectively. We interpret this to mean, that a sentence which is obtained from the implication "$p \rightarrow q$" is true if we substitute any false sentence for "p" and any true sentence for "q"; this is obviously in full agreement with the discussion of Section 8.—The variables "p" and "q" occurring in the tables can, of course, be replaced by any two distinct variables.

With the help of the two above tables, called FUNDAMENTAL TRUTH TABLES, we can construct a DERIVATIVE TRUTH TABLE for any compound

††Let us elaborate on the analogy with numbers and numerical functions (cf. the preceding footnote). In the latter case, we start by substituting numerical constants for variables, and instead of talking about these constants, we can also refer to numbers and use the term "*numerical values*". From these values we obtain the numerical values of the successive combinations, and finally, the numerical value of the original function.—We see that the the truth values of the expression "$(p \vee q) \rightarrow (p \wedge r)$" are determined in much the same way as the numerical values of, say, "$x^2 + 2x + y$".

sentential function. The table for the function "$(p \vee q) \rightarrow (p \wedge r)$", for instance, looks as follows:

p	q	r	$p \vee q$	$p \wedge r$	$(p \vee q) \rightarrow (p \wedge r)$
T	T	T	T	T	T
F	T	T	T	F	F
T	F	T	T	T	T
F	F	T	F	F	T
T	T	F	T	F	F
F	T	F	T	F	F
T	F	F	T	F	F
F	F	F	F	F	T

In order to explain the construction of this table, let us concentrate, say, on its fifth horizontal line (below the headings). We substitute true sentences for "p" and "q" and a false sentence for "r". According to the second fundamental table, we then obtain from "$p \vee q$" a true sentence and from "$p \wedge r$" a false sentence. From the whole function "$(p \vee q) \rightarrow (p \wedge r)$" we obtain then an implication with a true antecedent and a false consequent; hence, again with the help of the second fundamental table (in which we think of "p" and "q" as being for the moment replaced by "$p \vee q$" and "$p \wedge r$"), we conclude that this implication is a false sentence.

The horizontal lines of a table, consisting of the symbols "T" and "F", are called ROWS of the table, and the vertical lines are called COLUMNS. Each row, or rather, that part of each row which is on the left of the vertical bar, represents a certain substitution of true or false sentences for the variables. When constructing the matrix for a given function, we take care to exhaust all possible ways in which a combination of the symbols "T" and "F" could be correlated with the variables; and, of course, we never write in a table two rows which do not differ either in the number or in the order of the symbols "T" and "F". One can then see very easily that the number of rows in a table depends in a simple way on the number of distinct variables occurring in the function; if a function contains 1, 2, 3, ... variables of different shape, its matrix consists of $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, ... rows. As for the number of columns, it is equal to the number of partial sentential functions of different form which are contained in the given function (where each of the variables and the whole function are also counted among its partial functions).

* * *

Once the construction of the truth tables has been understood, one may proceed to decide whether or not a given sentence of sentential calculus is true. As we know, in sentential calculus there is no external difference between sentences and sentential functions; the only difference lies in the way the expressions are interpreted and depends on the rule, that the expressions which are considered to be sentences are always completed mentally by the universal quantifier. Accordingly, in order to determine whether the given sentence is true, we may treat it for the time being as a sentential function, and we construct the truth table for it. If in the last column of this table the symbol "F" does not occur, then every sentence which is obtainable from the function in question by substitution will be true, and therefore our original universal sentence (obtained from the sentential function by mentally prefixing the universal quantifer) is also true. If, however, the last column contains even one symbol "F", our sentence is false.

For instance, we have seen that in the above matrix for the function "$(p \vee q) \rightarrow (p \wedge r)$", the symbol "$F$" occurs four times in the last column. If, therefore, we should consider this expression to be a sentence (that is, if we prefixed to it the words "*for any p, q, and r*"), then this sentence would be false. On the other hand, one can easily verify with the help of the method of truth tables that all the laws of sentential calculus which are stated in Section 12, that is: the laws of simplification, identity, and so on, are true sentences. The table for the law of simplification:

$$(p \wedge q) \rightarrow p,$$

in particular, is as follows:

p	q	$p \wedge q$	$(p \wedge q) \rightarrow p$
T	T	T	T
F	T	F	T
T	F	F	T
F	F	F	T

We give here a number of other important laws of sentential calculus, whose truth can be ascertained by proceeding in a similar way:

$$\sim (p \wedge \sim p), \qquad p \vee \sim p,$$

$$(p \wedge p) \leftrightarrow p, \qquad (p \vee p) \leftrightarrow p,$$

$$(p \wedge q) \leftrightarrow (q \wedge p), \qquad (p \vee q) \leftrightarrow (q \vee p),$$

$$[p \wedge (q \wedge r)] \leftrightarrow [(p \wedge q) \wedge r], \qquad [p \vee (q \vee r)] \leftrightarrow [(p \vee q) \vee r].$$

The two laws in the first line are called the LAW OF CONTRADICTION and the LAW OF EXCLUDED MIDDLE; we next have the two LAWS OF TAUTOLOGY (for logical multiplication and logical addition); we then have the two COMMUTATIVE LAWS, and finally the two ASSOCIATIVE LAWS. One can easily see how obscure the meaning of the last two laws would become if we try to express them in ordinary language. They illustrate very clearly the value of logical symbolism as a precise instrument for expressing more complicated thoughts.

* * *

*It turns out that by the method of matrices we can establish as laws certain sentences, whose truth seems to be far from obvious until they are tested by this method. Here are a few examples of this kind‡:

$$p \rightarrow (q \rightarrow p),$$
$$\sim p \rightarrow (p \rightarrow q),$$
$$(p \rightarrow q) \vee (q \rightarrow p).$$

That these laws are not immediately evident is due mainly to the fact that they depend upon the peculiarity of implication which is characteristic of modern logic, that is, on the usage of implication in the material meaning.

These sentences appear especially paradoxical if they are formulated in words of common language, and if the implications are expressed with the help of phrases containing "*implies*" or "*follows*", for instance, if we state them in the following way:

if p is true, then p follows from any q (in other words: *a true sentence follows from every sentence);*

if p is false, then p implies any q (in other words: *a false sentence implies every sentence);*

for any p and q, p implies q, or q implies p (in other words: *at least one of any two sentences implies the other).*

In such formulations, these statements frequently gave rise to misunderstandings and superfluous discussions. We have therefore additional support for the remarks which were made at the end of Section 9.*

14 An application of laws of sentential calculus in inference

Almost all reasonings in any scientific domain involve explicitly or implicitly various laws of sentential calculus; we shall try to explain by means of an example how this takes place.

If we are given a sentence having the form of an implication, then, in addition to its converse of which we spoke already in Section 10, we can form two other sentences: the INVERSE SENTENCE (or, the INVERSE OF THE GIVEN SENTENCE) and the CONTRAPOSITIVE SENTENCE. The inverse sentence is obtained by replacing both the antecedent and the consequent

‡The second of these sentences illustrates a convention which governs the symbol "$\sim$" (and which was stated in a different way in the beginning of this section). Namely, this symbol negates the shortest expression which follows directly and which constitutes a sentential function. The second sentence therefore means: "$(\sim p) \rightarrow (p \rightarrow q)$", and not: "$\sim [p \rightarrow (p \rightarrow q)]$".

of the given sentence by their negations. The contrapositive is the result of interchanging the antecedent and the consequent in the inverse sentence; the contrapositive sentence is, therefore, the converse of the inverse sentence and also the inverse of the converse sentence. The converse, the inverse, and the contrapositive sentences, together with the original sentence, are referred to as CONJUGATE SENTENCES.—As an illustration, we may consider the following conditional sentence:

(I) *if x is a positive number, then $2x$ is a positive number,*

and form its three conjugate sentences:

if $2x$ is a positive number, then x is a positive number;

if x is not a positive number, then $2x$ is not a positive number;

if $2x$ is not a positive number, then x is not a positive number.

In this particular example, all the conjugate sentences which are obtained from a true sentence turn out to be likewise true. But in general this is not at all so; it is entirely possible not only for the converse sentence (as was mentioned already in Section 10) but also for the inverse sentence to be false, although the original sentence is true, and in order to see this, it suffices to replace "$2x$" by "x^2" in the above sentences.

We thus conclude, that from the truth of an implication nothing definite can be inferred about the truth of the converse or of the inverse sentence. The situation is quite different in the case of the fourth conjugate sentence; whenever an implication is true, the same applies to the corresponding contrapositive sentence. This fact may be confirmed by numerous examples, and it finds its expression in a general law of sentential calculus, namely in the so-called LAW OF CONTRAPOSITION, or, OF TRANSPOSITION.

In order to formulate this law in a precise way, we first observe that every implication may be given the schematic form:

if p, then q;

the converse, the inverse, and the contrapositive sentences will then assume the forms:

if q, then p; if not p, then not q; if not q, then not p.

The law of contraposition, according to which any conditional sentence implies the corresponding contrapositive sentence, may therefore be formulated as follows:

if: if p, then q, then: if not q, then not p.

In order to avoid the accumulation of the words "*if*", it is expedient to make a slight change in the formulation:

(II) *From: if p, then q, it follows that: if not q, then not p.*

* * *

When a statement has the form of an implication—as, for instance, statement (I)—then we can derive its contrapositive statement with the help of the law which has just been given, in the following way.

(II) remains true if we substitute arbitrary sentences or sentential functions for "p" and "q". In particular, we can substitute for "p" and "q" the expressions:

x is a positive number

and

$2x$ is a positive number.

Changing, for stylistic reasons, the position of the word "*not*", we obtain:

(III) *from: if x is a positive number, then $2x$ is a positive number, it follows that: if $2x$ is not a positive number, then x is not a positive number.*

Now compare (I) and (III): (III) has the form of an implication, (I) being its hypothesis. Since the whole implication as well as its hypothesis have been acknowledged as true, the conclusion of the implication must likewise be acknowledged as true; but that is just the contrapositive statement in question:

if $2x$ is not a positive number, then x is not a positive number.

Anyone who knows the law of contraposition can establish in this way the contrapositive sentence as true, provided he or she has already proved the original sentence. Moreover, as one can easily verify, the inverse sentence is contrapositive with respect to the converse of the original sentence (that is to say, the inverse sentence can be obtained from the converse sentence by replacing the antecedent and the consequent by their negations and then interchanging them); for this reason, if the converse of the given sentence has been proved, the inverse sentence may likewise be considered true. If, therefore, one has succeeded in proving two sentences—the original and its converse—a special proof for the two remaining conjugate sentences is superfluous.

It may be mentioned that several variants of the law of contraposition can be given; one of them is the converse of (II):

From: if not q, then not p, it follows that: if p, then q.

This statement allows one to derive the original sentence from the contrapositive, and the converse from the inverse sentence.

15 Rules of inference, complete proofs

Let us take another look at the mechanism of the proof by means of which sentence (IV) was demonstrated in the preceding section; this brings us to the following subject. Besides the rules of definition, of which we have already spoken, there are other rules of a somewhat similar character, namely, the RULES OF INFERENCE, or, RULES OF PROOF. These rules, which must not be mistaken for logical laws, amount to directions as to how sentences which are already known to be true may be transformed so as to yield new true sentences. In the above proof we made use of two such rules, namely, of the RULE OF SUBSTITUTION and the RULE OF DETACHMENT (also known as the MODUS PONENS RULE).

* * *

The content of the rule of substitution is as follows. If a universal sentence, which has already been accepted as true, contains sentential variables, and if these variables are replaced by other sentential variables or by sentential functions or by sentences—always substituting the same expression for a given variable throughout—, then the sentence obtained in this way may also be recognized as true. It was by applying this particular rule that we obtained sentence (III) from sentence (II). It should be emphasized that the rule of substitution may also be applied to other kinds of variables, for example, to the variables "x", "y", ... designating numbers: in place of these variables, any symbols or expressions denoting numbers may be substituted.[†]

* * *

*The formulation of the rule of substitution as given above is not quite precise. This rule refers to sentences which are composed of a universal quantifier followed by a sentential function, the latter containing variables which are bound by the universal quantifier. When we want to apply the rule of substitution, we omit the quantifier and substitute for the variables which were previously bound by this quantifier other variables or related compound expressions (e.g., sentential functions or sentences for the variables "p", "q", "r", ..., and expressions denoting numbers for the variables "x", "y", "z", ...); any other bound variables which may occur in the sentential function have to remain unaltered, and in the substituted expressions we cannot admit any variables having the same form as the

[†]The substitution rule reminds us of the replacements which were discussed in Chapter I. In that chapter, however, we were concerned with transforming sentential functions into sentences, rather than with rules of inference.—We remark that when considering the rule given in the text, the word "*substitution*" is traditional. The word "*replacement*" is commonly used in stating certain other rules (cf. Sections 17 and 45). Either word would be suitable in an informal discussion as in Chapter I.

bound ones; finally, if necessary, a universal quantifier is set in front of the expression which is obtained in this way, in order to turn it into a sentence.—Let us illustrate these considerations by applying the rule of substitution to the sentence:

for any number x there is a number y such that $x + y = 5$,

from which the following sentence can be obtained:

there is a number y such that $3 + y = 5$,

and also the sentence:

for any number z there is a number y such that $z^2 + y = 5$;

in these two examples we substituted for "x" while leaving "y" unaltered. We must not, however, substitute for "x" any expression containing "y"; for, although our original sentence was true, we might arrive in this way at a false sentence. For instance, by substituting "$3 - y$", we would obtain:

there is a number y such that $(3 - y) + y = 5$.*

* * *

The rule of detachment states that, if two sentences are accepted as true, one of which is an implication while the other is the antecedent of this implication, then the consequent of the implication may also be recognized as true.†† (We "detach", so to speak, the antecedent from the whole implication.) For instance, this rule allows the derivation of sentence (IV) from sentences (III) and (I).

* * *

One can now verify that in the proof of sentence (IV), as carried out above, each step consisted in applying a rule of inference to sentences which were previously accepted (or, recognized) as true. A proof of this kind will be called COMPLETE. A complete proof may also be characterized a little more precisely as follows. It consists in the construction of *a chain of sentences with these properties: the initial members are sentences which were previously accepted as true; every subsequent member is obtainable from preceding ones by applying a rule of inference; and finally, the last member is the sentence to be proved.*

It should be observed what an extremely elementary form—from the psychological point of view—all complete proofs assume, as a result of their

†† We should like to paraphrase some of the foregoing specifications. When we are dealing with a scientific theory, then its true sentences are just its laws (cf. Section 1). One can therefore say, that a rule of inference provides a method of transforming given laws into new ones.—Evidently, such a rule itself cannot be regarded as one of the laws, whether belonging to logic or to a particular theory. Compare the corresponding remark in the opening paragraph of this section.

being based on the knowledge of laws of logic and on the application of the rules of inference; complicated mental processes can be fully reduced to such simple tasks as the attentive observation of statements which were previously accepted as true, the perception of structural (purely external) connections among these statements, and the execution of mechanical transformations as prescribed by the rules of inference. It is obvious that, when such a procedure is adopted, the possibility of committing mistakes in a proof is virtually eliminated.[‡]

Exercises

1. Give examples of specifically mathematical expressions from the fields of arithmetic and geometry.

2. Distinguish in the following two sentences between the specifically mathematical expressions and those belonging to the domain of logic (the identity symbol "=" and the word "*same*" should be included among the latter):

(a) *for any numbers x and y, if $x > 0$ and $y < 0$, then there is a number z such that $z < 0$ and $x = y \cdot z$;*
(b) *for any points A and B there is a point C which lies between A and B and is at the same distance from A as from B.*

3. Form the conjunction of the negations of the following sentential functions:

$$x < 3$$

and

$$x > 3.$$

What number satisfies this conjunction?

4. State in which of its two meanings the word "*or*" occurs in the following sentences:

(a) *two ways were open to him: to betray his country or to die;*
(b) *if I earn a lot of money or win the sweepstake, I shall go on a long journey.*

[‡]The last two sections are perhaps too short to fully convey the following message: The sentential calculus is, first and foremost, a tool for making deductions. This message will be reinforced in the remainder of the book, where the reader will come upon a variety of applications of this calculus to mathematical arguments. On the other hand, the analysis of sentential connectives and the verification of laws by the method of truth tables (to which the bulk of this chapter is devoted) should be considered as preparation for the indicated role, and not as ends in themselves.

Give further examples where the word "*or*" is used in its first meaning, and examples where it is used in its second meaning.

***5.** Consider the following conditional sentences:

(a) *if today is Monday, then tomorrow is Tuesday;*
(b) *if today is Monday, then tomorrow is Saturday;*
(c) *if today is Monday, then the 25th of December is Christmas day;*
(d) *if wishes were horses, beggars could ride;*
(e) *if a number is divisible by 2 and by 6, then it is divisible by 12;*
(f) *if 18 is divisible by 3 and by 4, then 18 is divisible by 6.*

Which of the above implications are true and which are false from the point of view of mathematical logic? In which cases do the questions, of meaningfulness and of truth or falsity, raise any doubts from the standpoint of ordinary language? Direct special attention to sentence (b), and examine how its truth depends on the day of the week on which the sentence was uttered.

6. Put the following theorems into the form of ordinary conditional sentences:

(a) *for a triangle to be equilateral, it is sufficient that the angles of the triangle be congruent;*
(b) *the condition: x is divisible by 3, is necessary for x to be divisible by 6.*

Give further paraphrases of these two sentences.

7. Is the condition:

$$x \cdot y > 4$$

necessary or sufficient for:

$$x > 2 \text{ and } y > 2\,?$$

8. Give alternative formulations for the following sentences:

(a) *x is divisible by 10 if, and only if, x is divisible both by 2 and by 5;*
(b) *for a quadrangle to be a parallelogram it is necessary and sufficient that the point of intersection of its diagonals be at the same time the midpoint of each diagonal.*

Give further examples of theorems from the fields of arithmetic and geometry that have the form of equivalences.

9. Which of the following sentences are true?

(a) *a triangle is isosceles if, and only if, all the altitudes of the triangle are congruent;*
(b) *the condition that $x \neq 0$ is necessary and sufficient for x^2 to be a positive number;*

(c) *the fact that a quadrangle is a square implies that all its angles are right angles, and conversely;*
(d) *for x to be divisible by 8 it is necessary and sufficient that x be divisible both by 4 and by 2.*

10. Assuming the terms "*natural number*"† and "*product*" (or "*quotient*") to be known already, construct a definition of the term "*divisible*", giving it the form of an equivalence:

we say that x is divisible by y if, and only if,....

Likewise formulate a definition of the term "*parallel*"; what terms (from the domain of geometry) have to be presupposed for this purpose?

11. Translate the following symbolic expressions into ordinary language:

(a) $(\sim p \rightarrow p) \rightarrow p$,
(b) $(\sim p \vee q) \leftrightarrow (p \rightarrow q)$,
(c) $\sim (p \vee q) \leftrightarrow (p \rightarrow q)$,
(d) $\sim p \vee [q \leftrightarrow (p \rightarrow q)]$.

Direct special attention to the difficulty in distinguishing the last three expressions, one from another, when stated in ordinary language.

12. Formulate the following expressions in logical symbolism:

(a) *if not p or not q, then it is not the case that p or q;*
(b) *if p implies that q implies r, then p and q together imply r;*
(c) *if r follows from p and r follows from q, then r follows from p or q.*

13. Construct truth tables for all the sentential functions given in exercises 11 and 12. Assume that we interpret these functions as sentences (what does this mean?), and determine which of the sentences obtained in this way are true and which are false.

14. Verify by the method of truth tables that the following sentences are true:

(a) $\sim\sim p \leftrightarrow p$,
(b) $\sim (p \wedge q) \leftrightarrow (\sim p \vee \sim q)$,
$\sim (p \vee q) \leftrightarrow (\sim p \wedge \sim q)$,
(c) $[p \wedge (q \vee r)] \leftrightarrow [(p \wedge q) \vee (p \wedge r)]$,
$[p \vee (q \wedge r)] \leftrightarrow [(p \vee q) \wedge (p \vee r)]$.

Sentence (a) is the LAW OF DOUBLE NEGATION, sentences (b) are called DE MORGAN'S LAWS,[6] and sentences (c) are the DISTRIBUTIVE LAWS (for logical multiplication with respect to addition and for logical addition with respect to multiplication).

---

†We recall that a *natural number* is a number which is a positive integer or zero, i.e. one of the following: 0, 1, 2, etc.

[6]These laws were given by A. DE MORGAN (1806–1878), an eminent English logician.

15. For each of the following sentences, state the three corresponding conjugate sentences (the converse, the inverse, and the contrapositive sentence):

(a) *the hypothesis that x is a positive number implies that $-x$ is a negative number;*
(b) *if a quadrangle is a rectangle, then a circle can be circumscribed about it.*

Which of the conjugate sentences are true?

Why can one not give an example of four conjugate sentences which are all false?

16. Explain the following fact on the basis of the truth table for the function "$p \leftrightarrow q$": if in any sentence some of its parts which are themselves sentences are replaced by equivalent sentences, then the whole new sentence which is obtained in this way is equivalent to the original sentence (and we note that this conclusion can be extended to sentential functions). Some of our statements and remarks in Section 11 were dependent on this fact; indicate where this was the case.

17. Consider the following two sentences:

(a) *from: if p, then q, it follows that: if q, then p;*
(b) *from: if p, then q, it follows that: if not p, then not q.*

Suppose these sentences were logical laws; would it then be possible to apply them in mathematical proofs, as was done with the law of contraposition in Section 14? Which conjugate sentences would it be possible to derive from a given asserted implication? Consequently, can our supposition be maintained that sentences (a) and (b) are true?

18. Confirm the conclusion which was reached in exercise 17 by applying the method of truth tables to sentences (a) and (b).

19. Consider the following two statements:

the assumption that yesterday was Monday implies that today is Tuesday;
the assumption that today is Tuesday implies that tomorrow will be Wednesday.

What statement may be deduced from them in accordance with the law of the hypothetical syllogism (cf. Section 12)?

***20.** Carry out a complete proof of the statement which was obtained in the preceding exercise; use the statements given there and the law of the hypothetical syllogism, and apply—in addition to the rule of substitution and the rule of detachment—the following rule of inference: if any two sentences are accepted as true, then their conjunction may also be recognized as true.

This page intentionally left blank

III

On the Theory of Identity

16 Logical concepts not belonging to sentential calculus; the concept of identity

Sentential calculus, to which the preceding chapter was devoted, forms only one of the parts of logic. It constitutes undoubtedly the most fundamental part,—at least inasmuch as one necessarily uses its terms and laws in defining other terms, and in formulating and demonstrating logical laws that do not belong to this calculus. Sentential calculus by itself, however, does not form an adequate basis for the foundations of other sciences and, in particular, not for the foundations of mathematics; various concepts from other branches of logic are constantly encountered in mathematical definitions, theorems, and proofs. Some of them will be discussed in the present and in the following two chapters.

* * *

Among the logical concepts not belonging to sentential calculus, the concept of IDENTITY, or, of EQUALITY, is perhaps the one which has the greatest importance. It occurs in phrases such as:

x is identical with y,

x is the same as y,

x equals y.

We ascribe the same meaning to all three phrases; for the sake of brevity, they will be replaced by the symbolic expression:

$$x = y.$$

Moreover, instead of writing:

x is not identical with y

or

x is different from y

we employ the formula:

$$x \neq y.$$

The general laws governing the above expressions constitute a part of logic, which we shall call the THEORY OF IDENTITY.

17 Fundamental laws of the theory of identity

Among logical laws which involve the concept of identity, the most fundamental is the following:

I. $x = y$ *if, and only if, x has every property which y has, and y has every property which x has.*

We could also say more simply:

$x = y$ *if, and only if, x and y have every property in common.*

Other and perhaps more transparent, though less correct, formulations of the same law are known, for instance†:

$x = y$ *if, and only if, everything that may be said about one of the objects x or y may also be said about the other.*

Law I was first stated by LEIBNIZ[1] (although in somewhat different terms) and hence may be called LEIBNIZ'S LAW. It has the form of an equivalence, and enables us to replace the formula:

$$x = y,$$

which is the left side of the equivalence, by its right side, that is, by an expression no longer containing the symbol of identity. In view of its form, this law may be considered as the definition of the symbol "=", and it was so considered by LEIBNIZ himself. (Of course, it would make sense to regard LEIBNIZ's law as a definition only if the meaning of the symbol "=" seemed to us less evident than the meaning of the expressions on the

†Throughout this book the word "*object*" may refer not only to concrete, physical objects, but also to abstractions like numbers and points. Such abstractions (and more sophisticated ones) of course occur constantly in mathematics, and are also referred to as *entities*.—The reader may have observed that properties of objects played a secondary role in the preceding chapter. However, objects and their properties will be central for the discussions from now on.

[1]Cf. footnote 1 on p. 18.

corresponding right side, such as "*x has every property which y has*"; cf. Section 11.)

As a consequence of LEIBNIZ's law we have the following RULE OF REPLACEMENT OF EQUALS, which is of great practical importance: if, in a certain context, a formula having the form of an equation, e.g.:

$$x = y^2,$$

has been assumed or proved, then one is allowed to replace in any formula or sentence occurring in this context the left side of the equation by its right side, "x" by "y^2" in this case, and the right side by the left side. It is moreover understood that, should "x" occur at several places in a formula, it may be left unchanged at some places, and at others replaced by "y^2"; we have therefore an essential difference between the rule now under consideration and the rule of substitution discussed in Section 15, which does not permit such a partial replacement of one symbol by another.[††]

* * *

From LEIBNIZ's law we can derive a number of other laws which belong to the theory of identity, and which are frequently applied in various considerations and especially in mathematical proofs. The most important of these will be listed here, together with sketches of their proofs; we shall then see by way of concrete examples that there is no essential difference between reasonings in the field of logic and those in mathematics.

II. *Every object is equal to itself:* $x = x$.

Proof.[‡] Use the rule of substitution, and substitute in LEIBNIZ's law "x" for "y"; we obtain:

$x = x$ *if, and only if, x has every property which x has, and x has every property which x has.*

We can, of course, simplify this sentence by omitting its last part "*and x has...*" (this is a consequence of one of the laws of tautology which were stated in Section 13). The sentence assumes then the following form:

$x = x$ *if, and only if, x has every property which x has.*

[††]The reader no doubt remembers the rule of replacement of equals from his or her first lessons on the solution of equations, and appreciates its significance. There are reasons, however, for not including it among the basic rules of inference when discussing complete proofs (cf. Section 15).—We note that in the previous editions of *Introduction to Logic* this rule was introduced in an informal manner, without a special emphasis, and was hardly mentioned in subsequent arguments. LEIBNIZ's law was cited instead. In the present edition the law is cited as in the previous editions, and the rule is referred to on several occasions as well.

[‡]We recall again the notion of a complete proof. (Cf. the preceding footnote and Section 15.) Often, however, demonstrations of logic and mathematics are presented without strict adherence to rules such as were described, and this proof is an example. Such proofs are sometimes called INFORMAL.

Obviously, the right side of this equivalence is always satisfied (for, according to the law of identity in Section 12, if x has a certain property, then it has this property). Hence the left side of the equivalence must also be satisfied; in other words, we always have:

$$x = x,$$

which was to be proved.

III. *If* $x = y$, *then* $y = x$.

Proof. By substituting in LEIBNIZ's law "x" for "y" and "y" for "x", we obtain:

$y = x$ *if, and only if,* y *has every property which* x *has, and* x *has every property which* y *has.*

Let us compare this sentence with the original formulation of LEIBNIZ's law. We have here two equivalences, such that their right sides are conjunctions which differ only in the order of their members. Hence the right sides are equivalent (cf. the commutative law for logical multiplication in Section 13), and therefore the left sides, that is, the formulas:

$$x = y \text{ and } y = x$$

must also be equivalent. *A fortiori* we may assert that the second of these formulas follows from the first, and our law is established.

IV. *If* $x = y$ *and* $y = z$, *then* $x = z$.

Proof. By hypothesis, the two formulas:

(1) $$x = y$$

and

(2) $$y = z$$

are assumed valid.[‡‡] According to LEIBNIZ's law, it follows from formula (2) that every property of y applies also to z. Hence we may replace the variable

[‡‡]We should like to comment on the two words, "*valid*" and "*true*". The term "*true sentence*" occurs throughout Chapters I and II. This term and its opposite, "*false sentence*", have been the subject of many philosophical debates, but they are used in this book in a straightforward manner: with reference to sentences of elementary mathematical theories or to simple examples from everyday language. The term "*valid*", on the other hand, is broader in scope (and it has not given rise to such controversies). It is often used in mathematics in a similar way as "*true*". Thus, "*valid*" may mean "*true*" when referring to a special context (in particular, when referring to chosen models and interpretations, cf. Chapters VI and IX). In fact, situations sometimes arise when the two words are used interchangeably. There is, moreover, a certain preference for using "*valid*" when one deals with sentential functions, as in the above proof. Then "*valid*" means about the same as "*satisfied under every substitution of variables*". Furthermore, sometimes "*valid*" means the same as "*proved*" or "*provable*".—Other uses of this word are familiar, and one says, for instance, "*a valid argument*".

"y" by "z" in formula (1) (in fact, one could justify this step simply by referring to the rule of replacement of equals), and we obtain the desired conclusion:

$$x = z.$$

V. *If* $x = z$ *and* $y = z$, *then* $x = y$; in other words, *two objects which are equal to the same object are equal to each other.*

This law can be proved by an argument analogous to the preceding (it can also be deduced from Laws III and IV, without using LEIBNIZ's law).

Laws II, III, and IV are called the LAWS OF REFLEXIVITY, SYMMETRY, and TRANSITIVITY, respectively, for the relation of identity.

18 Identity of objects and identity of their designations; the use of quotation marks

*Although the meaning of such expressions as:

$$x = y \text{ and } x \neq y$$

appears to be evident, these expressions are sometimes misunderstood. It seems obvious, for instance, that the formula:

$$3 = 2 + 1$$

is a true assertion, and yet some people remain skeptical as to its truth. In their opinion, this formula seems to state that the symbols "3" and "$2+1$" are identical, which is obviously false since these symbols have entirely different shapes, and therefore, it is not true that everything which may be said about one of these symbols may also be said about the other (we might add, the first symbol is a single sign, while the second is not).

* * *

In order to dispel any lingering doubts of this kind, one should make clear a general and very important principle, which must be retained if a language is to be usefully employed. According to this principle, whenever we wish to state a sentence about a certain object, we have to use in this sentence not the object itself, but its name or designation.

The application of this principle does not give rise to any problems as long as the object in question is not a word or a symbol or, more generally, an expression of a language. Let us imagine, for example, that we have a small blue stone in front of us, and that we write the following sentence:

this stone is blue.

To no one, presumably, would it occur to replace in this sentence the words "*this stone*" (which together constitute the designation of the object) by the object itself, that is to say, to blot or cut these words out and to place in their stead the stone. For, in doing so, we would arrive at a whole consisting partly of a stone and partly of words, and therefore at something which would not be a linguistic expression, and all the more, would not be a true sentence.

This principle, however, is frequently violated if the object which is talked about happens to be a word or a symbol. And yet the application of the principle is indispensable also in this case; for otherwise we would arrive at a whole which, though being a linguistic expression, would fail to convey the thought intended by us, and might even be a meaningless aggregate of words. Let us consider, for example, the following two words:

good, Mary.

Clearly, the first consists of four letters, and the second is a proper name. But let us suppose that we would express these thoughts, which are quite correct, in the following manner:

(1) *good consists of four letters;*

(2) *Mary is a proper name;*

then, in talking about words, we would be using the words themselves and not their names. And if we examine expressions (1) and (2) more closely, we must admit that the first is not a sentence at all, since the subject of a sentence can only be a noun and not an adjective; the second might be considered a meaningful sentence, but, at any rate, would be a false one since no woman is a proper name.

In order to clarify these situations, we have to realize that when the words "*good*" and "*Mary*" occur in such contexts as those of (1) and (2), then their meanings differ from the usual ones, and that here these words function as their own names. In generalizing this viewpoint, we should have to admit that any word may, at times, function as its own name; to use the terminology of medieval logic, we may say that in such a case the word is used in SUPPOSITIO MATERIALIS, as opposed to its use in SUPPOSITIO FORMALIS, that is, in its ordinary meaning.† As a consequence, every word of common or scientific language could take on two or more different meanings, and one does not have to look far for examples of situations

†The reader may find it instructive to compare the preceding discussion with that of Section 9. Both discussions concern expressions, but the expressions are of different kinds. In essence, those of one kind denote objects, while those of the other refer to assertions, respectively.—With regard to the latter, i.e. to sentences and sentential functions, we say that we USE a sentence when making a direct assertion, while in an indirect formulation, as in *loc. cit.*, we MENTION the given sentence.

where serious doubts might arise as to which meaning was intended. We do not wish to reconcile ourselves with such ambiguities, and therefore we will make it a rule that every expression should differ (at least in writing) from its name.

* * *

The problem arises as to how we can go about forming names of words and expressions. There are various devices that can be used. The simplest of them is based on the convention, that one forms a name of an expression by placing the latter between quotation marks. On the basis of this convention, the thoughts tentatively expressed in (1) and (2) can now be stated correctly and without ambiguity in the following way:

(I) *"good" consists of four letters;*

(II) *"Mary" is a proper name.*

In the light of these remarks, all possible doubts as to the meaning and the truth of such formulas as:

$$3 = 2 + 1$$

are dispelled. This formula contains symbols designating certain numbers, but it does not contain the names of any such symbols. Therefore this formula states something about numbers and not about symbols designating numbers; the number 3 is obviously equal to the number $2 + 1$, so that the formula is a true assertion. We may, admittedly, replace this formula by a sentence which expresses the same idea but is about symbols, namely, by a sentence which asserts that the symbols "3" and "$2 + 1$" designate the same number. But this by no means implies that the symbols themselves are identical; for it is well known that a given object—and in particular, a given number—can have many different designations. The symbols "3" and "$2 + 1$" are, no doubt, different, and this observation can also be expressed by a new formula:

$$\text{“3”} \neq \text{“2 + 1”}$$

which, of course, does not contradict in any way the formula previously stated.[2]

[2]The convention concerning the use of quotation marks has been adhered to in this book pretty consistently. We deviate from it only in special cases, by way of a concession to traditional usage. For instance, we state formulas and sentences without quotation marks, if they are displayed in a separate line or if they occur in the formulation of mathematical or logical theorems; and we do not put quotation marks about expressions which are preceded by such phrases as "*is called*", "*is known as*", and so on. But other precautionary measures are taken in these cases; the expression in question is often preceded by a colon, and usually it is printed in a different kind of print (small capitals or italics). It should moreover be observed that, in everyday language, quotation marks are used also under certain other circumstances; and examples of such usage can be found in this book, too.

19 Equality in arithmetic and in geometry, and its relationship to logical identity

In this book we consider the notion of equality among numbers always as a special case of the general concept of logical identity.† One should add, however, that there have been (and perhaps still are) mathematicians who—as opposed to the standpoint adopted here—did not identify the symbol "=" occurring in arithmetic with the symbol of logical identity; they did not consider equal numbers to be necessarily identical, and therefore looked upon the notion of equality among numbers as a specifically arithmetical concept. In connection with this outlook, those mathematicians rejected LEIBNIZ's law in its general form, and instead, recognized various of its consequences, such as those which relate to numbers; the consequences in question are then of a less general character and are therefore regarded as laws which are specifically mathematical. Among these consequences there are Laws II to V of Section 17, as well as assertions to the effect that, whenever $x = y$ and x satisfies some formula which is constructed out of arithmetical symbols only, then y satisfies the same formula; the following law is an example:

$$\textit{if } x = y \textit{ and } x < z, \textit{ then } y < z.$$

In our opinion, this point of view can claim no particular theoretical advantages, while in practice, it entails considerable complications in the presentation of arithmetic. For one rejects here the general rule which allows us—on the assumption that an equation holds—to replace everywhere the left side of the equation by its right side; however, such a replacement is indispensable in various arguments, and so it becomes necessary in each particular case to give a special proof that this replacement is permissible.

To illustrate this kind of situation by an example, let us consider a system of equations in two variables, for instance:

$$x = y^2,$$
$$x^2 + y^2 = 2x - 3y + 18.$$

†Several footnotes of this chapter have dealt with the usage of words. We should like to consider now "*concept*", and specifically "*concept of identity*". In general, one can think of a concept as a kind of an inclusive idea, a synthesis resulting from several ingredients. Accordingly, if a broad interpretation is envisaged, then the concept of identity could encompass the material of the whole present chapter. However, the word "*concept*" is sometimes used in a more restrictive way, and one might then think of the concept of identity as simply the term "=", when this term is defined (or characterized) by LEIBNIZ's law. And speaking of words, instead of saying '*the term "="*', one could just as well say '*the (logical) constant "="*', or even, '*the symbol "="*'.—The usage of "*concept*" in this book is primarily of the latter, more restrictive kind.

In order to solve this system of equations by means of the so-called method of substitution, one has to form a new system of equations, which is obtained by leaving the first equation unchanged and by replacing in the second equation "x" by "y^2" throughout. And the question arises whether this transformation is permissible, that is, whether the new system is equivalent to the old. The answer is undoubtedly in the affirmative, no matter what conception of the notion of equality among numbers is adopted. But if the symbol "=" is understood to designate logical identity, and if LEIBNIZ's law is assumed, the answer is obvious; the assumption:

$$x = y^2$$

permits us to replace "x" everywhere by "y^2", and vice versa. Otherwise it would first be necessary to give reasons for the affirmative answer, and although such a justification would not involve any essential difficulties, it could at any rate be rather long and tedious (such proofs can be found in some of the less recent elementary textbooks).

* * *

As to the notion of equality in geometry, traditionally there has been a point of view of an entirely different kind. One sometimes says that two geometrical figures, such as two line segments, or two angles, or two polygons, are equal, in the sense of "*congruent*", and one does not intend to assert their identity by such a statement. Rather, he or she only wishes to state that the two figures have the same size and shape, in other words—to use a figurative if not quite correct mode of expression—, that each figure would exactly cover the other if it were placed on top of the other. For example, a triangle is capable of having two sides, or even three, which are equal in this sense—we shall say in the present discussion that the sides are then "geometrically equal"—and yet these sides are certainly not identical. On the other hand, there are also cases in which it is not a question of the geometrical equality of two figures, but of their logical identity; in an isosceles triangle, for instance, the altitude upon the base and the median to the base are not only geometrically equal, but they are simply the same line segment. Therefore, in order to avoid any confusion, it would be recommendable to avoid consistently the term "*equality*" in all those cases where it is not a question of logical identity, and instead, to speak of geometrically equal figures as congruent figures, replacing at the same time—as it is often done anyhow—the symbol "=" by a different one, such as "$\cong$".

20 Numerical quantifiers

By using the concept of identity, it is possible to give a precise meaning to certain phrases which are closely related (with regard to both content and form) to the universal and the existential quantifiers, and which are also classified as operators, but which are of a more special character. The expressions in question are such as:

> *there is at least one,* or *at most one,* or *exactly one, object x such that...,*
>
> *there are at least two,* or *at most two,* or *exactly two, objects x such that...,*

and so on; they might be called NUMERICAL QUANTIFIERS. It appears that specifically mathematical terms occur in these phrases, namely, the numerals "*one*", "*two*", and so on. A more detailed analysis shows, however, that the content of these phrases (if each is considered as a whole) is of a purely logical nature. Thus, in the expression:

> *there is at least one object satisfying the given condition*

the words "*at least one*" may simply be replaced by the article "*an*" without altering the meaning. The expression:

> *there is at most one object satisfying the given condition*

means the same as:

> *for any x and y, if x satisfies the given condition and if y satisfies the given condition, then $x = y$.*

The sentence:

> *there is exactly one object satisfying the given condition*

is equivalent to the conjunction of the two sentences just given:

> *there is at least one object satisfying the given condition, and at the same time there is at most one object satisfying the given condition.*

To the expression:

> *there are at least two objects satisfying the given condition*

we give the following meaning:

> *there are x and y, such that both x and y satisfy the given condition and $x \neq y$;*

each of these two expressions is, in fact, equivalent to the negation of:

there is at most one object satisfying the given condition.

In a similar way we explain the meanings of other expressions of this category.

For the purpose of illustration, a few true sentences of arithmetic may be listed here in which numerical quantifiers appear:

there is exactly one number x, such that $x + 2 = 5$;

there are exactly two numbers y, such that $y^2 = 4$;

there are at least two numbers z, such that $z + 2 < 6$.

* * *

That part of logic, in which general laws involving quantifiers are established, is known as PREDICATE CALCULUS or PREDICATE LOGIC.† In this theory the universal and the existential quantifiers are of primary concern, while numerical quantifiers of the above kind have been largely neglected.

Exercises

1. Prove Law V of Section 17 by using exclusively Laws III and IV, thus without the use of LEIBNIZ's law (or of the rule of replacement of equals). Hint: In Law V the formulas:

$$x = z \text{ and } y = z$$

are assumed valid by hypothesis. By virtue of Law III, interchange the variables in the second of these formulas, and then apply Law IV.

†The word "*predicate*" refers to features (in technical terms: properties and relations, cf. Chapters IV and V) which may be under discussion, as for instance: *is less than*, *is a perfect square*, *is congruent to*, *is a regular polygon*.—The two options, "*predicate calculus*" and "*predicate logic*", like the alternatives "*sentential calculus*" and "*sentential logic*" (cf. the beginning of Section 7) are not just a matter of choice of words, but rather, they indicate different attitudes toward the subject. When we say "*calculus*", we emphasize a structure which shows some resemblance to computations, even though arithmetical calculations might not be involved. The truth tables provide an illustration. On the other hand, the word "*logic*" points to the essential purpose of these disciplines, namely, to establish the nature of mathematical arguments. "*Logic*" is now widely used in naming these disciplines, while formerly the word "*calculus*" was predominant. In this book the name "*sentential calculus*" is retained from the previous editions.—In earlier editions of this book, furthermore, other (and older) names were mentioned: "*theory of deduction*" for sentential calculus and "*theory of apparent variables*" and "*functional calculus*" for predicate logic.

2. Prove the following law:

$$\textit{if } x = y, \ y = z, \textit{ and } z = t, \textit{ then } x = t,$$

using exclusively Law IV of Section 17.

3. Are the sentences true which are obtained by replacing in Laws III and IV of Section 17 the symbol "=" by "≠" throughout?

***4.** On the basis of the convention stated in Section 18 concerning the use of quotation marks, determine which of the following sentences are true:

(a) 0 *is an integer,*
(b) 0 *is a cipher having an oval shape,*
(c) "0" *is an integer,*
(d) "0" *is a cipher having an oval shape,*
(e) $1.5 = \frac{3}{2}$,
(f) "1.5" = "$\frac{3}{2}$",
(g) $2 + 2 \neq 5$,
(h) "2 + 2" ≠ "5".

***5.** In order to form a name of a word, we put that word in quotation marks; in order to form a name of this name we put, in turn, the name of this word in quotation marks, and thus the word itself in double quotation marks. Hence, of the following three expressions:

John, "*John*", " "*John*" ",

the second is a name of the first, and the third is a name of the second, while the first is a name of a man. Substitute in turn the three above expressions for "x" in the following sentential functions, and determine which of the twelve sentences obtained in this way are true:

(a) x *is a man,*
(b) x *is a name of a man,*
(c) x *is an expression,*
(d) x *is an expression containing quotation marks.*

***6.** In Section 9 we gave conditional sentences in various formulations which are encountered in mathematical books. Attention was also called to the fact that in some of these formulations we talk, not about numbers or properties of numbers, and so on, but about expressions (in particular, about sentences and sentential functions). It follows from remarks made in Section 18 that these latter formulations call for the use of quotation marks. Indicate the formulations, and the exact places in them, in which quotation marks should be used.

***7.** Having discussed the general principle concerning the use of names of objects in sentences which state something about these objects, we may now subject the last sentence but one of Section 12 ("Just as the arithmetical theorems...") to a certain scrutiny. We know that variables occurring in arithmetic stand for names of numbers. Do the variables which occur in sentential calculus stand for names of sentences or for sentences themselves?[†] May we therefore say, if we want to be exact, that the laws of this calculus assert something about sentences and their properties?

8. Consider a triangle with sides a, b, and c. Let h_a, h_b, and h_c be the altitudes upon the respective sides a, b, and c; similarly, let m_a, m_b, and m_c be the medians and s_a, s_b, and s_c, the bisectors of the angles of the triangle.

Assuming the triangle to be isosceles (with a as the base, and b and c as the sides of equal length), which of the twelve segments named are congruent (so to say, equal in the geometrical sense), and which are identical? Express the answer by means of formulas, using the symbol "$\cong$" to designate congruence, and the symbol "=" to designate identity.

Solve the same problem under the assumption that the triangle is equilateral.

9. Explain the meaning of the following expressions:

(a) *there are at most two objects satisfying the given condition;*
(b) *there are exactly two objects satisfying the given condition.*

10. Determine which of the following sentences are true:

(a) *there is exactly one number x such that* $x + 3 = 7 - x$;
(b) *there are exactly two numbers x such that* $x^2 + 4 = 4x$;
(c) *there are at most two numbers y such that* $y + 5 < 11 - 2y$;
(d) *there are at least three numbers z such that* $z^2 < 2z$;
(e) *for any number x there is exactly one number y such that* $x+y = 2$;
(f) *for any number x there is exactly one number y such that* $x \cdot y = 3$.

11. How can one make use of numerical quantifiers in order to express the fact that the equation:

$$x^2 - 5x + 6 = 0$$

has two roots?

12. What numbers x satisfy the sentential function:

there are exactly two numbers y such that $x = y^2$?

Distinguish in this function between free and bound variables. *Do numerical quantifiers bind variables?*

[†]We may remark that, in fact, two different kinds of variables can be introduced in the present context: those that stand for sentences (i.e., for expressions having a particular kind of structure), and those that stand for names of sentences, or, denote sentences. These two possibilities clearly relate to the use-mention juxtaposition. (Cf. footnote on p. 54.)

13. (a) Consider the following sentential function:

(1) *There are* x, y, *and* z *such that* $x < a$, $y < a$, $z < a$, $x \neq y$, $x \neq z$, *and* $y \neq z$.

Express (1) using numerical quantifiers. Formulate (1) by employing symbols which were introduced in Chapters I and II.

*(b) Translate the following formula (2) into ordinary language, and show that both (2) and (1) are true statements of arithmetic:

$$\forall_{x,y}\{(x < a \land y < a) \to \exists_z[z < a \land x \neq z \land y \neq z]\}. \tag{2}$$

IV

On the Theory of Classes

21 Classes and their elements

Apart from separate individual objects, which we shall also call INDIVIDUALS for short, logic is concerned with CLASSES of objects; in everyday life as well as in mathematics, classes are more often referred to as SETS. Arithmetic, for instance, frequently deals with sets of numbers, and in geometry our interest lies not so much in single points as in sets of points (that is, in geometrical configurations). Now, classes of individuals are called CLASSES OF THE FIRST ORDER. Relatively more rarely in our investigations we come upon CLASSES OF THE SECOND ORDER, that is, upon classes which consist, not of individuals, but of classes of the first order. Sometimes even CLASSES OF THE THIRD, FOURTH, and HIGHER ORDERS have to be dealt with. Here we shall be concerned almost exclusively with classes of the first order, and only exceptionally—as in Section 26—we shall have to deal with classes of the second order; however, our considerations can be applied with practically no changes to classes of any order.

In order to distinguish between individuals and classes (and also among classes of different orders), we employ as variables letters of different shape and belonging to different alphabets. It is customary to designate individual objects such as numbers, and classes of such objects, by lower-case and capital letters of the Latin alphabet, respectively. In elementary geometry the opposite notation is the accepted one, capital letters designating points and lower-case letters (of the Latin or the Greek alphabet) designating sets of points.

That part of logic in which classes and their properties are examined is called the THEORY OF CLASSES; sometimes this theory is also treated as an independent mathematical discipline under the name of the GENERAL

THEORY OF SETS.[1]†

* * *

Fundamental in the theory of classes are such phrases as:

the object x is an element (or, *a member) of the class K,*

the object x belongs to the class K,

the class K contains the object x as element (or, *as member);*

we consider these expressions as having the same meaning and, for the sake of expediency, replace them by the formula:

$$x \in K.$$

Thus, if I is the set of all integers, then the numbers $1, 2, 3, \ldots$ are among its elements, whereas the numbers $\frac{2}{3}$, $2\frac{1}{2}, \ldots$ do not belong to the set; hence, the formulas:

$$1 \in I, \quad 2 \in I, \quad 3 \in I, \ldots$$

are true, while the formulas:

$$\tfrac{2}{3} \in I, \quad 2\tfrac{1}{2} \in I, \ldots$$

are false.

[1]The beginnings of the theory of classes—or, to be more precise, of that part of this theory to which we shall refer below as the algebra of classes—are found already in the works of BOOLE (cf. footnote 1 on p. 18). The actual creator of the general theory of sets as an independent mathematical discipline was the great German mathematician G. CANTOR (1845–1918); in particular, we are indebted to him for the analysis of the concept of order, and of such concepts as equality in power, cardinal number, and infinity, which will be discussed at the end of the present and in the next chapter.—CANTOR's set theory is a mathematical discipline which for many years has been in a state of intensive development. Its ideas and lines of thought have penetrated into almost all branches of mathematics, and have exerted everywhere a most stimulating and fertilizing influence.

†In this book the words "*class*" and "*set*" are used as synonyms. More advanced considerations require, however, that we should distinguish between two different kinds of entities, and one then assigns somewhat different meanings to these two words. The collections which are commonly encountered in mathematics are then designated by the term "*sets*".—The symbol "$\in$" which occurs in the sequel is derived from the lower-case Greek letter epsilon: "ϵ". Sometimes this letter itself is used for denoting membership.

22 Classes and sentential functions with one free variable

We consider a sentential function with one free variable, for instance:

$$x > 0.$$

If we prefix the words:

(I) *the set of all numbers x such that*

to that function, we obtain the expression:

the set of all numbers x such that $x > 0$.

This expression designates a well-determined set, namely the set of all positive numbers; it is the set having as its elements those, and only those, numbers which satisfy the given sentential function. This set can be denoted by a special symbol, e.g. by "P"; our function then becomes equivalent to:

$$x \in P.$$

We may, in fact, apply an analogous procedure to any other sentential function. In arithmetic, for instance, we can describe in this way various sets of numbers, like the set of all negative numbers or the set of all numbers which are greater than 2 and less than 5 (that is, which satisfy the function "$x > 2$ *and* $x < 5$"). This procedure plays also an important role in geometry, especially in defining diverse geometrical figures†; the surface of a sphere is defined, for instance, as the set of all points in space which have a definite distance from a given point. It is customary in geometry to replace the words "*the set of all points*" by "*the locus of the points*".

* * *

We shall now put the above remarks into a general form. It is assumed in logic that, to every sentential function containing just one free variable, say "x", there is exactly one corresponding class having as elements those, and only those, objects x which satisfy the given function. We obtain a designation for that class by putting in front of the sentential function the following phrase, which is among the fundamental expressions of the theory of classes:

(II) *the class of all objects x such that.*

† In this book the term "*geometrical configuration*" is used as a synonym for "*set of points*" or "*point set*"; such usage was indicated in the beginning of this chapter. On the other hand, "*geometrical figure*" refers to a configuration which is, so to say, familiar and more regular, like a triangle or a sphere.

Moreover, if we denote the class in question by a single symbol, say "C", then the formula:

$$x \in C$$

will—for any x—be equivalent to the original sentential function.

Hence we see that any sentential function containing "x" as the only free variable can be transformed into an equivalent function of the form:

$$x \in K,$$

where in place of "K" we have a constant denoting a class; one may, therefore, consider the latter formula as the most general form of a sentential function with one free variable.

The phrases in (I) and (II) are sometimes replaced by symbolic expressions; we can, for instance, agree to use the following notation for this purpose:

$$\{x : \ldots\}.$$

* * *

*Let us now consider the following statement:

1 *belongs to the set of all numbers* x *such that* $x > 0$,

which can also be written in symbols only:

$$1 \in \{x : x > 0\}.$$

What we have here is obviously a sentence, in fact a true sentence; it expresses, in a more complicated form, the same thought as the simple formula:

$$1 > 0.$$

Consequently, the preceding expressions (involving "$x > 0$") cannot contain any free variables, and the variable "x" which occurs there must be a bound variable. On the other hand, since we do not find in the above expressions any quantifiers, we arrive at the conclusion that such phrases as the one in (I) or in (II) function like quantifiers, insofar as they bind variables, and must therefore be classified as operators (cf. Section 4).

It should be added that we frequently prefix an operator like (I) or (II) to sentential functions which contain—besides "x"—other free variables (this occurs in nearly all cases in which such operators are applied in geometry). The expressions which are obtained in this way, for instance:

the set of all numbers x *such that* $x > y$

do not designate any definite classes; rather, they are designatory functions in the sense established in Section 2, that is, they become designations of classes if we replace the free variables contained in them (but of course, not "x") by suitable constants, for instance "y" by "0" in the example just given.*

* * *

It is frequently said that a sentential function with one free variable expresses a certain property of objects,—a property pertaining to those, and only those, objects which satisfy the sentential function in question (the sentential function "*x is divisible by* 2", for example, expresses a certain property, namely divisibility by 2, or, the property of being even). The class corresponding to this function contains as its elements all objects possessing the given property, and no others. In this manner it is possible to correlate a uniquely determined class with every property of objects. And also, conversely, with every class there is correlated a property which is possessed exclusively by the elements of that class; it is simply the property of being an element of that class. Accordingly, in the opinion of numerous logicians, it is unnecessary to distinguish at all between the concept of a class and that of a property; in other words, a special "theory of properties" is dispensable,—the theory of classes being entirely capable of taking the place of the former.

As an application of these remarks, we shall give a new formulation of LEIBNIZ's law. The original one (in Section 17) contained the term "*property*"; in the following formulation, which is completely equivalent, we employ the term "*class*" instead:

$x = y$ *if, and only if, every class which contains one of the objects* x *and and* y *as element also contains the other as element.*

As this formulation of LEIBNIZ's law shows, it is possible to define the concept of identity by employing the terms of the theory of classes.

23 The universal class and the null class

As we already know, to any sentential function with one free variable there corresponds the class of all objects satisfying this function. This principle will now be applied to the following two particular functions:

(I) $$x = x, \qquad x \neq x.$$

The first of these functions is obviously satisfied by every individual (cf. Section 17). The corresponding class,

$$\{x : x = x\},$$

therefore contains as elements all individuals; we call this class the UNIVERSAL CLASS and denote it by the symbol "**V**" (or by "1"). The second sentential function, on the other hand, is satisfied by nothing. Consequently, the class corresponding to it,

$$\{x : x \neq x\},$$

called the NULL CLASS or the EMPTY CLASS, and denoted by "$\emptyset$" (or by "0"), contains no elements. Accordingly, we may replace the sentential functions in (I) by two equivalent sentential functions of the form:

$$x \in K,$$

namely by:

(II) $$x \in \mathbf{V}, \qquad x \in \emptyset,$$

the first of which is satisfied by every individual, and the second by none.

Instead of dealing with all individuals and specifying which of them fall within the framework of a particular mathematical theory, it can be more convenient to restrict the consideration from the beginning to the class of those individuals which the theory in question involves; such a class will be denoted again by "**V**" and will be called the UNIVERSE OF DISCOURSE of the theory. In arithmetic, for instance, it is the class of all numbers which forms the universe of discourse.[†]

* * *

*It should be emphasized that **V** is the class of all individuals but not the class containing as elements all possible objects, among which are classes of the first order, of the second order, and so on. The question arises whether such a class of all possible objects exists at all, and more generally, whether we may consider "inhomogeneous" classes not belonging to a particular order, and containing as elements individuals as well as classes of various orders. This question is closely related to some of the intricate subjects of contemporary logic, namely, to the so-called ANTINOMY OF RUSSELL and to the THEORY OF LOGICAL TYPES.[2] [††] A discussion of this question would

[†]The important notion of universe of discourse was introduced into logic by DE MORGAN (see footnote 6 on p. 46).—On the other hand, the idea of the null class goes back to earlier times. This idea was brought out already in medieval logic and philosophy, in the 13th and 14th centuries, when linguistic expressions were extensively studied and analyzed. (Cf. in this connection Section 18.) In particular, one considered expressions which name non-existent entities, and which therefore (as we could say today) determine the null class.

[2]The concept of logical type, which was introduced by RUSSELL, is akin to that of the order of a class, and in fact, can be conceived as a generalization of the latter,—a generalization which refers not only to classes but also to other objects, for instance to relations (these will be considered in the next chapter) . The theory of logical types was systematically developed in *Principia Mathematica* (cf. footnote 1 on p. 18).

[††]There are alternate ways of setting up a theory of sets (or of classes), where the theory of logical types is bypassed, and where the problems which have just been

go beyond the intended limits of this book. We shall only remark here that the need for considering "inhomogeneous" classes occurs hardly ever in the whole of mathematics (except for the general theory of sets), and even more rarely in other sciences.*

24 The fundamental relations among classes

Various relations may hold between two classes K and L. It may happen, for instance, that every element of the class K is at the same time an element of the class L, in which case the class K is said to be a SUBCLASS OF THE CLASS L, or, to be INCLUDED IN THE CLASS L, or, to HAVE THE RELATION OF INCLUSION TO THE CLASS L; and the class L is said to COMPREHEND or INCLUDE or CONTAIN THE CLASS K AS A SUBCLASS (and we say, under analogous circumstances, that a given set K is a SUBSET of the set L, etc.). This situation is expressed, briefly, by either of the formulas:

$$K \subseteq L, \quad L \supseteq K.$$

By saying that K is a subclass of L, one does not preclude the possibility of L also being a subclass of K. In other words, K and L may be subclasses of each other and thus have all their elements in common; in this case it follows from a law (given below) of the theory of classes that K and L are identical. However, if the converse relation does not hold, so that every element of the class K is an element of the class L, but not every element of the class L is an element of the class K, then the class K is said to be a PROPER SUBCLASS or a PART OF THE CLASS L, and L is said to COMPREHEND K AS A PROPER SUBCLASS or AS A PART. For example, the set of all integers is a proper subset of the set of all rational numbers; a line comprehends each of its segments as a part.†

Two classes K and L are said to OVERLAP or to INTERSECT if they have at least one element in common and if, at the same time, each contains elements which are not contained in the other. If two classes have at least one element each (that is, if they are not empty), but if they have no element in common, they are then called MUTUALLY EXCLUSIVE or DISJOINT. A

---

raised appear in a different perspective. Such alternatives are preferred nowadays, and in fact, logical types are only rarely considered in introductory textbooks.

†The following notation is sometimes used for stating that the class K is a proper subclass of L:

$$K \subset L \quad \text{or} \quad L \supset K.$$

The reader may observe some analogies between the two formulas, "$K \subset L$" and "$x < y$" (the latter referring to numbers), and also between "$K \subseteq L$" and "$x \leq y$". However, the analogy is imperfect: Various formulas for inequalities among numbers do not carry over to classes.

circle, for instance, intersects any straight line drawn through its center, but it is disjoint from any straight line whose distance from the center is greater than the radius. The set of all positive numbers and the set of all rational numbers overlap, but the set of positive and the set of negative numbers are mutually exclusive.

* * *

Let us give some examples of laws which involve the foregoing relations among classes.

For any class K, $K \subseteq K$.

If $K \subseteq L$ and $L \subseteq K$, then $K = L$.

If $K \subseteq L$ and $L \subseteq M$, then $K \subseteq M$.

If K is a non-empty subclass of L, and if the classes L and M are disjoint, then the classes K and M are disjoint.

The first of these statements is called the LAW OF REFLEXIVITY for inclusion or the class-theoretical LAW OF IDENTITY. The third is known as the LAW OF TRANSITIVITY for inclusion; together with the fourth and with others of a similar structure, they form a group of statements which are called LAWS OF THE CATEGORICAL SYLLOGISM.

A characteristic property of the universal and the null classes relates to the concept of inclusion, and is expressed in the following law:

For any class K, $\mathbf{V} \supseteq K$ and $\emptyset \subseteq K$.

This statement, particularly in view of its second part referring to the null class, seems to many people somewhat paradoxical. In order to demonstrate the second part, let us consider the implication:

if $x \in \emptyset$, then $x \in K$.

Whatever object we substitute here for "x", and whatever class for "K", the antecedent of the implication will be a false sentence, and hence the whole implication a true sentence (the implication—as mathematicians sometimes say—is satisfied "vacuously"). We may therefore say, that whatever is an element of the class $\emptyset$ is also an element of the class K, and hence, by the definition of inclusion, that $\emptyset \subseteq K$.—In an analogous way the first part of the law can be demonstrated.

It is easy to see that between any two classes one, and only one, of the relations which we have considered has to hold; the following law provides a detailed statement of this conclusion:

For any classes K and L, either $K = L$, or K is a proper subclass of L, or K comprehends L as a proper subclass, or K and L overlap, or finally K and L are disjoint; no two of these relations can hold simultaneously.

In order to get a clear intuitive understanding of this law, it is best to think of the classes K and L as geometrical figures, and to imagine all possible ways in which one of these two figures may be arranged with respect to the other.

* * *

The relations which have been dealt with in this section may be called the FUNDAMENTAL RELATIONS AMONG CLASSES.[3]

* * *

The old traditional logic (cf. Section 6) can be reduced almost entirely to the theory of the fundamental relations among classes, and therefore, to a small fragment of the theory of classes. Outwardly these two disciplines differ by the fact that, in the old logic, the concept of a class does not appear explicitly. Instead of saying, for instance, that the class of horses is contained in the class of mammals, one used to say in the old logic that the property of being a mammal belongs to all horses, or simply, that every horse is a mammal. The most important laws of traditional logic are those of the categorical syllogism; they correspond precisely to laws of the theory of classes which are like the two which we stated above and named accordingly. For example, the first of the foregoing laws of syllogism assumes the following form in the old logic:

If every M is P and every S is M, then every S is P.

This is the most famous of the laws of traditional logic, and is known as the law of the syllogism BARBARA.

25 Operations on classes

We shall now concern ourselves with certain operations which, when performed on given classes, yield new classes.

Given any two classes K and L, one can form a new class M which contains as elements those, and only those, objects which belong to at least one of the classes K and L; the class M, one might say, results from the class K by adjoining to it the elements of the class L. This operation is called ADDITION OF CLASSES, and the class M is referred to as the SUM or UNION OF THE CLASSES K AND L and is denoted by the symbol:

$$K \cup L \text{ (or } K + L).$$

[3]These relations were first investigated in an exhaustive manner by the French mathematician J. D. GERGONNE (1771–1859).

Another operation on two classes K and L, called MULTIPLICATION OF CLASSES, depends on forming a new class M whose elements are those, and only those, objects which belong to both K and L; this class M is called the PRODUCT or INTERSECTION OF THE CLASSES K AND L, and is designated by the symbol:

$$K \cap L \text{ (or } K \cdot L).$$

These two operations are frequently applied in geometry; sometimes it is very convenient to define with their help new kinds of geometrical figures. Suppose, for instance, that we know already what is meant by a pair of supplementary angles; then a half-plane—or, a straight angle—may be defined as the union of two supplementary angles (an angle here being considered as the corresponding angular region, that is, as the part of the plane which is bounded by the two half-lines, called the legs of the angle). Or, if we take an arbitrary circle and an angle whose vertex lies at the center of the circle, then the intersection of these two figures is a figure which is called a circular sector.

Let us add two examples from the field of arithmetic: the sum of the set of all positive numbers and the set of all negative numbers is the set of all numbers different from 0; the intersection of the set of all even numbers and the set of all prime numbers is the set having as its sole element the number 2, this number being the only even prime number.[†]

* * *

The addition and the multiplication of classes are governed by various laws. Some of these are entirely analogous to certain laws of arithmetic, which relate to addition and multiplication of numbers—and it is just for this reason that the terms "*addition*" and "*multiplication*" have been chosen for the above operations; as examples we mention the COMMUTATIVE and the ASSOCIATIVE LAWS for addition and multiplication of classes:

For any classes K and L, $K \cup L = L \cup K$ *and* $K \cap L = L \cap K$.

For any classes K, L, and M, $K \cup (L \cup M) = (K \cup L) \cup M$ *and* $K \cap (L \cap M) = (K \cap L) \cap M$.

The analogy with the corresponding laws of arithmetic becomes striking when we replace the symbols "$\cup$" and "$\cap$" by the usual signs of addition and multiplication, "+" and "$\cdot$".

Other laws, however, deviate considerably from those of arithmetic; the LAW OF TAUTOLOGY constitutes a characteristic example:

For any class K, $K \cup K = K$ *and* $K \cap K = K$.

[†] We recall: A prime number is a natural number (cf. footnote on p. 46) which is greater than one and whose only divisors are one and the number itself.

This law becomes obvious when one reflects upon the meaning of the combinations "$K \cup K$" and "$K \cap K$"; for instance, if one augments the class K by adjoining the elements of the same class, one does not really add anything, and the resulting class is again K.

* * *

We want to mention one other operation, which can be performed on one class rather than on two, and which differs in this respect from addition and from multiplication. If a class K is given, this operation consists in forming the so-called COMPLEMENT OF THE CLASS K, which is the class of all objects not belonging to K; the complement of the class K is denoted by:

$$K'.$$

For instance, if K is the set of all integers, and our universe of discourse is the set of real numbers, then all proper fractions and all irrational numbers belong to the set K'.

As examples of laws which govern the concept of complement and establish its connection with entities considered earlier, we give the following two statements:

For any class K, $K \cup K' = \mathbf{V}$.

For any class K, $K \cap K' = \emptyset$.

The first of these is called the class-theoretical LAW OF EXCLUDED MIDDLE, and the second, the class-theoretical LAW OF CONTRADICTION.[††]

* * *

The relations among classes and the operations on classes with which we have become acquainted, and also the concepts of the universal class and

[††]These two laws correspond directly to the two laws of sentential calculus with analogous names: "$p \vee \sim p$", "$\sim (p \wedge p)$". One can easily verify, in fact, that the above class-theoretical laws are direct consequences of the two laws just stated. (See also exercise *14.)—It may be opportune to return at this point to a former topic, namely, to the general considerations about sentences in Chapter II. A fundamental hypothesis which is implicit in those discussions is that a given sentence has to be either true or false. One could therefore say, that these two laws of sentential calculus bring us back to the original premisses. Now, examples of sentences have certainly been considered where the truth and the falsity appear indeterminate; for instance, sentences relating to the future might be of this kind. Then the validity of the law of excluded middle would likewise be questioned.—It is of interest in this connection, that variants to the true-false alternative have been considered not only in philosophical deliberations, but also in logical theories. To be more precise, the usual sentential calculus was then modified so as to admit three or more truth values (cf. Section 13; such theories were first suggested by ŁUKASIEWICZ, mentioned in footnote 2 on p. 18). Such theories could be extended, in principle, to include e.g. the calculus of classes; then various class-theoretical laws, such as the class-theoretical law of excluded midddle, would have to be modified as well. In practice, however, the true-false alternative remains the standard basis of logic and mathematics.

the null class, are treated in a special branch of the theory of classes; since the laws which apply to these relations and operations tend to have the character of simple formulas reminiscent of those of arithmetic, this branch of the theory is known as the ALGEBRA OF CLASSES and as the CALCULUS OF CLASSES.

26 Equinumerous classes, the cardinal number of a class, finite and infinite classes; arithmetic as a part of logic

*Among the remaining topics which form the subject matter of the theory of classes, there is one which deserves particular attention and which comprises such concepts as equinumerous classes, the cardinal number of a class, finite and infinite classes. These concepts are, unfortunately, rather involved and can be only superficially discussed here.

* * *

As an example of two EQUINUMEROUS OR EQUIVALENT CLASSES, we may consider the set of the fingers of the right hand and that of the left; these two sets are equinumerous, because it is possible to pair off the fingers of both hands in such a manner that (i) every finger belongs to just one pair, and (ii) every pair contains just one finger of the left hand and just one finger of the right hand. In a similar way, the following three sets are equinumerous: the set of all vertices, the set of all sides, and the set of all angles of a given polygon. Later, in Section 33, we shall be able to give a precise and general definition of this concept of equinumerous classes.

Now let us consider an arbitrary class K; there certainly exists a property belonging to all classes equinumerous to K and to no other classes, namely, the property of being equinumerous with K; this property is called the CARDINAL NUMBER OF THE CLASS K, or, its NUMBER OF ELEMENTS, or, its POWER. This can also be expressed more briefly and more precisely, though perhaps in a somewhat abstract manner: the cardinal number of a class K is the class of all classes equinumerous with K (where we should add the provision: the cardinal number is a second-order class, whose elements are first-order classes). This means, in particular, that two (first-order) classes K and L have the same cardinal number if, and only if, they are equinumerous.

With regard to the number of their elements, classes are classified into finite and infinite ones. Among the former, we identify the classes which consist of no elements, of one element, of two, of three elements, and so on. The new terms are most easily definable on the basis of arithmetic. Indeed, let n be an arbitrary natural number (that is, a non-negative integer);

then we shall say that THE CLASS K CONSISTS OF n ELEMENTS, if K is equinumerous with the class of all natural numbers less than n. In particular, a class consists of 2 elements if it is equinumerous with the class of all natural numbers less than 2, that is, with the class consisting of the numbers 0 and 1. Similarly, a class consists of 3 elements if it is equinumerous with the class containing the numbers 0, 1, and 2 as elements. In general, we shall call a class K FINITE if there exists a natural number n such that the class K consists of n elements, otherwise INFINITE.

It has been recognized, however, that there is an alternative possibility. All the terms which have just been considered can be defined in purely logical terms, without resorting at all to any expressions belonging to the field of arithmetic. For instance, we may say that the class K consists of one element, if this class satisfies the following two conditions: (i) there is an x such that $x \in K$; (ii) for any y and z, if $y \in K$ and $z \in K$, then $y = z$ (these two conditions may also be replaced by a single one: "*there is exactly one x such that $x \in K$*"; cf. Section 20).† In an analogous way, we can define the phrases: "*the class K consists of two elements*", "*the class K consists of three elements*", and so on. The problem becomes much more difficult when we turn to the question of defining the terms "*finite class*" and "*infinite class*"; but also in these cases one has succeeded in resolving the problem positively (cf. Section 33), and as a result, all the concepts under consideration have been brought within the scope of logic.

This circumstance has a most interesting consequence of far-reaching importance; for it turns out that the notion of number itself, and all other arithmetical concepts as well, are definable within the field of logic. It is, indeed, easy to establish the meaning of symbols designating individual natural numbers, such as "0", "1", "2", and so on. The number 1, for instance, can be defined as the number of elements—or, as the cardinal number—of a class which consists of one element. (A definition of this kind seems to be incorrect and apparently contains a vicious circle, since the word "*one*", which is to be defined, occurs in the definiens; but actually no error is committed, because the phrase "*consists of one element*" is considered as a whole and its meaning has already been established.) Nor is it hard to define the general concept of a natural number: a natural number is the cardinal number of a finite class. We are, moreover, in a position to define all operations on natural numbers, and to extend the concept of number by the introduction of fractions, negative numbers, and

†We may remark at this time, that here as elsewhere one has to distinguish between an object and the one-element class whose element is the given object; such one-element classes are sometimes called SINGLETONS. Note, in particular, that the set of all individual objects is the universe of discourse, while the set of corresponding singletons is a second-order set which we may designate here by the phrase: "*cardinal-number one*".

irrational numbers, without having to go beyond the limits of logic at any point. Furthermore, it is possible to prove all theorems of arithmetic on the basis of laws of logic alone (with the provision that the system of logical laws must first be enriched by the inclusion of a statement which is intuitively less evident than others, that is, by the inclusion of the so-called AXIOM OF INFINITY, which states that there are infinitely many distinct objects[††]). This entire construction is very abstract, it cannot easily be popularized, and it does not fit into the framework of an elementary presentation of arithmetic; in this book we also do not attempt to adapt ourselves to this conception, and we treat numbers as individuals and not as properties or classes of classes. *But the mere fact that it has been possible to develop the whole of arithmetic, including the disciplines erected upon it—algebra, calculus and analysis, and so on—, as a part of pure logic, constitutes one of the most remarkable achievements of logical investigations.*[4*]

Exercises

1. Let K be the set of all numbers less than $\frac{3}{4}$; which of the following formulas are true:

$$0 \in K, \quad 1 \in K, \quad \tfrac{2}{3} \in K, \quad \tfrac{3}{4} \in K, \quad \tfrac{4}{5} \in K\ ?$$

2. Consider the following four sets:

(a) the set of all positive numbers,
(b) the set of all numbers less than 3,
(c) the set of all numbers x such that $x + 5 < 8$,
(d) the set of all numbers x satisfying the sentential function "$x < 2x$".

Which of these sets are identical, and which are distinct?

3. What name is given in geometry to the set of all points in space whose distance from a given point (or from a given straight line) does not exceed the length of a given line segment?

[††]A more precise way of stating the axiom of infinity is as follows: *There exists an infinite set.* In other words, not only do we need infinitely many distinct objects, but we must also be able to combine them into a set (which, in turn, will obey the laws of the general theory of sets).

[4]The fundamental ideas in this field are due to FREGE (cf. footnote 2 on p. 18), who developed them for the first time in his interesting book: *Die Grundlagen der Arithmetik* (Breslau 1884). FREGE's ideas found their systematic and exhaustive realization in WHITEHEAD and RUSSELL's *Principia Mathematica* (cf. footnote 1 on p. 18).

4. Let K and L be two concentric circles, the radius of the first being smaller than that of the second. Which of the relations discussed in Section 24 holds between these circles? Does the same relation hold between the circumferences of the circles?

5. For each of the following conditions, draw two squares K and L in such a way that the condition be satisfied:

(a) $K = L$,
(b) the square K is a part of the square L,
(c) the square K comprehends the square L as a part,
(d) the squares K and L overlap,
(e) the squares K and L are disjoint.

Which of these cases are eliminated (i) if the squares have to be congruent, or (ii) if not the squares but only their perimeters are considered?

6. Let x and y be two arbitrary numbers, with $x < y$. We recall that the set of numbers which are not smaller than x and not larger than y is called the INTERVAL (more precisely, the CLOSED INTERVAL) with the endpoints x and y; it is denoted by the symbol "$[x, y]$".

Which of the formulas below are correct (that is, true)?

(a) $[3, 5] \subseteq [3, 6]$,
(b) $[4, 7] \subseteq [5, 10]$,
(c) $[-2, 4] \supseteq [-3, 5]$,
(d) $[-7, 1] \supseteq [-5, -2]$.

Which of the fundamental relations holds between the given intervals in each of the following?

(e) $[2, 4]$ and $[5, 8]$,
(f) $[3, 6]$ and $[3\frac{1}{2}, 5\frac{1}{2}]$,
(g) $[1\frac{1}{2}, 7]$ and $[-2, 3\frac{1}{2}]$.

7. Is the following sentence (which has the same structure as the laws of the categorical syllogism, given in Section 24) true:

for any classes K, L, and M,
if K is disjoint from L and L is disjoint from M,
then K is disjoint from M ?

8. Translate the following formulas into ordinary language:

(a) $x = y \leftrightarrow \forall_K (x \in K \leftrightarrow y \in K)$,
(b) $K = L \leftrightarrow \forall_x (x \in K \leftrightarrow x \in L)$.

Which of the laws mentioned in Sections 22 and 24 find their expression in these formulas? What alterations on each side of equivalence (b) would be required in order to arrive at definitions of the symbols "$\subseteq$" and "$\supseteq$"?

9. Let ABC be an arbitrary triangle, and D an arbitrary point lying on the segment BC. What figures are formed by the sum (or, union) of the two triangles ABD and ACD and by their product? Express the answer in formulas.

10. Represent an arbitrary square:

(a) as the sum of two trapezoids,
(b) as the intersection of two triangles.

11. Which of the formulas below are true (compare exercise 6)?

(a) $[2, 3\frac{1}{2}] \cup [3, 5] = [2, 5]$,
(b) $[-1, 2] \cup [0, 3] = [0, 2]$,
(c) $[-2, 8] \cap [3, 7] = [-2, 8]$,
(d) $[2, 4\frac{1}{2}] \cap [3, 5] = [2, 3]$.

In those formulas which are false, correct the expression on the right of the symbol "=".

12. Let K and L be two arbitrary classes. What classes are $K \cup L$ and $K \cap L$ in case $K \subseteq L$? In particular, what classes are $K \cup \mathbf{V}$, $K \cap \mathbf{V}$, $\emptyset \cup L$, and $\emptyset \cap L$?

Hint: In answering the second question, keep in mind a law of Section 24 concerning the classes $\mathbf{V}$ and $\emptyset$.

13. Try to show that any classes K, L, and M satisfy the following formulas:

(a) $K \subseteq K \cup L$ and $K \supseteq K \cap L$,
(b) $K \cap (L \cup M) = (K \cap L) \cup (K \cap M)$ and
$K \cup (L \cap M) = (K \cup L) \cap (K \cup M)$,
(c) $(K')' = K$,
(d) $(K \cup L)' = K' \cap L'$ and $(K \cap L)' = K' \cup L'$.

Formulas (a) are called the LAWS OF SIMPLIFICATION (for addition and multiplication of classes); formulas (b) are the DISTRIBUTIVE LAWS (for multiplication of classes with respect to addition and for addition with respect to multiplication); formula (c) is the LAW OF DOUBLE COMPLEMENT; and, finally, formulas (d) are the class-theoretical LAWS OF DE MORGAN.[5] Which of these laws correspond to laws of arithmetic?

Hint: In order to prove the first of formulas (d), for instance, it suffices to show that the classes $(K \cup L)'$ and $K' \cap L'$ consist entirely of the same elements (cf. Section 24). For this purpose we need the definitions of Section 25, and have to make clear to ourselves when an object x belongs to the class $(K \cup L)'$ and when it belongs to $K' \cap L'$.

[5]Cf. footnote 6 on p. 46.

***14.** Between the laws of sentential calculus given in Sections 12 and 13 and in exercise 14 of Chapter II, on the one hand, and the laws of the algebra of classes given in Sections 24 and 25 and in the preceding exercise, on the other, there subsists a far-reaching similarity in structure (which is also indicated by the analogy in their names). Describe in detail wherein this similarity lies, and try to find a general explanation of this phenomenon.

In Section 14 we became acquainted with the law of contraposition of sentential calculus; formulate the analogous law of the algebra of classes.

15. Using the notation:

$$\{x: \ldots\}$$

introduced in Section 22, we can write the definition of the sum of two classes in the following way:

$$K \cup L = \{x : x \in K \vee x \in L\};$$

but it is also possible to restate this definition in the usual form of an equivalence (without the use of that notation):

$$[x \in (K \cup L)] \leftrightarrow [x \in K \vee x \in L].$$

Formulate analogously in two ways the definitions of the universal class, the null class, the product of two classes, and the complement of a class.

16. The difference of two given classes K and L, in symbols: $K - L$, is the class M defined by the formula:

$$M = \{x : x \in K \wedge \sim (x \in L)\}.$$

If K is the class of all integers and L is the class of all negative integers, what are $K - L$ and $L - K$? Which of the following are satisfied by all classes:

$$(K - L) \cup L = K \cup L, \quad K - (K - L) = L,$$
$$\text{and } L - (K - L) = L?$$

***17.** Is there a polygon in which the set of all sides is equinumerous with the set of all diagonals?

***18.** John works on Mondays, Wednesdays, and Fridays while Peter works on Mondays, Tuesdays, Wednesdays, and Thursdays. Let K be the set of days of the week when John works and L the set when Peter works. What percentage is the number of elements in $K \cap L$ of the number of elements in $K \cup L$? Answer the same question after replacing $K \cap L$ by $K - L$ (see exercise 16).

***19.** For each of the following two expressions, write an equivalent formula using only logical symbols:

(a) *the class K consists of two elements,*
(b) *the class K consists of three elements.*

***20.** Which of the following sets are finite and which are infinite?

(a) the set of all natural numbers x such that $0 < x$ and $x < 4$,
(b) the set of all rational numbers x such that $0 < x$ and $x < 4$,
(c) the set of all irrational numbers x such that $0 < x$ and $x < 4$.

21. For each of the following expressions state whether it is a sentence, a sentential function, a designation, or a designatory function, and specify which variables occur free and which occur bound:

(a) $\{x : x^2 = 9 \wedge 0 < x\}$,
(b) $\forall_{K,L}(K \subseteq L' \rightarrow \exists_z[z \in K \wedge \sim (z \in L)])$,
(c) $K \cup \{y : y \in L \wedge \sim (y \in K)\}$,
(d) $9 \in \{x : \exists_y(x = y^2)\}$.

V

On the Theory of Relations

27 Relations, their domains and counter-domains; relations and sentential functions with two free variables

Already in the previous chapters we came upon a few RELATIONS among objects. As examples of a relation involving two objects, we may take identity (equality) and diversity. Indeed, sometimes we read the formula:

$$x = y$$

as follows:

x has the relation of identity to y

or else:

the relation of identity holds between x and y,

and we say that the symbol "$=$" designates the relation of identity. In an analogous way, the formula:

$$x \neq y$$

is sometimes read:

x has the relation of diversity to y

or:

the relation of diversity holds between x and y,

and one says that the symbol "$\neq$" designates the relation of diversity. We have encountered, moreover, certain other relations which hold among classes, in particular the relations of inclusion, overlapping, and disjointness. We shall now discuss several concepts belonging to the general THEORY OF RELATIONS, which constitutes a separate and very important branch of logic; in this theory, relations of arbitrary character are considered, and general laws concerning them are established.[1]

To facilitate our considerations, we introduce special variables: "R", "S", ... which serve to denote relations. In place of such phrases as:

the object x has the relation R to the object y

and:

the object x does not have the relation R to the object y

we shall employ symbolic abbreviations:

$$xRy$$

and (to use the negation sign of sentential calculus, cf. Section 13)

$$\sim (xRy),$$

respectively; we shall also use abbreviative expressions like the following: "*the relation R holds between x and y*".

If an object x has the relation R to an object y, we call x a PREDECESSOR WITH RESPECT TO THE RELATION R; any object y for which there is an object x such that

$$xRy$$

is called a SUCCESSOR WITH RESPECT TO THE RELATION R. The class of all predecessors with respect to the relation R is known as the DOMAIN, and the class of all successors, as the COUNTER-DOMAIN (or, the CONVERSE DOMAIN) OF THE RELATION R. For example, every individual is both a predecessor and a successor with respect to the relation of identity, so that the domain as well as the counter-domain of this relation is the universal class.

* * *

In the theory of relations—just as in the theory of classes—we may distinguish among relations of different orders. The RELATIONS OF THE FIRST ORDER are those which hold among individuals; the RELATIONS OF THE SECOND ORDER are those which hold among objects of the first order, whether classes or relations; and so on. Here the situation is somewhat more complicated since often we have to consider a "mixed" relation, whose predecessors are, say, individuals and whose successors are classes, or such that its predecessors are, for instance, classes of the first order and its

---

[1] DE MORGAN and PEIRCE (cf. footnotes 6 on p. 46 and 2 on p. 12) were the first to develop the theory of relations, especially that part of it which is known as the algebra of relations (cf. Section 28). Their work was systematically expanded and completed by the German logician E. SCHRÖDER (1841–1902). SCHRÖDER's *Algebra und Logik der Relative* (Leipzig 1895), which appeared as the third volume of his comprehensive work *Vorlesungen über die Algebra der Logik*, is still the only exhaustive account of the algebra of relations.

successors, classes of the second order. The most important example of a relation of this kind, is the relation which holds between an element and a class to which the element belongs; as we recall from Section 21, this relation is denoted by the symbol "$\in$".—As was the case with classes, our considerations will refer primarily to relations of the first order, although the concepts which are discussed here can also be applied to relations of higher orders, and in a few cases will be so applied.

* * *

We assume that, to every sentential function with two free variables "x" and "y", there corresponds a relation which a given object x has to a given object y if, and only if, x and y satisfy the given sentential function; one therefore says that a sentential function with two free variables, say "x" and "y", expresses a relation. For instance, the sentential function:

$$x + y = 0$$

expresses the relation of having the opposite sign or, briefly, of being opposite; that is, the numbers x and y have the relation of being opposite if, and only if, $x + y = 0$. If we denote this relation by the symbol "$\mathcal{O}$", then the formulas:

$$x\mathcal{O}y$$

and

$$x + y = 0$$

are equivalent. In a similar way, if we have any sentential function containing the symbols "x" and "y" as the only free variables, we can refer to the corresponding relation; the sentential function may then be transformed into an equivalent formula of the form:

$$xRy$$

where, in place of "R", we have a constant which designates the corresponding relation. The formula:

$$xRy$$

may therefore be considered as the general form of a sentential function with two free variables, just as the formula:

$$x \in K$$

could be looked upon as the general form of a sentential function with one free variable (cf. Section 22).

28 The algebra of relations

The theory of relations is an extensively developed branch of mathematical logic. One part of it, the ALGEBRA (or, the CALCULUS) OF RELATIONS, is akin to the algebra of classes, and its principal aim is to establish formal laws governing the operations by means of which new relations can be constructed from given ones.

* * *

In the algebra of relations we consider, in the first place, a group of concepts which are the exact analogues of those of the algebra of classes; they are usually denoted by the same symbols and are governed by analogous laws. (In order to avoid ambiguity we might, of course, employ a different set of symbols in the algebra of relations; for instance, we might take the symbols of the algebra of classes and place a dot over each.)

We have thus in the algebra of relations two particular relations, the UNIVERSAL RELATION $\mathbf{V}$ and the NULL RELATION $\emptyset$, the first of which holds between any two individuals, and the second, between none.

We have, moreover, various relations among relations, for instance, the RELATION OF INCLUSION; we say that the relation R is INCLUDED in the relation S, in symbols:

$$R \subseteq S,$$

if, whenever R holds between two objects, S holds between them likewise; in other words, if for any x and y, the formula:

$$xRy$$

implies:

$$xSy.$$

For instance, we know from arithmetic that, whenever

$$x < y,$$

then

$$x \neq y;$$

hence the relation of being smaller is included in the relation of diversity.

If, at the same time,

$$R \subseteq S \text{ and } S \subseteq R,$$

that is, if the relations R and S hold for the same pairs of objects, then they are identical:

$$R = S.$$

We have, further, the SUM or UNION OF TWO RELATIONS R AND S, in symbols:

$$R \cup S,$$

and the PRODUCT or INTERSECTION OF R AND S, in symbols:

$$R \cap S.$$

The first, $R \cup S$, holds between two objects if, and only if, at least one of the relations R and S holds between them; in other words, the formula:

$$x(R \cup S)y$$

is equivalent to the condition:

$$xRy \quad or \quad xSy.$$

The product of two relations is defined similarly, by using the word "*and*" instead of "*or*". For example, if R is the relation of fatherhood (that is, the relation which holds between two persons x and y if, and only if, x is the father of y), and S the relation of motherhood, then $R \cup S$ is the relation of parenthood, while $R \cap S$ is, in this case, the null relation.

We have, finally, the NEGATION or COMPLEMENT OF A RELATION R, which is denoted by:

$$R'.$$

It is the relation which holds between two objects if, and only if, the relation R does not hold between them; in other words, for any x and y, the formulas:

$$xR'y \text{ and } \sim(xRy)$$

are equivalent. It should be noted that, if a relation is designated by a constant symbol, then its complement is frequently denoted by the symbol which is obtained from the former by crossing it with a vertical or oblique bar. The negation of the relation $<$, for example, is usually denoted by "$\nmid\!<$" or by "$\not<$" (and not by "$<'$").

* * *

In the algebra of relations, however, there occur entirely new concepts, which have no analogues in the algebra of classes.

We have here, first, two particular relations: IDENTITY and DIVERSITY among individuals (which, incidentally, are familiar to us from earlier considerations).[†] In the algebra of relations they are denoted by special symbols or letters, e.g. by "$\mathcal{I}$" and "$\mathcal{D}$", and not by the symbols "$=$" and "$\neq$" which are used elsewhere in logic. We write, thus:

$$x\mathcal{I}y \text{ and } x\mathcal{D}y$$

[†] We should like to comment on the terms which have just been mentioned. We consider the relation of *identity* and that of *equality* as the same, in accordance with the discussion of Chapter III. With regard to the complementary relation, the usual term is "*diversity*", as in the text. One must, in particular, be careful to distinguish between the concepts of *diversity* and *inequality*; the latter will be introduced later (in Section 46), and it involves specifically arithmetical entities.

instead of:

$$x = y \text{ and } x \neq y.$$

The symbols "=" and "≠" are used in the algebra of relations only to denote the identity and the diversity among relations.

We next have a very interesting and important new operation, with the help of which we form, from two relations R and S, a third relation called the RELATIVE PRODUCT OR COMPOSITION OF R AND S (the ordinary product sometimes is opposed to it and called the ABSOLUTE PRODUCT). The relative product of R and S is denoted by the symbol:

$$R \mid S;$$

it holds between two objects x and y if, and only if, there exists an object z such that we have at the same time:

$$xRz \text{ and } zSy.$$

For instance, if R is the relation of being husband and S the relation of being daughter, then $R \mid S$ holds between two persons x and y if there is a person z such that x is husband of z and z is daughter of y; the relation $R|S$ therefore coincides with the relation of being son-in-law.—There is, in addition, another operation of a similar character, whose result is called the RELATIVE SUM OF TWO RELATIONS. This operation does not play a great role and will not be defined here.

Finally, we have an operation which involves only one relation R, and in this respect it is analogous to the forming of R'; with its help we obtain a new relation, called the CONVERSE OF R and denoted by:

$$\breve{R}.$$

The relation $\breve{R}$ holds between x and y if, and only if, R holds between y and x. If a relation is denoted by a constant symbol, then for denoting its converse we often employ the same symbol but printed in the opposite direction. For instance, the converse of the relation $<$ is the relation $>$, since, for any x and y, the formulas:

$$x < y \text{ and } y > x$$

are equivalent.[††]

* * *

In view of the rather specialized character of the algebra of relations, we shall not go here into further details.

[††]The two new operations are useful when one wishes to describe the properties of particular kinds of relations, or of definite relations. (Cf. in particular exercises 16 and *17, also 11 and *12 of Chapter VII, as well as Section 45.) One can, furthermore, easily find some general laws which involve these operations (cf. exercise 4).

29 Several kinds of relations

We now turn to that part of the theory of relations whose task is to describe and investigate various kinds of relations (and the associated special properties), especially those kinds of relations which are frequently met in mathematics or in other sciences.

We shall call a relation R REFLEXIVE IN THE CLASS K, if every element x of the class K has the relation R to itself:

$$xRx;$$

if, on the other hand, no element of this class has the relation R to itself:

$$\sim (xRx),$$

then the relation R is said to be IRREFLEXIVE IN THE CLASS K. The relation R is classified as SYMMETRIC IN THE CLASS K if, for any two elements x and y of the class K, the formula:

$$xRy$$

always implies that:

$$yRx.$$

If, however, the formula:

$$xRy$$

always implies:

$$\sim (yRx),$$

then the relation R is said to be ASYMMETRIC IN THE CLASS K. The relation R is classified as TRANSITIVE IN THE CLASS K if, for any three elements x, y, and z of the class K, the conditions:

$$xRy \text{ and } yRz$$

always imply:

$$xRz.$$

If, finally, for any two distinct elements x and y of the class K, at least one of the formulas:

$$xRy \text{ and } yRx$$

holds, that is, if the relation R subsists between two arbitrary distinct elements of K in at least one direction, the relation is said to be CONNECTED IN THE CLASS K.

In case K is the universal class (or, at any rate, the universe of discourse of the science in which we happen to be interested—cf. Section 23) we usually speak, not of relations which are reflexive in the class K, symmetric in the class K, and so on, but simply of reflexive relations, symmetric relations, and so on.

30 Relations which are reflexive, symmetric, and transitive

The properties of relations which we have just described are frequently encountered in combinations. Very common, for instance, are those relations which are reflexive, symmetric, and transitive as well. A familiar example of this type is the relation of identity; Law II of Section 17 states that this relation is reflexive, by Law III identity is a symmetric relation, and according to Law IV it is transitive (and in view of this circumstance these laws were given their respective names). Numerous other examples of relations of this kind may be found within the field of geometry. The congruence relation, for instance, is reflexive in the set of all line segments (or of arbitrary geometrical figures), since every segment is congruent to itself; it is symmetric, since, if a segment is congruent to another segment, the other is congruent to the first; and finally, it is transitive, since, if the segment A is congruent to the segment B, and B to C, then the segment A is also congruent to the segment C. The same three properties belong to the relations of similarity among polygons and parallelism among straight lines (assuming any line to be parallel to itself), or—outside the domain of geometry—to the relation of being equally old among people and to that of synonymity among words.

A relation which is at the same time reflexive, symmetric, and transitive is usually thought of as some kind of equality. Instead of saying, therefore, that such a relation holds between two objects, one can also say (in the same sense) that these objects are equal in such and such a respect, or—using a more precise mode of speech—that certain properties of these objects are identical. Thus, instead of stating that two segments are congruent, or two people equally old, or two words synonymous, we may just as well state that the segments are equal with respect to their length, that the people have the same age, or that the meanings of the words are identical.

* * *

*We shall indicate by means of an example how one can establish a logical basis for such a mode of expression. For this purpose let us consider the relation of similarity among polygons. We shall refer to the set of all polygons which are similar to the given polygon P (or, to say differently, to the common property which belongs to all polygons similar to P and to no others, cf. the remarks at the end of Section 22) as the shape of the polygon P. Thus shapes are certain sets of polygons (or, certain properties of polygons). Making use of the fact that the relation of similarity is reflexive, symmetric, and transitive, we can easily show that every polygon belongs to one and only one such set, that two similar polygons always belong to the same set, and that two polygons which are not similar belong to different sets. Consequently the two statements:

the polygons P and Q are similar

and

the polygons P and Q have the same shape (that is, *the shapes of P and Q are identical)*

are equivalent.

The reader will notice immediately that once before our discussion depended upon a procedure of this kind, namely in Section 26, when making the transition from the expression:

the classes K and L are equinumerous

to the equivalent one:

the classes K and L have the same cardinal number.

It can be shown without much difficulty, that such a procedure is applicable to any relation which is reflexive, symmetric, and transitive. There is in fact a logical law, called the PRINCIPLE OF ABSTRACTION, that supplies a general theoretical foundation for the procedure which we have been considering, but here we shall forego a precise formulation of this principle.*

* * *

There is a term which is widely used for denoting relations which are at the same time reflexive, symmetric, and transitive. They are generally called EQUIVALENCE RELATIONS, or for short, EQULVALENCES†; sometimes one refers to such relations as equalities.†† In particular, the term "*equality*" might be occasionally adopted for a certain relation which has the three properties in question, and two objects would then be called equal if

---

†The associated adjective, "*equivalent*", is very common in logical and in mathematical texts. In particular, if a new equivalence relation E has been introduced in a discussion, then two objects are often called equivalent if E holds between them, and by this device one is spared the task of finding a new name. We shall see an example of such usage in Section 37. Sometimes, moreover, "*equivalent*" is used in addition to other terms, as a synonym; compare "*equivalent classes*" in Section 26. (One also speaks of equivalent sentences, but then other concepts are involved; cf. Section 10 and also footnote on pp. 153–154.)

††The following may be noted: If the object x is identical with the object y, then every equivalence relation necessarily holds between x and y, as the reader can easily demonstrate. (We can say this more concisely: *If R is an equivalence relation, then* $\mathcal{I} \subseteq R$.) However, the fact that an equivalence relation holds between two objects does not in general imply their identity. The relation of identity is therefore the most restrictive among equivalence relations.

the relation holds between them. For instance, in geometry, as was pointed out in Section 19, congruent segments are at times referred to as equal segments. We will emphasize here once more that it is preferable to avoid such expressions altogether; their use brings about ambiguities, and it violates the convention according to which the terms "*equality*" and "*identity*" should be considered as synonymous.

31 Ordering relations; examples of other relations

Another very common kind of relation is represented by those which are asymmetric, transitive, and connected in a given class K (they must then also be irreflexive in this class, as can be easily shown). Of a relation with these properties we say that it ESTABLISHES AN ORDER IN THE CLASS K, or that it is an ORDERING RELATION IN THE CLASS K; we say also that the class K is ORDERED BY THE RELATION R. Consider, for example, the relation of being smaller (or, the relation *less than*, as we shall say on occasion); it is asymmetric in any set of numbers, for, if x and y are any two numbers and if

$$x < y,$$

then

$$y \nless x, \quad \text{that is:} \quad \sim (y < x);$$

it is transitive, since the formulas:

$$x < y \text{ and } y < z$$

always imply:

$$x < z;$$

finally, it is connected, since, of any two distinct numbers, one must be smaller than the other (and it is also irreflexive, since no number is smaller than itself). Any set of numbers, therefore, is ordered by the relation of being smaller. Moreover, the relation of being greater is likewise an ordering relation for any set of numbers.[†]

Let us now consider the relation of being older. One can easily verify that this relation is irreflexive, asymmetric, and transitive in any given

[†]Sometimes, when considering an abstract concept, one finds a mathematical entity which adequately characterizes the concept in question. It has become a style of speaking in mathematics, that under such circumstances the entity simply "is" the concept under consideration. For instance, one might say: "the relation of being greater is an order among numbers", even though we tend to envisage this order as a concept which differs from a (binary) relation. The reader may find such a way of speaking somewhat artificial, and may therefore prefer to say that a relation "establishes an order", as in the text.—Similar remarks could be made regarding the concept of shape, which was considered in the previous section.

set of people. However, it is not necessarily connected; for it can happen, perchance, that the set contains two people having exactly the same age, that is to say, who were born at the same moment, so that the relation of being older does not hold between them in either direction. If, on the other hand, we consider a set of people in which no two are of exactly the same age, then the relation of being older establishes an order in that set.

* * *

Many relations are known that are neither ordering relations nor equivalences. Let us consider a few examples.

The relation of diversity is irreflexive in any set of objects, since nothing is different from itself; it is symmetric, for, if

$$x \neq y,$$

then we also have

$$y \neq x;$$

it fails to be transitive, since the formulas:

$$x \neq y \quad \text{and} \quad y \neq z$$

do not imply the formula:

$$x \neq z;$$

on the other hand, it is connected, as can be seen at once.

The relation of inclusion among classes, by the law of identity and one of the laws of syllogism (cf. Section 24), is reflexive and transitive; it is, however, neither symmetric nor asymmetric, since the formula:

$$K \subseteq L$$

neither implies nor excludes the formula:

$$L \subseteq K$$

(these two formulas are satisfied simultaneously if, and only if, the classes K and L are identical); finally, it can be seen with ease that it is not connected. The relation of inclusion therefore differs in its properties from the other relations which were considered above.

32 Many-one relations or functions

We shall now deal in some detail with a category of relations which is particularly important. A relation R is called a MANY-ONE or FUNCTIONAL

RELATION, or simply a FUNCTION, if, to every object x, there corresponds at most one object y such that xRy; in other words, if the formulas:

$$xRy \text{ and } xRz$$

always imply:

$$y = z.$$

The predecessors with respect to the relation R, that is, those objects x for which there actually are objects y such that

$$xRy,$$

are the ARGUMENT VALUES, the successors are the VALUES OF THE FUNCTION R, or simply, the FUNCTION VALUES. Let now R be an arbitrary function, and x, any one of its argument values; let us denote the unique value y of the function corresponding to the value x of the argument by the symbol "$R(x)$"; we may then replace the formula:

$$xRy$$

by:

$$R(x) = y.$$

It has moreover become the custom, especially in mathematics, to use letters such as "f", "g", ... (and not the variables "R", "S", ...) when denoting functional relations, so that we find formulas like these:

$$f(x) = y, \quad g(x) = y, \ldots;$$

the formula:

$$f(x) = y,$$

for instance, is read as follows:

the function f assigns (or, *correlates) to the argument value x the value y.*

One often states such formulas with the function value first. The last formula would thus be written in the equivalent form:

$$y = f(x),$$

and would then be read:

y is that value of the function f which corresponds to (or, *is correlated with) the argument value x.*

* * *

In many of the less recent textbooks of elementary algebra one finds a definition of the concept of a function that is quite different from the definition adopted here. The functional relation is characterized there as a relation between two "variable" quantities or numbers: the "independent variable" and the "dependent variable", which depend upon each other insofar as a change of the first effects a change of the second. Definitions of this kind should no longer be employed today, since they are incapable of standing up to logical criticism; they are the vestiges of a period when one tried to distinguish between "constant" and "variable" quantities (cf. Section 1). A person who desires to comply with the requirements of contemporary science and yet does not wish to break away completely from tradition may, however, retain the old terminology; besides using the terms "*argument value*" and "*function value*", he or she may also refer to the "value of the independent variable" and to the "value of the dependent variable".†

* * *

The simplest example of a functional relation is represented by the familiar relation of identity. As an example of a function from everyday life, let us take the relation which is expressed by the sentential function:

father of x is y.

(This example is like that of Section 28, but now the roles of x and y are interchanged. The phrase just given is awkward but illustrative.) This relation is functional, since, to every person x, there exists but one person y who is father of x. In order to indicate the functional character of this relation, we insert the word "*the*" in the above formulation:

the father of x is y,

instead of which we might also write:

y is the same as the father of x.

Such an alteration of the original expression, involving the insertion of the definite article and further rephrasing, serves in ordinary language the same purpose as the transition (in our symbolism) from the formula:

$$xRy$$

to the formula:

$$R(x) = y$$

†Such a way of speaking, with mention of an independent variable and a dependent variable and changes in each, is also sometimes encountered in introductory courses of calculus. We should therefore like to illustrate by means of an example how the concepts of this section allow a more satisfactory formulation.—Let the function f be determined by the formula: "$y = x^2$". We consider the difference, or change:

$$(x+1)^2 - x^2 = 2x + 1,$$

which we interpret by introducing several new functions: first g, corresponding to "$y = x + 1$", then a combination of f and g which corresponds to "$y = (x+1)^2$". Finally, there is the difference-function, whose function value is $2x + 1$ at x.

and to the formula:

$$y = f(x).$$

We should also note, that the concept of a function plays a very basic role in the mathematical sciences. There are whole branches of higher mathematics which are devoted exclusively to the study of various kinds of functions. But also in elementary mathematics, especially in algebra and trigonometry, we find an abundance of functional relations. As examples we have the relations which are expressed by the following formulas:

$$x + y = 5,$$
$$y = x^2,$$
$$y = \log_{10} x,$$
$$y = \sin x.$$

Let us consider the second of these more closely. To every number x there corresponds only one number y such that $y = x^2$, so that the formula really does represent a functional relation. Argument values of this function are arbitrary numbers, while function values can only be non-negative numbers. By introducing now the symbol "f" for this function, we can cast the equation:

$$y = x^2$$

into the above form:

$$y = f(x).$$

This notation can be used not only for indicating that we have here a functional relation, but also when we refer to definite numbers. Since, for instance,

$$4 = (-2)^2,$$

we may assert that

$$4 = f(-2);$$

the last formula states that 4 is the value of the function f corresponding to the argument value -2.

On the other hand, already in elementary mathematics we encounter numerous relations which are not functions. For example, the relation of being smaller is certainly not a function, since, for every number x, there are infinitely many numbers y such that

$$x < y.$$

Nor is the relation between the numbers x and y which is expressed by the formula:

$$x^2 + y^2 = 25$$

a functional relation, since, to one and the same number x, there may correspond two different numbers y for which the formula is satisfied; corresponding to the number 4, for instance, there are the two numbers 3 and

-3. We may note that such relations among numbers, which are expressed by equations and correlate with one number x two or more numbers y, are sometimes called in mathematics two-valued or many-valued functions (in opposition to single-valued functions, that is, to functions in the ordinary meaning). It seems inexpedient, however—at least on an elementary level—to speak of such relations as functions, for this only tends to blot out the essential difference between the notion of a function and the more general one of a relation.

* * *

Functions are of particular significance in the applications of mathematics to empirical sciences. Whenever we inquire into the dependence between two kinds of quantities occurring in the external world, we strive to put this dependence into the form of a mathematical formula, e.g. an equation, which would permit us to determine the quantity of the one kind when the corresponding quantity of the other kind is given; such an equation then represents a certain functional relation between the quantities of the two kinds. As example let us mention the well-known formula from physics:

$$s = 16.1t^2$$

expressing the dependence of the distance s, which is covered by a freely falling body, upon the time t of its fall (the distance being measured in feet and the time in seconds).

* * *

*In concluding our remarks on functional relations, we want to emphasize that the concept of a function which we are now considering differs in an essential way from the concepts of a sentential and a designatory function which are known to us from Section 2. Strictly speaking, the terms "*sentential function*" and "*designatory function*" do not belong to the domain of logic or of mathematics; they refer to certain categories of expressions which serve to compose logical and mathematical statements, but they do not denote objects which are treated in those statements (cf. Section 9). The term "*function*" in its new sense, on the other hand, is an expression of a purely logical character; it designates an object of a type that one deals with in logic and mathematics. There is, of course, a connection between these concepts, which may be described roughly as follows. If the variable "y" is joined by the symbol "$=$" to a designatory function containing "x" as the only variable, e.g. to "x^2+2x+3", then the resulting formula (which is a sentential function):

$$y = x^2 + 2x + 3$$

expresses a functional relation; to say in other words: the relation, which holds between two numbers x and y if, and only if, they satisfy this formula, is a function in the new sense. This is one of the reasons why these concepts are so often confused.*

33 One-one relations or bijective functions, and one-to-one correspondences

From the totality of functional relations, a special attention deserve the so-called ONE-ONE RELATIONS, also known as BIJECTIVE FUNCTIONS, having not only the property that to every argument value x only one function value y is correlated, but also the converse property, that only one argument value x corresponds to every value y of the function; they might also be characterized as those relations which are functional and whose converses are also functional (while a precise definition of these terms would involve the domains as well; for the concept of converse relation see Section 28).

If f is a bijective function, K an arbitrary class of its argument values, and L the class of function values which are correlated with the elements of K, we say that the function f MAPS THE CLASS K ONTO THE CLASS L IN A ONE-TO-ONE MANNER, or, that it ESTABLISHES A ONE-TO-ONE CORRESPONDENCE BETWEEN THE ELEMENTS OF K AND THOSE OF L.[†]

* * *

We shall give a few examples. Suppose we have a half-line issuing from the point O, with a segment which is marked off and which specifies the unit of length. Let X be any point on the half-line. Then the segment OX can be measured, that is to say, one can correlate with it a certain non-negative number y called the length of the segment. Since this number

[†] The reader may find it useful to acquaint himself with some additional terminology, which is used a great deal in more advanced work. Since we say (as above): "*f maps the class...*", a function is also called a MAP or a MAPPING. The counter-domain of a function is commonly called its RANGE. Consider next the variable "x" whose values can be all elements of a certain set (or, class) K. Then one says "x" RANGES OVER THE SET K. This phrase is commonly used in discussing functional relations as in "$y = f(x)$", and then K is the domain of f, but it can also be used under more general circumstances, e.g. in connection with arbitrary relations. (Phrases like ' *"x" denotes elements of K*' are used in this book in a similar sense. Note also that the meanings of "*the range*" and "*to range*" are not analogous, and should not be confused.)—Now, let the domain of f be K, as before, and let the function values be in a certain set M. I.e., the range could be M itself, but it could also be a proper subset of M. Then we say: f is A MAP FROM K TO M. The following example shows that it is often convenient to make an assertion which would allow either option: Let h be the function defined by the formula "$y = x^n$", where "x" ranges over the set $\mathbf{R}$ of real numbers, and where n is a positive integer which is left unspecified. Then h *is a map from* $\mathbf{R}$ *to* $\mathbf{R}$. If n is odd then the range is $\mathbf{R}$ itself, but if n is even then the range is a proper subset of $\mathbf{R}$.—If the set M containing the function values is also the range, then we say that f, as a map from K to M, is SURJECTIVE or ONTO. (Otherwise we can say that *f maps K into M*.) If a function from K to M is such that every function value corresponds to only one argument value (but the given function need not be surjective), then we say that it is INJECTIVE or ONE-ONE. A function which is both injective and surjective is therefore bijective. Such functions are also called ONE-ONE ONTO. (Still another name, "*biunique*", was used for such functions in older editions of *Introduction to Logic*.)

depends exclusively on the position of the point X, we may denote it by the symbol "$f(X)$"; we consequently have:

$$y = f(X).$$

But, conversely, to every non-negative number y we may also correlate a uniquely determined segment OX on the half-line under consideration, whose length equals y; in other words, to every y there corresponds exactly one point X such that

$$y = f(X).$$

The function f is therefore bijective; it establishes a one-to-one correspondence between the points of the half-line and the non-negative numbers (and it would be just as simple to set up a one-to-one correspondence between the points of the entire line and all real numbers). Another example is provided by the relation which is expressed by the formula:

$$y = -x.$$

This is a bijective function since, to every number y, there is only one number x satisfying the given formula; it can be seen at once that this function maps, for instance, the set of all positive numbers onto the set of all negative numbers in a one-to-one manner. As the last example let us consider the relation which is expressed by the formula:

$$y = 2x$$

under the assumption that the symbol "x" here denotes natural numbers only (or, as one says, that "x" ranges over the set of natural numbers). Again we have a bijective function; it correlates with every natural number x the even number $2x$; and vice versa—to every even natural number y there corresponds just one number x such that $2x = y$, namely, the number $x = \frac{1}{2}y$. The function thus establishes a one-to-one correspondence between all natural numbers and even natural numbers.—Numerous examples of bijective functions, or equivalently, of one-to-one mappings, can be drawn from the field of geometry (mappings preserving a given symmetry, collinear mappings, and so on).

* * *

Now that the notion of a one-to-one correspondence has been introduced and is at our disposal, we are in a position to give a precise definition of a concept which, previously, we had been able to characterize only in an intuitive way. It is the concept of equinumerous classes (see Section 26). We shall now say that two classes K and L are equinumerous, or, that they have the same cardinal number, if there exists a function which establishes a one-to-one correspondence between the elements of the two classes. From this definition, together with the examples considered above, we conclude that the set of all points of an arbitrary half-line is equinumerous with the set of all non-negative numbers; and likewise, that the set of positive numbers and the set of negative numbers are equinumerous, and that the same holds for the set of all natural numbers and the set of all even natural numbers. The last example is particularly instructive; for it shows that a class may be equinumerous with a proper subclass of itself. To many readers this fact may seem highly paradoxical at first glance, because they have been used to comparing only finite classes with respect to the numbers of their elements; and a finite class, as one expects, always has a greater cardinal number than any of its parts. The paradox disappears when we realize that the set of natural numbers is infinite, and that we are not entitled, under any pretense, to ascribe to infinite classes those properties which we have observed when dealing exclusively with finite classes.—It is noteworthy that the foregoing property of the set of natural numbers, that of being equinumerous with one of its parts, is shared by all infinite classes. This property is therefore characteristic of infinite classes, and it permits us to distinguish them from finite classes; one can set up a definition whereby a class is finite if, and only if, it is not equinumerous with any one of its proper subclasses.[2] (However, this definition entails a certain logical difficulty, a discussion of which we shall not undertake here.)

34 Many-place relations; functions of several variables and operations

So far we have considered exclusively BINARY or TWO-PLACE RELATIONS, that is, relations involving two objects. However, one also frequently encounters THREE-PLACE or TERNARY, and in general, MANY-PLACE RELATIONS within various sciences. For instance, the relation of betweenness in geometry constitutes a typical example of a three-place relation; it may hold for three points on a line, and is expressed symbolically by the formula:

$$A/B/C$$

which is read:

the point B lies between the points A and C.

---

[2] One of the the first to call attention to the property of infinite classes which is considered here was the Austrian-Bohemian philosopher and mathematician B. BOLZANO (1781–1848); in his book *Paradoxien des Unendlichen* (Leipzig 1851, posthumously published) we already find the first beginnings of the contemporary theory of sets. The above property was later employed by PEIRCE (cf. footnote 2 on p. 12) and others in order to formulate a precise definition of a finite and of an infinite class.

Arithmetic, too, supplies numerous examples of three-place relations; it may suffice to mention the relation which involves three numbers x, y, and z and asserts that the last number is the sum of the other two:

$$x + y = z,$$

as well as similar relations, which are expressed by the following formulas:

$$x - y = z,$$
$$x \cdot y = z,$$
$$x \div y = z.$$

As an example of a four-place relation, let us mention the relation holding for the four points A, B, C, and D if, and only if, the distance between the first two equals the distance between the last two, or, in other words, if the segments AB and CD are congruent. Another example is the relation holding among the numbers x, y, z, and t whenever they form a proportion:

$$x : y :: z : t.$$

* * *

Of particular importance among the many-place relations are the many-place functional relations, which are analogous to the binary functional relations. For reasons of simplicity we shall restrict our discussion to three-place relations of this type. R is called a THREE-PLACE FUNCTIONAL RELATION if, to any two objects x and y, there corresponds at most one object z having this relation to x and y. We denote this uniquely determined object, provided it exists at all, either by the expression:

$$R(x, y)$$

or else by the expression:

$$xRy$$

(which now assumes a different meaning from what it had in the theory of binary relations). Accordingly, in order to state that z stands to x and y in the functional relation R, we have two formulas at our disposal:

$$R(x, y) = z \quad \text{and} \quad xRy = z.$$

Corresponding to this twofold symbolism we have a twofold mode of expression. When using the notation:

$$R(x, y) = z,$$

the relation R is called a FUNCTION. In order to distinguish between two-place and three-place functional relations, we speak in the first case of FUNCTIONS OF ONE VARIABLE, or, of FUNCTIONS WITH ONE ARGUMENT, and in the second, of FUNCTIONS OF TWO VARIABLES or of FUNCTIONS WITH TWO ARGUMENTS. Similarly, four-place functional relations are called FUNCTIONS OF THREE VARIABLES or FUNCTIONS WITH THREE ARGUMENTS, and so on. In designating functions with any number of arguments it is customary to employ, as before, the variables "f", "g", ... (and the symbol which denotes the function value is often shown first); the formula:

$$z = f(x, y)$$

can be read:

z is that value of the function f which is correlated with the argument values x and y.

When the symbolism:

$$z = xRy$$

is employed, the relation R is usually referred to as an OPERATION or, specifically, a BINARY OPERATION, and the above formula is read as follows:

z is the result of the operation R carried out on x and y;

in this case, in place of the letter "R" we tend to use other letters, especially the letter "O". The four fundamental arithmetical operations of addition, subtraction, multiplication, and division may serve as examples, and also such logical operations as addition and multiplication of classes or relations (see Sections 25 and 28). The content of the two concepts, of a function of two variables and of a binary operation, is evidently exactly the same. It should perhaps be noted that functions of one variable are sometimes also called operations, and, in particular, UNARY OPERATIONS; in the algebra of classes, for instance, we usually think of the complement K' as the result of an operation carried out on K (and usually we do not speak of a function which assigns the value K' to the argument value K).

* * *

Although many-place relations play an important role in various sciences, the general theory of these relations is as yet in its initial stage; when speaking of a relation, or of the theory of relations, one usually has only binary relations in mind. A detailed study has been made so far only of one particular category of three-place relations, namely, of a category of binary operations, as the prototype of which we may consider the ordinary arithmetical addition. Many of these investigations were carried out within the framework of a special mathematical discipline known as the theory of groups. We shall become acquainted with certain concepts from the theory of groups—and thereby also with certain general properties of binary operations—in the Second Part of this book.

35 The importance of logic for other sciences

We have discussed the most important concepts of contemporary logic, and in doing so we have become acquainted with certain laws (very few, by the way) involving these concepts. It was not our intention, however, to give a complete list of all logical concepts and laws which are applied in scientific arguments. This, incidentally, is not necessary for the study and development of other sciences, not necessary even for the study of mathematics, which is especially closely related to logic.—Let us say this in a different way: logic is justly considered the basis[†] of all other sciences, even if only for the reason that in every argument we employ concepts taken from the field of logic, and that every correct inference proceeds in accordance with its laws. But this does not mean that a thorough knowledge of logic is essential for correct thinking; even professional mathematicians, who, in general, do not commit errors in their inferences, usually do not know logic to such an extent as to be conscious of all logical laws of which they make subconscious use. All the same, there can be no doubt that the knowledge of logic is of considerable practical importance for everyone who desires to think and to infer correctly, since it enhances the innate and the acquired faculties in this direction and, in unusual cases, prevents the committing of mistakes. Moreover, insofar as the construction of mathematical theories is concerned, logic plays a role of far-reaching importance also from a more fundamental point of view; this aspect will be discussed in the next chapter.

Exercises

1. Give examples of relations from the fields of arithmetic, geometry, physics, and everyday life.

2. Consider the relation of being father, that is to say, the relation expressed by the sentential function:

$$x \textit{ is father of } y.$$

Do all human beings belong to the domain of this relation? And do they all belong to the counter-domain?

[†] It is of interest that the author chose to say "*the basis of*" rather than "*basic to*". The two phrases refer to different attitudes that one can adopt toward logic. For the author, logic was the very foundation of mathematics, and not only an essential tool.

3. Consider the following seven relations among people, namely, of being father, mother, child, brother, sister, husband, wife; cf. the previous exercise. We denote these relations by the symbols "$\mathcal{F}$", "$\mathcal{M}$", "$\mathcal{C}$", "$\mathcal{B}$", "$\mathcal{S}$", "$\mathcal{H}$", "$\mathcal{W}$". By applying various operations defined in Section 28 to these relations, we obtain new relations for which we sometimes find simple names in ordinary language; "$\mathcal{H}|\mathcal{C}$", for instance, denotes the relation of being son-in-law (see a similar example in Section 28). Find, if possible, simple names for the following relations:

$$\breve{\mathcal{B}},\quad \breve{\mathcal{H}},\quad \mathcal{H}\cup\mathcal{W},\quad \mathcal{F}\cup\mathcal{B},\quad \mathcal{F}|\mathcal{M},\quad \mathcal{M}|\breve{\mathcal{C}},\quad \mathcal{B}|\breve{\mathcal{C}},$$
$$\mathcal{F}|(\mathcal{H}\cup\mathcal{W}),\quad (\mathcal{B}|\breve{\mathcal{C}})\cup[\mathcal{H}|(\mathcal{S}|\breve{\mathcal{C}})].$$

Express with the help of the symbols "$\mathcal{F}$", "$\mathcal{M}$", and so on, together with the symbols of the algebra of relations, the relations of being parent, sibling, grand-child, daughter-in-law, and mother-in-law.

Explain the meanings of the following formulas, and determine which of them are true:

$$\mathcal{F}\subseteq\mathcal{M}',\quad \breve{\mathcal{B}}=\mathcal{S},\quad \mathcal{F}\cup\mathcal{M}=\breve{\mathcal{C}},\quad \mathcal{H}|\mathcal{M}=\mathcal{F},\quad \mathcal{B}|\mathcal{S}\subseteq\mathcal{B},\quad \mathcal{S}\subseteq\mathcal{C}|\breve{\mathcal{C}}.$$

4. Consider the following two formulas of the algebra of relations:

$$R|S=S|R \quad\text{and}\quad \overset{\smile}{(R|S)}=\breve{S}|\breve{R}.$$

Show by means of an example that the first is not always satisfied, and try to show that the second is satisfied by arbitrary relations R and S.

Hint: Consider what it means to say that the relation $\overset{\smile}{(R|S)}$ (that is, the converse of the relation $R|S$) or that the relation $\breve{S}|\breve{R}$ holds between two objects x and y.

5. Formulate in symbols the definitions of all the terms of the algebra of relations that were discussed in Section 28.

Hint: The definition of the sum of two relations, for instance, has the following form:

$$[x(R\cup S)y]\leftrightarrow[(xRy)\vee(xSy)].$$

6. Consider the following relations; which of the properties of relations discussed in Section 29 does each possess?

(a) the relation of divisibility in the set of natural numbers;
(b) the relation of being relatively prime in the set of natural numbers (two natural numbers being called relatively prime to each other if their greatest common divisor is 1);
(c) the relation of congruence in the set of polygons;
(d) the relation of being longer in the set of line segments;
(e) the relation of being perpendicular in the set of straight lines in a plane;

(f) the relation of intersecting in the set of geometrical configurations;
(g) the relation of simultaneity in the class of physical events†;
(h) the relation of temporally preceding in the class of physical events†;
(i) the relation of being related in the class of human beings;
(j) the relation of fatherhood in the class of human beings.

7. Is it true that every relation is either reflexive or irreflexive (in a given class), and either symmetric or asymmetric? Give examples.

8. We shall call a relation R INTRANSITIVE IN THE CLASS K if, for any three elements x, y, and z of K, the formulas:

$$xRy \text{ and } yRz$$

imply the formula:

$$\sim (xRz).$$

Which of the relations listed in exercises 3 and 6 are intransitive? Give other examples of intransitive relations. Is every relation either transitive or intransitive?

***9.** Show how to make the transition from the expression:

the lines a and b are parallel

to the equivalent one:

the directions of the lines a and b are identical,

and how to define the expression "*the direction of a line*".

Solve the same problem for the following two expressions:

the segments AB and CD are congruent

and

the lengths of the segments AB and CD are equal.

What logical law has to be applied here?

Hint: Compare the remarks in Section 30 concerning the concept of similarity.

---

† A reader who has studied modern physics will realize that the relations which are mentioned in (g) and (h) have different characteristics, depending on whether one refers to the classical description of space and time, or to a relativistic refinement. The familiar classical description should be assumed for the purposes of this exercise. (We may add, that at the time of the early editions of *Introduction to Logic*, when this exercise was formulated, the appreciation of intricacies of theories of relativity was not as widespread as it is today.)

10. Let us agree to call two signs, or two expressions consisting of several signs, EQUIFORM, if they do not differ with regard to their external appearance, but may differ with respect to their position in space, that is, with respect to the place at which they are printed; otherwise let us call them NON-EQUIFORM. For instance, in the formula:

$$x = x,$$

the variables on the two sides of the equality sign are equiform, whereas we have non-equiform variables in the formula:

$$x = y.$$

Of how many signs does the formula:

$$x + y = y + x$$

consist? Into how many groups can these signs be divided, such that two equiform signs belong to the same group and two non-equiform signs belong to different groups?

Which of the properties discussed in Section 29 belong to the relations of equiformity and non-equiformity?

***11.** Explain, on the basis of the results of the preceding exercise, why one may say that equiform signs are equal with respect to their FORM, or, that they have the same form, and explain how the term "*the form of the given sign*" is to be defined (cf. exercise *9).††

It is a very common usage to call equiform signs simply equal, and even to treat them as if they were one and the same sign. For instance, one often says that in an expression like:

$$x + x$$

the same variable occurs on both sides of the symbol "+". How should this be expressed with greater precision?

***12.** The imprecise mode of speech which was pointed out in exercise *11 has also been employed several times in this book (after all, we do not want to contend over deeply rooted usages). Show that inexactitudes of this kind occur on pp. 11 and 51, and explain how they could be avoided.

Another example of an inexact mode of speech of this kind is the following: when speaking of sentential functions with one free variable, one means functions in which all free variables are equiform. How can the expression:

sentential functions with two free variables

††The reader may observe that here the word "*form*" means about the same as "*external appearance*" in exercise 10, and as "*shape*" in Sections 13 and 21. (Cf. "*shape of the variables*", "*letters of different shape*".) On the other hand, in Section 30 the word "*shape*" is introduced with a somewhat different meaning. The over-all situation is perhaps a bit confusing. However, it illustrates the kind of difficulty that is frequently encountered when one tries to find terms that would be clear as well as intuitive.

be formulated more precisely?

13. Given a point in a plane, consider the set of all circles in that plane which have the given point as their common center. Show that this set is ordered by the relation of being a part. Would this be also true, if the circles did not lie in the same plane, or if they were not concentric?

14. We consider a relation among words of the English language which will be called the relation of PRECEDING (IN LEXICOGRAPHICAL ORDER). The meaning of this term can be best explained by means of examples. The word "*and*" precedes the word "*can*", since the first begins with "*a*", the second with "*c*", and "*a*" has an earlier place in the Latin alphabet than "*c*". The word "*air*" precedes the word "*ale*", since they have the same first letter (or, rather, equiform first letters—cf. exercise 10), while the second letter of the first word, that is "*i*", has an earlier place in the Latin alphabet than the second letter of the second word, that is "*l*". In the analogous way, "*each*" precedes "*eat*", and "*timber*" precedes "*time*". Finally, "*war*" precedes "*warfare*", since the first three letters of the two words are the same, and the first word has only these letters, while the second has more than these; and analogously "*mean*" precedes "*meander*".

Write the following words in a line so that, of any two words, the one on the left precedes the one on the right:

care, arm, salt, art, car, sale, trouble, army, ask.

Try to define the foregoing relation of preceding among words precisely and in a quite general way. Show that this relation establishes an order in the set of all English words (without hyphen or apostrophe). Point out some practical applications of this relation, and explain why this particular order sometimes is called "*lexicographical*".

15. Consider an arbitrary relation R and its negation R', and also a class K. Show that the following statements of the theory of relations are true:

(a) *if the relation R is reflexive in the class K, then the relation R' is irreflexive in that class;*
(b) *if the relation R is symmetric in the class K, then the relation R' is also symmetric in that class;*
(c) *if the relation R is asymmetric in the class K, then the relation R' is reflexive and connected in that class;*
(d) *if the relation R is transitive and connected in the class K, then the relation R' is transitive in that class.*

Are the converses of these statements likewise true?

16. Show that, if the relation R has any of the properties discussed in Section 29, then the converse relation $\breve{R}$ possesses the same properties.

***17.** The properties of relations which were introduced in Section 29 can easily be expressed in terms of the algebra of relations, provided the class K to which they refer is the universal class. The formulas:

$$R|R \subseteq R \quad \text{and} \quad \mathcal{D} \subseteq R \cup \breve{R},$$

for instance, assert that the relation R is transitive and connected, respectively. Explain why; recall the meaning of the symbol "$\mathcal{D}$" from Section 28. Give similar formulas asserting that the relation R is symmetric, asymmetric, or intransitive (cf. exercise 8). Of the properties of relations discussed in the present chapter, which is expressed by the formula:

$$\breve{R}|R \subseteq \mathcal{I}?$$

18. Which of the relations expressed by the following formulas are functions?

(a) $2y + 3x = 12$,
(b) $y^2 = x^2$,
(c) $y + 2 > x - 3$,
(d) $y + x = x^2$,
(e) y *is mother of* x,
(f) y *is daughter of* x.

Which of the relations considered in exercise 3 are functions?

19. Consider the function which is expressed by the formula:

$$y = x^2 + 1.$$

What is the set of all argument values, and what is the set of all function values?

***20.** Which of the functions in exercise 18 are bijective? Give other examples of bijective functions.

***21.** Consider the function expressed by the formula:

$$y = 3x + 1.$$

Show that this is a bijective function, and that it maps the interval $[0, 1]$ onto the interval $[1, 4]$ in a one-to-one manner (cf. exercise 6 of Chapter IV). What conclusion may be drawn from this concerning the cardinal numbers of those intervals?

***22.** Consider the function expressed by the formula:

$$y = 2^x.$$

Use this function to show, along the lines of the preceding exercise, that the set of all numbers and the set of all positive numbers are equinumerous.

***23.** Show that the set of all natural numbers and the set of all odd numbers are equinumerous.

24. Give examples of many-place relations from the fields of arithmetic and geometry.

25. Which of the three-place relations expressed by the following formulas are functions?

(a) $x + y + z = 0$,
(b) $z \cdot x > 2y$,
(c) $z^2 = x^2 + y^2$,
(d) $z + 2 = x^2 + y^2$.

26. Name a few laws of physics that are expressed with the help of functional relations; include laws which involve two, three, and four quantities.

27. We consider the relation of betweenness, expressed symbolically by the formula "$A/B/C$", where A, B, and C are three distinct points in a given plane (cf. Section 34). Write the following two sentences in symbols:

(a) *for any points A, B, C, and D, if B lies between A and C and also between A and D, then C lies between A and D;*
(b) *for any points A and B, if A and B are distinct, then there is a point C such that neither B lies between A and C nor C lies between B and A nor A lies between B and C.*

Also translate the following formulas into ordinary language:

(c) $\forall_{A,B,C,D}[(B \neq C \wedge A/B/D \wedge A/C/D) \rightarrow (A/B/C \vee C/B/D)]$;
(d) $\exists_{A,C}[A \neq C \wedge \forall_B \sim (A/B/C)]$.

Which of sentences (a)–(d) are true? (Do not assume that the points involved lie necessarily on the same line.)

***28.** Consider the following three formulas involving a binary operation O:

$$\forall_{x,y}(xOy = yOx), \quad \forall_{x,y,z}[(xOy)Oz = xO(yOz)],$$
$$\forall_{x,y}\exists_z(x = yOz).$$

In these formulas substitute successively for "O" the four arithmetical symbols: "+", "−", "·", and "÷". Which of the resulting sentences are true?

VI

On the Deductive Method

36 Fundamental constituents of a deductive theory—primitive and defined terms, axioms and theorems

We shall now attempt to give an exposition of the fundamental principles and methods that should be applied in the construction of logical and of mathematical theories.[1] The detailed analysis and the critical evaluation of these principles are the tasks of a special discipline, called the METHODOLOGY OF DEDUCTIVE SCIENCES, or, the METHODOLOGY OF MATHEMATICS. Now, a familiarity with the methods which are employed in the construction of a certain science is undoubtedly essential to anyone who is involved in its study or advancement; and we shall see that, in the case of mathematics, having a firm understanding of those methods is of particularly far-reaching significance, for otherwise it is impossible to discern the nature of this subject.

* * *

The principles with which we shall become acquainted serve the purpose of assuring for mathematics (as well as for logic) the highest possible degree of clarity and certainty. From this point of view a method or a procedure might be considered ideal, if it allowed us to explain the meaning of every expression occurring in this science and to justify each of its assertions. It is easy to see that this ideal can never be realized. In fact, when one tries to explain the meaning of an expression, then one necessarily uses other

[1]Ideas which are closely related to those presented in this section can be found in earlier literature. See, for instance, the opusculum (posthumously published), *De l'esprit géométrique et de l'art de persuader*, of the great French philosopher and mathematician B. PASCAL (1623–1662).

expressions; and in turn, to explain the meaning of these other expressions without entering into a vicious circle, one has to resort to still further expressions, and so on. We thus have the beginning of a process which can never be brought to an end, a process which, figuratively speaking, may be characterized as an INFINITE REGRESS—a *regressus in infinitum*. The situation is quite analogous with regard to the justification of the asserted statements of the science; for, in order to establish the validity of a statement, it is necessary to refer back to other statements, and (if no vicious circle is to occur) this leads again to an infinite regress.

* * *

By way of a compromise between that unattainable ideal and the realizable possibilities, certain principles regarding the construction of mathematical disciplines have emerged, and they may be described as follows.

When we set out to construct a given discipline, we distinguish, first of all, a certain small group of expressions of this discipline that seem to us to be immediately understandable; we call the expressions of this group PRIMITIVE TERMS or UNDEFINED TERMS, and we employ them without explaining their meanings. At the same time we adopt the principle: not to use any other expression of the discipline under consideration, unless its meaning has first been determined, with the help of primitive terms and of such expressions of the discipline whose meanings were explained previously. The sentences which determine the meanings of terms in this way are called DEFINITIONS, and the expressions themselves whose meanings are thereby determined are accordingly known as DEFINED TERMS.

We proceed similarly with respect to the asserted statements of the discipline in question. Some of these statements, whose truth appears to us evident, are chosen for the so-called PRIMITIVE STATEMENTS or AXIOMS (also often referred to as POSTULATES, but we shall not use the latter term in this technical meaning here); we accept them as true without establishing them in any way. On the other hand, we agree to accept any other statement as true only if we have established its validity, and if we used for this purpose nothing but the axioms, the definitions, and those statements of the discipline which were established previously. As is well known, statements which are justified in such a way are called PROVED STATEMENTS or THEOREMS, and the processes of justifying them are called PROOFS. More generally, if within logic or mathematics we establish one statement by starting with others, we refer to this process as a DERIVATION or DEDUCTION, and the statement in question is said to be DERIVED or DEDUCED from the other statements, or, to be their CONSEQUENCE.

* * *

Contemporary mathematical logic is one of those disciplines which have been constructed in accordance with the principles just stated (we shall use the term DEDUCTIVE LOGIC when referring to such a construction); unfortunately, it has not been possible within the narrow framework of this book to give due prominence to this important fact. If any other discipline is constructed in accordance with these principles, it is already based upon logic; logic, so to speak, is then already presupposed. This means that all expressions and all laws of logic are treated on an equal footing with the primitive terms and the axioms of the discipline in question; in particular, logical terms are used in the formulation of the axioms, theorems, and definitions without an explanation of their meaning, and logical laws are applied in proofs without having been established first. Sometimes, moreover, it is convenient not only to use logic in the construction of a discipline, but to presuppose in the same sense certain mathematical disciplines which were constructed previously; these theories, together with logic, may be characterized briefly as the DISCIPLINES PRECEDING THE GIVEN DISCIPLINE. Thus logic itself does not presuppose any preceding discipline; in constructing arithmetic as a special mathematical discipline, one presupposes logic as the only preceding discipline; on the other hand, in the case of geometry it is expedient—though not unavoidable—to presuppose not only logic but also arithmetic.

In view of the last remarks, it is necessary to make certain corrections in the formulation of the principles stated above. Before undertaking the construction of a discipline, one has to enumerate the disciplines that are to precede the one to be constructed; all requirements applying to the definitions of terms and to the proofs of statements will then be limited to those terms and statements which are specific for the discipline to be constructed, that is, to those which do not belong to the preceding disciplines.

* * *

The method of constructing a discipline in strict accordance with the foregoing principles is known as the DEDUCTIVE METHOD, and the disciplines constructed in this manner are called DEDUCTIVE THEORIES.[2] The

[2]The deductive method certainly is not an achievement of recent times. Already in the *Elements* of the Greek mathematician EUCLID (about 300 B.C.) we find a presentation of geometry which does not leave much to be desired from the standpoint of the methodological principles stated above. For some 2200 years mathematicians saw in EUCLID's work the ideal and the prototype of scientific exactitude. An essential advance in this field took place only after 1890, when the foundations of the basic mathematical disciplines of geometry and arithmetic were laid in accordance with the requirements of the present-day methodology of mathematics. Among the works to which we are indebted for this progress are the following two, which have already acquired historic importance: the collective work *Formulaire de Mathématiques* (Torino 1895–1908), whose editor and main author was the Italian mathematician and logician G. PEANO (1858–1932), and *Grundlagen der Geometrie* (Leipzig and Berlin 1899) by the renowned German mathematician D. HILBERT (1862–1943); in particular, the latter work stimulated further research in the foundations of geometry.

view has become more and more widely accepted, that *the deductive method is the only essential feature which distinguishes the mathematical disciplines from all other sciences; not only is every mathematical discipline a deductive theory, but also, conversely, every deductive theory is a mathematical discipline* (and according to this view, deductive logic likewise is to be regarded as a mathematical discipline). We shall not enter here into a discussion of the reasons in favor of this view, but only remark that one can put forward ponderable arguments in its support.

37 Models and interpretations of a deductive theory

If one adheres systematically to the principles which were presented in the preceding section, then the deductive theories acquire a number of interesting and important features. Since some of the questions which arise are rather involved and abstract, we shall try to elucidate them first by means of a concrete example.

Suppose we are interested in general facts about the congruence of line segments, and we intend to establish this fragment of geometry as a special deductive theory. We accordingly stipulate that the variables "x", "y", "z", ... denote segments. As primitive terms we choose the symbols "S" and "$\cong$". The former is an abbreviation of the term "*the set of all segments*"; the latter designates the relation of congruence, so that the formula:

$$x \cong y$$

is to be read as follows:

the segments x and y are congruent.

Furthermore, we adopt only two axioms:

AXIOM 1. *For any element x of the set S, $x \cong x$ (*in other words: *every segment is congruent to itself).*

AXIOM 2. *For any elements x, y, and z of the set S, if $x \cong z$ and $y \cong z$, then $x \cong y$ (*in other words: *two segments congruent to the same segment are congruent to each other).*

Various theorems on the congruence of segments may be derived from these axioms, for instance:

THEOREM 1. *For any elements y and z of the set S, if $y \cong z$, then $z \cong y$.*

THEOREM 2. *For any elements x, y, and z of the set S, if $x \cong y$ and $y \cong z$, then $x \cong z$.*

The proofs of these two theorems are very easy. For instance, let us sketch the proof of the first.

Putting in Axiom 2 "z" for "x", we obtain:

for any elements y and z of the set S, if $z \cong z$ and $y \cong z$, then $z \cong y$.

In the hypothesis of this statement we have the formula:

$$z \cong z$$

which, on the basis of Axiom 1, is undoubtedly valid, and hence may be omitted. We thus arrive at the theorem in question.

* * *

In connection with these simple considerations, we want to make the following remarks.

Our miniature deductive theory rests upon a suitably selected system of primitive terms and axioms. Our knowledge of the objects which are denoted by the primitive terms, that is, of the set of segments and of its congruence relation, is very comprehensive and is by no means exhausted by the adopted axioms. But this extended knowledge is, so to speak, our private concern which does not exert the least influence on the construction of our theory. In particular, in deriving theorems from the axioms we make no use whatsoever of this knowledge; we behave as though we did not understand the concepts which are involved in our considerations, and as if we knew nothing about the objects except for what has been expressly asserted in the axioms. We disregard, as it is commonly put, the meaning of the primitive terms adopted by us, and direct our attention exclusively to the form of the axioms in which these terms occur.

This leads to a very significant and interesting consequence. Let us replace the primitive terms in all the axioms and in all theorems of our theory by suitable variables, for instance, the symbol "S" by the variable "K" denoting classes, and the symbol "$\cong$" by the variable "R" denoting relations (and in order to simplify the considerations, we shall disregard here definitions and any theorems which contain defined terms). The statements of our theory will then be no longer sentences, but will become sentential functions which contain two free variables, "K" and "R", and which express, generally speaking, the condition that the relation R has this or that property in the class K (one can also say: that a certain relation holds between K and R; cf. Section 27). For instance, as is easily seen, Axiom 1 and Theorems 1 and 2 will now state that the relation R is reflexive, symmetric, and transitive, respectively, in the class K. Axiom 2 will express a property for which we do not have any special name and to which we shall refer as property $\mathcal{P}$; this is the following property:

for any elements x, y, and z of the class K, if xRz and yRz, then xRy.

In the proofs of our theory, we do not use any properties of the class of segments and of the relation of congruence other than those which were explicitly stated in the axioms; every proof can therefore be considerably generalized, for it can be applied to any class K and any relation R having those properties. As a result of such a generalization of the proofs, we can correlate with any theorem of our theory a general law belonging to the domain of logic, in particular to the theory of relations, and stating that every relation R which is reflexive and has the property $\mathcal{P}$ in a given class K also has the property which is expressed in the theorem under consideration. In this way we are led to the following two laws of the theory of relations, which correspond to Theorems 1 and 2[†]:

I. *Every relation R which is reflexive in a given class K and has the property $\mathcal{P}$ in that class is also symmetric in K.*

II. *Every relation R which is reflexive in a given class K and has the property $\mathcal{P}$ in that class is also transitive in K.*

If a relation R is reflexive and has the property $\mathcal{P}$ in a class K, we say that K and R together form a MODEL, or, a REALIZATION of the axiom system of our theory, or simply, that they satisfy the axioms. For instance, one model of the axiom system is formed by the class of all segments and the relation of congruence, that is, by the objects which are denoted by the primitive terms; of course, this model also satisfies all theorems which can be deduced from the axioms. (To be precise, we ought to say that a model satisfies not the statements of the theory themselves, but the sentential functions which are obtained from them by replacing the primitive terms by variables.) However, this particular model does not play any privileged role in the construction of the theory. We can say, in fact, the following: universal logical laws like I and II illustrate the general conclusion, that any model of the axiom system satisfies every theorem which can be deduced from these axioms. For this reason, a model of the axiom system of our theory is also referred to as a MODEL OF THE THEORY itself.

We are able to exhibit many different models for our axiom system, even in the domains of logic and elementary mathematics. Such a model can be obtained by selecting within any other deductive theory (to which we shall refer as "new") two suitable constants, say "$\mathcal{K}$" and "$\mathcal{R}$" (the former denoting a class, the latter a relation), then replacing "$\mathcal{S}$" by "$\mathcal{K}$" and "$\cong$" by "$\mathcal{R}$" everywhere in Axioms 1 and 2, and finally showing that the sentences which are obtained in this way are theorems, or possibly axioms, of the new theory. If we have succeeded in doing so, we say that we have found an INTERPRETATION of the axiom system within the new deductive

[†]One could adopt here the pattern which was introduced in Section 3 and express these laws in the following way: *For any class K and any relation R, ...* However, logicians, like other people, enjoy having some variety in speaking and writing!

theory; but the new theory is assumed to apply to the class $\mathcal{K}$ and to the relation $\mathcal{R}$, and therefore these objects determine a model; and the model as well as the interpretation will extend from the axiom system to our whole deductive theory. For, if we replace the primitive terms "$\mathcal{S}$" and "$\cong$" by "$\mathcal{K}$" and "$\mathcal{R}$" not only in the axioms, but also in the theorems of our theory, we can be sure in advance that all sentences which are obtained in this way will be asserted sentences of the new theory, and therefore true sentences.

* * *

We shall give here two concrete examples of models and interpretations of our miniature theory. Let us replace in Axioms 1 and 2 the symbol "$\mathcal{S}$" by the symbol of the universal class "**V**", and the symbol "$\cong$" by the identity sign "$=$". As one can see immediately, the axioms will then become logical laws (to be specific, Laws II and V of Section 17 in a slightly modified form). The universal class and the relation of identity constitute therefore a model of the axiom system, and our theory has found an interpretation within logic. Consequently, if in Theorems 1 and 2 we replace the symbols "$\mathcal{S}$" and "$\cong$" by the symbols "**V**" and "$=$", we are sure to arrive at sentences which again are logical laws (in fact, we already are familiar with them—cf. Laws III and IV of Section 17).

Next, let us consider the set of all numbers, or any other set of numbers, denoting it by "$\mathcal{N}$". Let us call two numbers x and y equivalent, in symbols:

$$x \equiv y,$$

if their difference $x - y$ is an integer; we have, for example:

$$1\tfrac{1}{4} \equiv 5\tfrac{1}{4},$$

whereas it is not the case that

$$3 \equiv 2\tfrac{1}{3}.$$

If now, in both axioms, the primitive terms are replaced by "$\mathcal{N}$" and "$\equiv$", it can be easily shown that the resulting sentences are asserted statements of arithmetic. Our theory therefore has an interpretation within arithmetic, while the chosen set of numbers $\mathcal{N}$ and the relation of equivalence $\equiv$ constitute a model of the axiom system. And again, without any additional arguments, we are sure that Theorems 1 and 2 will become true arithmetical statements if they are subjected to the same transformation as the axioms.

* * *

The general features which we described above have many interesting applications in methodological research. Here we shall illustrate such applications by a single example; we shall show how it may be proved—on the basis of the foregoing observations—that certain sentences cannot be deduced from our axiom system.

Let us consider the following sentence A (which is formulated in logical terms and in the primitive terms of our theory only):

A. *There exist two elements x and y of the set S for which it is not the case that $x \cong y$* (in other words: *there exist two segments which are not congruent).*

This sentence seems to be undoubtedly true. Nevertheless, all attempts to prove it on the basis of Axioms 1 and 2 alone fail to give a positive result. Thus the conjecture arises that Sentence A simply cannot be deduced from our axioms. In order to confirm this conjecture, we argue in the following way. If Sentence A could be deduced from our two axioms, then, as we know, every model of this system would satisfy that sentence; therefore, if we succeed in describing a model of the axiom system which does not satisfy Sentence A, we shall have obtained a proof that this sentence cannot be deduced from Axioms 1 and 2. Now, it turns out that producing such a model is not at all difficult. Let us consider, for instance, the set of all integers or any other set of integers, to be denoted by $\mathcal{I}$ (e.g. the set consisting of only the numbers 0 and 1 would be suitable), and the relation of equivalence $\equiv$ among numbers which was discussed above. We already know from the preceding remarks that such a set $\mathcal{I}$ and the relation $\equiv$ constitute a model of our axiom system; Sentence A however is not satisfied by this model, for there are no two integers x and y which are not equivalent, that is, whose difference is not an integer. Another model appropriate to this purpose is formed by an arbitrary class of individuals and the universal relation $\mathbf{V}$, which holds between any two individuals (this relation, of course, should not be confused with the universal class).

The type of reasoning which has just been applied is known as the METHOD OF PROOF BY EXHIBITING A MODEL; a variant relates directly to interpretations, and is called the METHOD OF PROOF BY INTERPRETATION.

* * *

The observations and the concepts which were discussed above can be extended, without essential change, to other deductive theories. In the next section we shall try to describe them in a quite general way.

38 The law of deduction; formal character of deductive sciences

*We consider an arbitrary deductive theory, which is based upon a certain system of primitive terms and axioms. In order to simplify our consideration, we assume that this theory presupposes logic only, in other words, that logic is the only theory preceding the given theory (cf. Section 36).

Let us imagine that in all statements of our theory the primitive terms are replaced by suitable variables throughout (as in Section 37; again we disregard definitions and theorems containing defined terms for the sake of simplicity). The statements of the theory in question then become sentential functions containing as free variables those symbols by which the primitive terms had been replaced, and not containing any constants but those belonging to logic. If certain objects, that is, classes, relations, etc., are given, one can try to find out whether they satisfy all the axioms of our theory, or rather, to be precise, whether they satisfy the sentential functions which are obtained from these axioms in the manner just described (that is, whether the names or designations of those objects, when put in place of the free variables, render true sentences out of the sentential functions; cf. Section 2). If this turns out to be the case, we shall say that the objects under consideration form a MODEL OR REALIZATION OF THE AXIOM SYSTEM of our deductive theory; we also say sometimes that they form a MODEL OF THE DEDUCTIVE THEORY itself. In an entirely analogous manner we can try to find out whether given objects satisfy any other chosen system of statements of our theory, and whether, therefore, they form a model of this system (where the statements can be axioms or not, and where the system may consist of a single statement).

A model of the axiom system is formed, for instance, by those objects which are denoted by the primitive terms of the given theory, since we assume that all the axioms are true sentences; this model then satisfies, of course, all theorems of our theory. But as far as the construction of our theory is concerned, this model has no distinguished place within the totality of models. When deducing this or that theorem from the axioms, we do not think of any specific properties of this model, and we make use of only those properties which are explicitly stated in the axioms and which, therefore, belong to every model of the axiom system. Consequently, a proof of any particular theorem of our theory can be extended to every model of the axiom system and, moreover, can be transformed (by replacing the primitive constants by variables, as before) into a much more general argument which belongs no longer to our theory but to logic; and as a result of this generalization we obtain a general logical statement (like laws I and II of the preceding section) which asserts that the theorem in question, when suitably interpreted, is satisfied by every model of our axiom system. The final conclusion at which we arrive in this way can be put in the following form:

> *Every theorem of a given deductive theory is satisfied by any model of the axiom system of this theory; and moreover, to every theorem there corresponds a general statement which can be formulated and proved within the framework of logic, and which asserts the fact that the theorem in question is satisfied by any such model.*

We have here a general law from the domain of the methodology of deductive sciences which, when formulated in a somewhat more precise way, is known as the LAW OF DEDUCTION (or, as the DEDUCTION THEOREM).[3]

The great practical importance of this law follows from the fact that we are usually able to exhibit numerous models of the axiom system of a particular theory, without ever leaving the domain of deductive sciences (that is, of mathematics). One way of looking for such models was suggested in the previous section and is of a special interest to us; it depends upon selecting certain constants from some other deductive theory (which can be logic or a theory presupposing logic, and the constants can be primitive or defined), putting them in the axioms in place of the primitive terms, and showing that the sentences which are obtained in this way are asserted statements of the other theory. We say in such a case that we have found an INTERPRETATION OF THE AXIOM SYSTEM OF THE ORIGINAL THEORY WITHIN THE OTHER THEORY.[†] (It may happen, in particular, that the chosen constants belong to the theory originally considered, in which case some of the primitive terms may even remain unchanged; one then says that the given axiom system has a new interpretation within the theory under discussion.)—Let us also subject the theorems of the original theory to an analogous transformation, replacing the primitive terms throughout by those constants that had been employed in the interpretation of the axioms. The law of deduction applies to such a case as well, and we can be sure in advance that the resulting sentences are asserted statements of the other theory. We can formulate our conclusion in the following way:

All theorems which are proved on the basis of a given axiom system remain valid for any interpretation of this system.

In other words, it is redundant to give a special proof for any of these

---

[3]This law was found independently by the French logician J. HERBRAND (1908–1931) and by the author.

[†]The reader should observe that here, as before, we have been dealing with two concepts: that of a model and that of interpretation of one (deductive) theory in another. We emphasize that these concepts are distinct: The former, regarding models, involves particular objects, such as classes, relations, etc., while interpretation depends on asserted statements. In practice, however, these ideas tend to merge, since the interpretation of one theory in another usually suggests a model (for instance, one can take the objects to which the new theory originally referred), while if we start with a model, this may lead directly to an associated theory, and so, to an interpretation of the original theory.—Let us recall an example from the preceding section. There we considered a set $\mathcal{N}$ of numbers and the equivalence relation $\equiv$, and these objects determine a model of the miniature theory. If we now describe this set and this relation by referring to arithmetic, then we are led to an interpretation of the miniature theory within arithmetic.—We may also remark that when discussing models, one sometimes uses the word "*interpretation*" in a different sense. For instance, one might say that the new terms (like "$\mathcal{N}$" and "$\equiv$") provide a new interpretation of the original terms (here, "$\mathcal{S}$" and "$\cong$"). Examples of such usage can be found in footnotes of Chapter VIII.

transformed theorems; it would in any case be a task of a purely mechanical nature, for it would suffice to transfer the corresponding deduction from the domain of the original theory, by subjecting it to the same transformations that had been carried out with respect to the axioms and theorems. Every proof within a deductive theory contains—potentially, so to speak,—an unlimited number of analogous proofs.

* * *

The facts described above illustrate the considerable value of the deductive method from the point of view of economy of human thought. They are also of far-reaching theoretical importance, even if only for the reason that they provide a foundation for various arguments and investigations within the methodology of deductive sciences. In particular, the law of deduction is the theoretical basis for all proofs known as proofs by exhibiting a model or by interpretation; we already encountered one example in the preceding section, and we shall see other examples in the Second Part of this book.

For precision it may be added, that the considerations which we sketched here are applicable to any deductive theory in whose construction logic is presupposed, but their application to logic itself brings about certain complications which we would rather not discuss here. If a deductive theory presupposes not only logic but also other theories, some of the foregoing ideas then require a formulation which is somewhat more involved.

* * *

The common source of the methodological phenomena just discussed is the precept which we emphasized in the preceding section, namely that, in constructing a deductive theory, we disregard the meaning of the axioms and take into account only their form. It is for this reason that when referring to those phenomena, people speak about the purely FORMAL CHARACTER of deductive sciences and of all reasonings within these sciences.

From time to time one finds statements which point to the formal character of mathematics in a paradoxical and exaggerated way; although fundamentally correct, these statements may become a source of obscurity and confusion. Thus one hears and even reads occasionally, that no definite content may be ascribed to mathematical concepts; that in mathematics we do not really know what we are talking about, and that we are not interested in whether our assertions are true. One should approach such judgments rather critically. If, in the construction of a theory, one behaves as if one did not understand the meaning of the terms of this discipline, this is not at all the same as denying that those terms have any meaning. Admittedly, sometimes we develop a deductive theory without ascribing a definite meaning to its primitive terms, thus dealing with the latter as with variables; under such circumstances we say that we treat the theory as a FORMAL SYSTEM. But this kind of situation (which was not taken into account in our general characterization of deductive theories, given in Section 36) arises only if several models or interpretations for the axiom system of this theory are available to us, that is, if we are concerned with several ways of ascribing concrete meaning to the terms occurring in the theory, but we do not desire to give preference in advance to any one of these ways. On the other hand, a formal system for which one could not give a single model would, presumably, be of interest to nobody.

* * *

In conclusion we shall bring to the attention of the reader certain remarkable examples of interpretations of mathematical disciplines, examples which are much more important than those given in Section 37.

An axiom system for arithmetic may be interpreted within geometry: given an arbitrary straight line, it is possible to define certain relations among its points and certain operations on its points so as to satisfy all the axioms—and hence also all theorems—of arithmetic, while in the original formulation these axioms and theorems concern the corresponding relations among numbers and operations on numbers. (This is closely connected with a circumstance which was mentioned in Section 33, namely, with the possibility of establishing a one-to-one correspondence between all points of a line and all numbers.) In a reciprocal manner, axiom systems of geometry have interpretations within arithmetic. The uses to which these two facts can be put are manifold. Geometrical objects, for instance, may be employed in order to give visual images of various facts in the field of arithmetic,—this is the so-called graphical method; on the other hand, it is possible to investigate geometrical features with the help of arithmetical or algebraic methods,—and there is a special branch of geometry, known as analytic geometry, which is concerned with investigations of this type.

Arithmetic, as we have already seen, may be constructed as a part of logic (cf. Section 26). But if we treat arithmetic as a separate deductive theory, resting upon its own system of primitive terms and axioms, its relation to logic can be described as follows: arithmetic admits an interpretation within logic (with the understanding that the axiom of infinity be included in logic,—cf. Section 26); in other words, it is possible to define certain logical concepts which satisfy all the axioms of arithmetic, and hence also all theorems. If we now remember that geometry has an interpretation within arithmetic, we arrive at the conclusion that geometry as well as arithmetic can be interpreted within logic. All these facts are of great significance from the methodological point of view.*

39 Selection of axioms and primitive terms; their independence

We shall now turn to the discussion of a few problems which are somewhat specialized, but which, nevertheless, concern certain fundamental components of the deductive method: namely, the choice of primitive terms and axioms on one hand, and the construction of definitions and proofs on the other.

* * *

It is important to realize that we have a large amount of freedom in selecting the primitive terms and the axioms; it would be quite erroneous to believe that, in a given theory, certain expressions cannot possibly be defined, or that certain statements cannot in any way be derived, and that, therefore, they have to be regarded as primitive terms or as axioms, respectively. This circumstance leads to the following notion: let us call two systems of sentences of a given theory EQUIPOLLENT,† if each sentence of the first system can be derived from the sentences of the second, together with theorems of any theories which precede the given theory, and, conversely, if every sentence of the second system can be derived from the sentences of the first (together with theorems of any preceding theories; any sentences which occur in both systems do not, of course, have to be derived). Let us imagine, further, that a deductive theory has been established on the basis of a certain axiom system, and that in the course of its construction we come across a system of statements which turns out to be equipollent in the sense just described to the original axiom system. (A concrete example can be given in the case of the miniature theory of the congruence of segments, discussed in Section 37: it is easy to show that its axiom system is equipollent to the system of sentences consisting of Axiom 1 together with Theorems 1 and 2.) If a situation of this kind arises, then, from the theoretical point of view, one could reconstruct the entire theory in such a manner that the statements of the new system are taken as axioms, while the former axioms are proved as theorems. Even such a circumstance, that the new axioms may appear at first much less evident, is inessential; for every sentence becomes evident to a certain degree, once it has been derived in a convincing manner from other evident sentences. All this applies likewise—*mutatis mutandis*—to the primitive terms of a deductive theory; the system of these terms may be replaced by any other system of terms of the theory in question, provided that the two systems

† In this discussion the word "*equipollent*" is used in two different meanings. If greater precision is needed, then we say that two systems of sentences can be EQUIPOLLENT IN MEANS OF PROOF, while two systems of primitive terms (cf. below) can be EQUIPOLLENT IN MEANS OF EXPRESSION. Usually, however, it is clear from context which meaning of "*equipollent*" is intended.

are EQUIPOLLENT, in the sense that each term of the first system can be defined by means of the terms of the second together with terms taken from the preceding theories, and vice versa.—When we select a certain system of primitive terms and axioms in preference to other possible equipollent systems, it is not for reasons of theoretical or fundamental character (or, at least, not only for such reasons); other factors play a role here,—practical, didactic, even esthetic ones. Sometimes it is a question of choosing the simplest possible primitive terms and axioms, then again it may be desirable to get along with as few of them as possible; finally, we may prefer such primitive terms and axioms as would enable us, in the simplest possible way, to define those terms and to prove those statements of a given theory in which we are especially interested.

* * *

Another problem relates directly to these remarks. Fundamentally, our goal is to arrive at an axiom system which does not contain any superfluous statements, that is, statements which can be derived from the remaining axioms, and which therefore might be regarded also as theorems of the theory under consideration. An axiom system of this kind is called INDEPENDENT (or, a SYSTEM OF MUTUALLY INDEPENDENT AXIOMS). We likewise attempt to see to it that the system of primitive terms is INDEPENDENT, that is, that it does not contain any superfluous terms, which can be defined by means of the others. Often, however, one does not insist on these methodological postulates (or, principles) for practical, expository reasons, particularly in cases where the omission of a superfluous axiom or primitive term would bring about great complications in the construction of the theory.[††]

40 Formalization of definitions and proofs, formalized deductive theories

The deductive method is justifiably considered the ideal among all methods which are employed in the construction of sciences. It eliminates to a large extent the possibility of obscurities and errors, without resorting to an infinite regress; and thanks to this method, any uncertainties regarding the content of concepts or the truth of assertions of a given theory are considerably reduced; these uncertainties concern at most the few primitive terms and axioms.

[††] As we see, the author uses the word "*postulates*" in the sense of "*guiding principles*" or "*specifications*". Some of the postulates may require a strict adherence, while some others, as we have just seen, less so.

One reservation has to be added to this statement, however. The application of the deductive method will give the desired results only if all definitions and all proofs fulfill their tasks completely, that is, if the definitions make clear, beyond doubt, the meaning of all the terms to be defined, and if the proofs convince us fully of the validity of all the theorems which were to be proved. It might not be easy to determine whether particular definitions and proofs actually comply with these requirements; it may well be, for instance, that an argument which seems entirely convincing to one person is not even comprehensible to another. In order to eliminate any uncertainty in this respect, the present-day methodology endeavors to replace subjective scrutiny of definitions and proofs by criteria of an objective nature, in such a way that the decision regarding the correctness of given definitions or proofs would depend exclusively upon their structure, that is, upon their exterior form. For this purpose, special RULES OF DEFINITION and RULES OF PROOF (or, OF INFERENCE) are introduced. The first tell us what form the sentences should have which are used as definitions in the theory under consideration, and the second describe the kind of transformations to which statements of this theory may be subjected in order to derive other statements from them; each definition has to be constructed in accordance with the rules of definition, and each proof must be COMPLETE, that is, it must consist of successive applications of rules of proof to sentences which were previously recognized as true (cf. Sections 11 and 15).—These new methodological postulates may be designated as postulates of the FORMALIZATION OF DEFINITIONS AND PROOFS; a discipline constructed in accordance with these new specifications is called a FORMALIZED DEDUCTIVE THEORY.[4]

* * *

*Through the postulates of formalization, the formal character of mathematics is enhanced considerably. Already at an earlier stage of our inquiry into the deductive method we were supposed to disregard the meaning of all expressions which are specific to a discipline that was being constructed, and we were to behave as if in place of these expressions we had variables which are void of any independent significance. But at least, the concepts of logic then retained their customary meanings. The axioms and the theorems of a mathematical discipline could therefore be treated, if not as sentences, then at least as sentential functions, that is, as expressions having the grammatical form of sentences and expressing certain properties of objects or relations among objects. To derive a theorem from accepted ax-

[4]The first attempts to present deductive theories in a formalized form are due to FREGE, who has already been quoted twice (cf. footnote 2 on p. 18, and also footnote 4 on p. 76). A very high level of formalization was achieved in the works of the Polish logician S. LEŚNIEWSKI (1886–1939); one of his achievements is a precise and exhaustive formulation of the rules of definition.

ioms (or from theorems which were proved previously) was then the same as to show convincingly that all objects which satisfy the axioms also satisfy the theorem in question; and under such circumstances mathematical proofs would not altogether differ very much from considerations of everyday life. Now, however, the meaning of all the terms which are encountered in the given discipline is to be disregarded without exception, and in developing the deductive theory we are supposed to think of sentences as configurations of signs, void of any content; each proof of a formalized theory consists of subjecting axioms or previously proved theorems to a series of purely external transformations.*

* * *

In the light of modern requirements, logic has become the basis of mathematical sciences in a much more substantial way than it used to be. We may no longer be satisfied with the conviction that—due to our innate or acquired capacity for correct thinking—our argumentations are in accordance with the rules of logic. In order to give a complete proof of a theorem, it is necessary to apply the transformations that are prescribed by the rules of proof not only to the statements of the theory with which we are concerned, but also to those of logic (and of other preceding theories); and we clearly need to have at our disposal a list of logical laws which is adequate for constructing the proofs.

It is only by virtue of the development of deductive logic that, theoretically at least, we are able today to present every mathematical discipline in formalized form. In practice, however, this still involves considerable complications; a gain in exactitude and methodological correctness is accompanied by a loss in clarity and intelligibility. The whole problem, after all, is fairly new, the relevant investigations are not yet definitively concluded, and there is reason to hope that further progress will eventually bring about essential simplifications.† It would therefore be premature to comply fully with the postulates of formalization at the present time, if one deals with a popular presentation of any branch of mathematics. In particular, it would not be reasonable to demand that the proofs in an ordinary

†Improvements and simplifications of various kinds have certainly been made since the time of the early editions of *Introduction to Logic*. Nevertheless, a formalized presentation of a mathematical discipline would still retain the drawbacks which are mentioned: considerable complications, loss of intelligibility. Such a presentation could serve some special purposes, but now as in former days, it would hardly be suitable for an introduction to the discipline in question.—We should like to note here that, besides eliminating uncertainties (as discussed in the text), formalization plays also another important role. In fact, various methodological problems (such as those which will be considered in Sections 41–42 and 59–60) require for their investigation a scrutiny of the following question: Just what is a proof? Or, in other words: What are its defining properties? Evidently, such a question can be answered in a precise way only when one is dealing with formalized theories.

textbook of some mathematical discipline be given in complete form; one should, however, expect the author of a textbook to be intuitively certain that all of his or her proofs can be brought into that form, and even to carry the considerations to such a point, that a reader who has some experience in deductive thinking and sufficient knowledge of contemporary logic would be able to fill the remaining gaps without much difficulty.

41 Consistency and completeness of a deductive theory; the decision problem

We shall now consider two methodological concepts which are of great importance from the theoretical point of view, while in practical matters they are of little significance. They are the concepts of CONSISTENCY and COMPLETENESS.

A deductive theory is said to be CONSISTENT or NON-CONTRADICTORY if no two asserted statements of this theory contradict each other, or, in other words, if of any two contradictory sentences (cf. Section 7) at least one cannot be proved. A theory is called COMPLETE, on the other hand, if of any two contradictory sentences which are formulated by employing exclusively the terms of the theory under consideration (and of the theories preceding it), at least one sentence can be proved within this theory. Now, if a sentence is such that its negation can be proved in a given theory, one usually says that it can be DISPROVED in that theory. With the help of this terminology we can say, that a deductive theory is consistent if no sentence can be both proved and disproved within it; a theory is complete if every sentence which is formulated by employing the terms of this theory can be proved or disproved within it. Both terms "*consistent*" and "*complete*" are commonly applied not only to the theory itself, but also to the axiom system upon which the theory is based.

Let us try to get a clear idea of the import of these two notions. Every discipline, even one which is constructed entirely correctly in every methodological respect, loses its value in our eyes if we have reason to suspect that not all assertions of this discipline are true. On the other hand, the value of a discipline will be greater, if more true sentences can be derived within it. From this point of view, a discipline might be considered ideal, if it contains among its asserted statements all true sentences which are from the domain in question, and not a single false one. A sentence is here considered as being "from the domain in question" if it is formulated by employing only the terms of the discipline under consideration and the terms of its preceding disciplines; after all, one cannot expect to prove in arithmetic, say, all true sentences containing concepts of chemistry or biology.—Let us now imagine that a deductive theory is inconsistent, that is to say, that two contradictory sentences occur among its axioms and theorems; from the law of contradiction (cf. Section 13) it follows that one of these sentences must be false. On the other hand, if the theory is incomplete, then there exist two contradictory sentences (from the domain in question), neither of which can be proved within that discipline; and yet, by another logical law, namely by the law of excluded middle, one of the two sentences must be true. We see therefore that a deductive theory certainly falls short of our ideal unless it is both consistent and complete. (We are not implying here that every consistent and complete discipline must, *ipso facto*, be a realization of our ideal, so that its axioms and theorems necessarily include all true sentences from the domain in question, and only such sentences; for, also a theory with a false axiom might be consistent and complete.)

There is yet another way of looking at the notions which we have been considering. The development of any deductive science consists in the formulation of problems which are of the type "*is such and such the case?*" by employing the terms of this science, and then, in attempting to resolve these problems on the basis of the axioms that have been assumed. Any problem of this type may clearly be resolved in one of two possible ways: in the affirmative or in the negative. In case of the first alternative, the answer runs: "*such and such is the case*"; and in case of the second: "*such and such is not the case*". The consistency and the completeness of the axiom system of a deductive theory now give us a guarantee, that every problem of the kind here described can actually be resolved within the theory, and moreover resolved in one way only; the consistency excludes the possibility that any problem may be resolved in two ways, both affirmatively and negatively, while the completeness assures us that it can be resolved in at least one way.

* * *

Closely connected with the problem of completeness is another problem, a more general one, which concerns incomplete as well as complete theories. This problem depends on finding, for a given deductive theory, a general mechanical method, that is, an ALGORITHM, which would enable us to establish whether or not any particular sentence (from the domain in question, cf. above) can be proved within that theory. One then speaks of DECIDING whether the sentence can be proved or not, and this important problem is known as the DECISION PROBLEM[5][†]; theories for which this

[5]The significance of the concepts discussed in this section—and especially the import of the consistency proofs and of the decision problem—were emphasized by HILBERT (cf. footnote 2 on p. 111), who greatly stimulated many important investigations into the foundations of mathematics. Upon his instigation, these concepts and problems became the subject of intensive research by a number of mathematicians and logicians.

[†]In footnote 5 above, the author speaks of a general decision problem for mathematics. Indeed, early methodological studies led to intriguing questions like the following: Will one be able to find an algorithm for solving problems of all kinds in mathematics?

problem has a positive solution are said to be DECIDABLE.††

* * *

There are only a few deductive theories which are known to be consistent and complete. They are, as a rule, elementary theories of a simple logical structure and of a modest stock of concepts. An example is given by sentential calculus, which was discussed in Chapter II, provided that it is considered as an independent deductive theory and not as a part of logic (however, when the term "*complete*" is applied to this theory, it is to be understood in a slightly modified sense). Perhaps the most interesting example of a consistent and complete theory is that of elementary geometry; we have here in mind the geometrical theory which is limited to those confines, within which it was taught in schools for centuries as a part of elementary mathematics; in other words, it is a discipline in which the properties of various special kinds of geometrical figures such as lines, planes, triangles, and circles are investigated, but into which the general concept of a geometrical configuration (a point set) does not enter.[6]

The situation changes in an essential way as soon as one goes over to such sciences as arithmetic or non-elementary geometry. Probably no one working in these sciences doubts their consistency; and yet, as methodological investigations have shown, any attempt to prove their consistency will necessarily involve great difficulties of a fundamental nature. The situation with regard to the problem of completeness is even worse; it turns out that arithmetic and non-elementary geometry are incomplete; for it has been possible to set up problems, of a purely arithmetical or geometrical character, that can be neither positively nor negatively resolved within these disciplines. One might suppose that this fact merely reflects the imperfection of the axiom systems and of the methods of proof which are at our disposal at the present time, and that a suitable modification (for instance, an extension of the axiom system) may yield complete systems in the future. Certain profound methodological studies, however, have shown that these theories are incomplete in a very deep sense: *it will never be possible*

Will such problems therefore become like exercises in calculation with numbers, not requiring any creativity? We now know that such an algorithm cannot be constructed; cf. the subsequent text. Accordingly, the general decision problem has given way to separate decision problems for diverse mathematical disciplines.

††Consideration of the decision problem leads us to a different subject. This problem by its nature requires the investigation of various algorithms, or, mechanical procedures; but algorithms of all kinds are of interest when one works with computers. In fact, there are various links and analogies between methodological studies and computer science, and decision problems provide an example.

[6]For the first proof of completeness of sentential calculus (and therefore, for the first positive result dealing with completeness) we are indebted to the American logician E. L. POST (1897–1954). The proof of completeness (and of decidability) of elementary geometry originates with the author.

to construct a consistent and complete deductive theory which contains as its theorems all true sentences of arithmetic or of non-elementary geometry. Moreover, it turns out that the decision problem likewise does not admit a positive solution in the case of these disciplines; it is impossible to set up a general method which would allow us to distinguish mechanically between those sentences which can be proved within these disciplines and those which cannot be proved.—All of these results extend to many other deductive theories, and in particular, to all theories which presuppose the arithmetic of integers or that of natural numbers (that is, the theory of the four basic arithmetical operations on the respective numbers) or contain ingredients which are adequate for developing such an arithmetic. *For instance, these results can be applied to the general theory of classes (as follows from the discussion in the latter part of Section 26).*[7] [‡]

In view of these last remarks, it is understandable that the concepts of consistency and completeness—notwithstanding their theoretical importance—exert little influence in practice upon the construction of deductive theories.[‡‡]

42 The widened conception of the methodology of deductive sciences

The investigations of consistency and completeness were especially important in bringing about a major expansion of the domain, and of the scope,

[7]These exceedingly important achievements are due to the American logician (of Austrian origin) K. GÖDEL (1906–1978). His results concerning the decision problem were later extended by another American logician, A. CHURCH (1903–).

[‡]The following recapitulation may be helpful. Incompleteness and the related results were first established for the arithmetic of natural numbers (and they extend easily to the integers). Since this arithmetic can be developed within the general theory of sets (or within that of classes), these results apply here as well. Furthermore, these results apply to theories which presuppose the theory of sets; they are the so-called non-elementary theories, and non-elementary geometry is an example.

[‡‡]This simple conclusion, that the two concepts exert little influence in practice, should perhaps be qualified. This conclusion certainly applies when one attempts a new construction for an established theory, and when the belief in its consistency is supported by years of experience, even if no proof is available. But the situation changes when a new theory is suggested, and when the experience is lacking. In case of such a theory, a novice might suppose that one could tolerate a few sentences which are asserted together with their negations. However, a basic methodological law (cf. exercise *11) then shows that in an inconsistent theory every sentence is asserted, so that the theory is devoid of meaning.—On the other hand, incompleteness by itself does not disqualify a theory. Therefore, when constructing a new theory, one should take utmost care to avoid an inconsistency, but one should not be upset if the resulting theory turns out to be incomplete.

of methodological studies; these investigations even led to a fundamental change in the whole character of the methodology of deductive sciences. That conception of methodology which was described at the beginning of the present chapter became, in the course of the historical development of the subject, too narrow. The analysis and the critical evaluation of the methods, which are applied in the construction of deductive sciences, ceased to be the exclusive or even the main task of methodology. *The methodology of the deductive sciences became a general theory of deductive sciences, in a sense analogous to that, in which arithmetic is the theory of numbers and geometry is the theory of geometrical figures.* In contemporary methodology we investigate deductive theories in their entirety as well as individual sentences which constitute them; we consider the symbols and the expressions of which such sentences are composed, properties and sets of expressions and of sentences, relations holding among them (such as the relation of consequence), and even relations among expressions and the objects which the expressions "talk about" (such as the relation of designation); we establish general laws which govern these concepts.†

* * *

*In this connection one should observe the following: terms which denote expressions occurring in deductive theories, as well as terms denoting properties of these expressions, or relations among them, belong to the methodology of deductive sciences, but not to the domain of logic. This applies in particular to various terms which were introduced and employed in the previous chapters of this book, such as "*variable*", "*sentential function*", "*quantifier*", "*consequence*", and many others. In order to make clearer to ourselves the difference between logical and methodological terms, let us consider such a pair of words as "*or*" and "*disjunction*". The word "*or*" belongs to sentential calculus—and therefore to logic—, although it is also used in all other sciences, and thus in particular in methodology. The word "*disjunction*", on the other hand, denotes a sentence which is constructed with the help of the word "*or*", and is a typical instance of a methodological term.

The reader will perhaps be surprised when he or she realizes that in the chapters which dealt with logic we employed so many methodological terms. The explanation of this, however, is relatively simple. On the one hand, a certain circumstance plays a role here, as we already indicated in Section 9: there is a widespread custom among logicians as well as mathematicians— sometimes for purely stylistic reasons—of using phrases

†It may be mentioned here, that the author's celebrated investigation of truth finds its natural place as a part of such methodology. Indeed, when dealing with the problem of truth, we are concerned with relating expressions (in this case, sentences) and the objects to which the expressions refer, or, which they "talk about".

which contain methodological terms and which are intended as synonyms for expressions of a purely logical or mathematical character; to some extend we have complied with this custom in the present work. But, on the other hand, a more important factor is also involved here; we have not attempted in this book so far to construct logic in a systematic way, but rather, we have talked about logic and have discussed and commented on its concepts and on its laws. We know however (from Section 18) that when talking about logical expressions, we must use the names of these expressions, and these names are terms which belong to methodology. Should we develop logic in the form of a deductive theory, without making any comments on it at all, then methodological terms would occur only in the formulation of rules of definition and inference.*

* * *

In the course of the evolution through which methodology had passed, there arose a need for applying new, more subtle, and more precise methods of inquiry in this field. Methodology took on the form of a deductive discipline—and in fact, became like the sciences which constitute its primary concern. Moreover, in view of the extended domain of investigations, the expression "*the methodology of deductive sciences*" itself ceased to appear appropriate enough; indeed, "*methodology*" means, essentially, "*the science of method*". Consequently this expression is now often replaced by others—especially by the terms "METALOGIC" and "METAMATHEMATICS", which mean about the same as "*the science of logic*" and "*the science of mathematics*". Another term has also been used, "SYNTAX AND SEMANTICS OF DEDUCTIVE SCIENCES", which stresses the analogy between the methodology of deductive sciences and the grammar and the interpretation of everyday language.[8] ††

---

[8]Methodology of the deductive sciences in the widened sense is a discipline which is younger than the others considered in this book. Its intensive development began only after 1920—simultaneously (and, it seems, independently) in two different centers: in Göttingen, under the influence of HILBERT and of the Swiss logician P. BERNAYS (1888–1977), and in Warsaw where LEŚNIEWSKI and ŁUKASIEWICZ worked (cf. footnotes 2 on p. 111, 4 on p. 123, and 2 on p. 18). The fundamental work of HILBERT and BERNAYS in this domain is *Grundlagen der Mathematik* (Berlin 1934, 1939). The conceptual apparatus and philosophical aspects of the new discipline are discussed in the writings of the American philosopher and logician (of German origin) R. CARNAP (1891–1970); his earliest work in this direction is *Logische Syntax der Sprache* (Vienna 1934).

††The methodological sciences (metamathematics, metalogic) have been growing steadily, and now several branches are recognized, among them: THEORY OF MODELS and PROOF THEORY. The first deals with realizations, or, models of axiom systems (one can say, with semantics), and the second, with problems of deduction (and so, with matters of form and syntax).

Exercises

1. Exhibit several interpretations within arithmetic and within geometry of the axiom system considered in Section 37.

Is the set of all numbers, together with the relation *less than* among numbers, a model of this axiom system? Is the set of all straight lines and the relation of parallelism among lines such a model?

2. In that fragment of geometry which was discussed in Section 37, the relation of being shorter among segments can be defined in the following way:

> *we say that x is shorter than y*, in symbols: $x < y$, *if x and y are segments and if x is congruent to a segment which is a part of y;* in other words, *if $x \in \mathcal{S}$, $y \in \mathcal{S}$, and if there exists an object z such that $z \in \mathcal{S}$, $z \subseteq y$, $z \neq y$, and $x \cong z$.*

Distinguish in this sentence between the definiendum and the definiens; determine the disciplines (or the branches of logic, as the case may be) to which the terms occurring in the definiens belong. Does this definition comply with the general methodological principles of Section 36 and with the rules of definition of Section 11?

3. Is the proof of Theorem 1, as given in Section 37, a complete proof, if only those rules of proof are taken into consideration that were stated in Section 15?

4. In addition to Theorems 1 and 2, the following theorems can be derived from the axioms of Section 37:

THEOREM 3. *For any elements x, y, and z of the set $\mathcal{S}$, if $x \cong y$ and $x \cong z$, then $y \cong z$.*

THEOREM 4. *For any elements x, y, and z of the set $\mathcal{S}$, if $x \cong y$ and $y \cong z$, then $z \cong x$.*

THEOREM 5. *For any elements x, y, z, and t of the set $\mathcal{S}$, if $x \cong y$, $y \cong z$, and $z \cong t$, then $x \cong t$.*

Give a careful proof that the following systems of sentences are equipollent, in the sense established in Section 39, to the system consisting of Axioms 1 and 2 (and that each could therefore be chosen as a new axiom system):

(a) the system consisting of Axiom 1 and Theorems 1 and 2;
(b) the system consisting of Axiom 1 and Theorem 3;
(c) the system consisting of Axiom 1 and Theorem 4;
(d) the system consisting of Axiom 1 and Theorems 1 and 5.

5. Along the lines of the remarks which were made in Section 37, formulate general laws of the theory of relations that represent a generalization of the results obtained in the preceding exercise.

Hint: These laws may be given in the form of equivalences, which could begin with the words:

> *for a relation R to be reflexive and to have the property $\mathcal{P}$ in a class K, it is necessary and sufficient that....*

6. Consider system (a) of sentences, as given in exercise 4. Exhibit models satisfying

(a) the first two sentences of the system, but not the last;
(b) the first and the third sentence, but not the second;
(c) the last two sentences, but not the first.

What conclusion may be drawn from the existence of such models, with respect to the possibility of deriving any one of the three sentences from the other two? Are these sentences mutually independent? (Cf. Sections 37 and 39.)

7. There have occasionally been complaints to the effect that there is a certain discrepancy among various textbooks of geometry, inasmuch as sentences which are treated as theorems in some textbooks are adopted as axioms, and thus without proof, in others. Are these complaints justified?

***8.** In Section 13 we became acquainted with the method of truth tables, which enables us to decide whether any given sentence of sentential calculus is true, and whether it may therefore be accepted as a law of this calculus. When applying this method, we may entirely forget about the meaning which was ascribed to the symbols "T" and "F" occurring in the truth tables; instead, we may assume that this method reduces to applying two rules in the construction of sentential calculus, where the first rule is close to the rules of definition, and the second, to the rules of proof. According to the first rule, if we want to introduce into sentential calculus a constant term, we have to begin by constructing the fundamental truth table for the simplest sentential function containing this term, and containing as many distinct variables as this term can affect. According to the second rule, if we want to accept a sentence as a law of sentential calculus (where the sentence contains only those constants for which the fundamental truth tables have already been constructed), we must construct the derivative truth table for this sentence, and verify that only the symbol "T" occurs in the last column of this table.

When sentential calculus is constructed exclusively by means of these two rules, it assumes a character which is close to that of formalized deductive theories. Justify this statement on the basis of the considerations of Section 40. Notice, however, some differences between this method of constructing sentential calculus and the general principles for constructing deductive theories, which were discussed in Section 36. If the method of truth tables is adopted, is it possible to distinguish in sentential calculus between primitive and defined terms? What other distinction is then lost?

***9.** The method of truth tables, as it was described in the preceding exercise, allows us to introduce into sentential calculus new terms which were not discussed in Chapter II. We can, for instance, introduce the symbol "$\triangle$", such that the sentential function:

$$p \triangle q$$

is considered as the symbolic form of the expression:

neither p nor q.

Construct the fundamental truth table for this function, which would correspond to the intuitive meaning which has been assigned to the symbol "$\triangle$"; then verify with the help of derivative truth tables, that the following sentences are true and may therefore be accepted as laws of sentential calculus:

$$\sim p \leftrightarrow (p \triangle p),$$
$$(p \vee q) \leftrightarrow [(p \triangle q) \triangle (p \triangle q)],$$
$$(p \rightarrow q) \leftrightarrow \{[(p \triangle p) \triangle q] \triangle [(p \triangle p) \triangle q]\}.$$

***10.** The method of constructing sentential calculus which is discussed in exercise *8 provides an immediate solution of the decision problem (cf. Section 41) for this calculus, and enables us to show very easily that sentential calculus is a consistent theory. How can these features be shown?

***11.** One of the laws of sentential calculus is the following:

For any p and q, if p and not p, then q.

On the basis of this logical law, establish the following methodological law (cf. exercise *20 of Chapter II regarding a rule for conjunctions):

> *If the axiom system of a deductive theory which presupposes sentential calculus is inconsistent, then every sentence which is formulated by employing the terms of this theory can be derived from that system.*

***12.** It is known that the following methodological law holds:

> *If the axiom system of a deductive theory is complete, and if any sentence which can be formulated but not proved within that theory is added to the system, then the axiom system extended in this manner is no longer consistent.*

Why is this the case?

***13.** Single out all the terms appearing in Chapter II which belong to the field of the methodology of deductive sciences, according to the remarks made in Section 42.

* * *

Supplementary exercises.—The following exercises, numbered S1–S19, are intended to give the reader a glimpse at formalized theories and at complete proofs.† We consider here three theories: the theory of congruent line segments, a fragment of the algebra of classes, and sentential calculus.

In case of the first two, we assume sentential calculus as the only preceding theory. This calculus will then be regarded as a theory which is determined by truth tables; cf. exercise *8. (We emphasize this point, since the exercises of the last group deal with an alternative formulation of this calculus.)—There will be three rules of inference, allowing: (i) substitution of sentential functions for sentential variables, (ii) substitution of an INDIVIDUAL VARIABLE by another (these are the variables that refer to the individual objects under discussion, that is, to line segments and to classes in the two examples), and (iii) detachment. For simplicity we shall disregard the universal quantifiers; a way of handling them was explained in Section 15.

The problem in each of the following exercises is to prove the theorem which is given (unless otherwise stated). In case of the first two theories, the proofs require various laws of sentential calculus with which the reader might not be familiar, and which are usually given in the "hints". The reader should convince himself or herself that these statements are indeed valid, even if he or she does not make a detailed verification.

* * *

In the first exercises we describe the theory of congruent segments, along the line of Section 37 and exercises 3 and 4. The only primitive term is "$\cong$" (except for terms of sentential calculus, and for the symbol "S" of the universe of discourse, but the latter does not enter explicitly into the formulas). There are two axioms:

AXIOM 1. $x \cong x$.
AXIOM 2. $(x \cong z \land y \cong z) \to x \cong y$.

S1. Prove:

(a) $y \cong z \to z \cong y$;
(b) $y \cong z \leftrightarrow z \cong y$.

†These exercises (which were arranged by the editor) are offered as a replacement for several exercises which were included in the previous editions of the book but are now eliminated.—For reference we note the numbers of eliminated exercises: 1–3, 13–16, and a part of no. 17 (out of 20). The remaining exercises were renumbered.

Hint: For (a), substitute "z" for "x" in Axioms 1 and 2 and use the law:

$$[(p \wedge q) \rightarrow r] \rightarrow [p \rightarrow (q \rightarrow r)].$$

Substitute here "$z \cong z$" for "p", and appropriate expressions for "q" and "r" (so that "$q \rightarrow r$" becomes the law to be proved), and detach in turn the transformed Axioms 2 and 1. For (b), substitute in (a) "z" for "y" and vice versa. This will give the converse of (a). Then use the law:

$$(p \rightarrow q) \rightarrow [(q \rightarrow p) \rightarrow (p \leftrightarrow q)].$$

Substitute for "p" and "q" and detach: first (a), then its converse.

S2. $(x \cong y \wedge y \cong z) \rightarrow x \cong z.$

Hint: First we want to replace "$y \cong z$" by "$z \cong y$" in Axiom 2. Intuitively, this is justified by the equivalence S1(b). A complete proof depends on the following law:

$$(p \leftrightarrow q) \rightarrow \{[(r \wedge p) \rightarrow s] \rightarrow [(r \wedge q) \rightarrow s]\}.$$

Substitute for the sentential variables so that "$p \leftrightarrow q$" becomes the expression in S1(b), and "$(r \wedge p) \rightarrow s$", the expression in Axiom 2. Then detach twice. After the detachments make a change of individual variables.

Note: First, the reader should find it instructive to compare the proofs of exercise S1 with the proof of Theorem 1 in Section 37. Second, the reader should take a close look at the proof of S2, especially at the law of sentential calculus. We have here a method of justifying the replacement of one equivalent expression by another, and there will be opportunities to utilize this method also in the exercises that follow.

For additional exercises, the reader can formulate Theorems 3–5 of exercise 4 in symbols, and try to construct complete proofs.

* * *

For our second theory, we refer to a universe of discourse which is a certain collection of non-null classes; this collection itself is a class which we can denote by, say, "$\mathcal{C}$". There are two other primitive terms: "$\subseteq$" (meaning: "*is included in*") and ")(" (meaning: "*is disjoint from*"). We assume five axioms:

AXIOM I. $K \subseteq K.$

AXIOM II. $(K \subseteq L \wedge L \subseteq M) \rightarrow K \subseteq M.$

AXIOM III. $M \,)(\, N \rightarrow N \,)(\, M.$

AXIOM IV. $(K \subseteq L \wedge L \,)(\, M) \rightarrow K \,)(\, M.$

AXIOM V. $K \,)(\, L \rightarrow \sim K \subseteq L.$

The reader may find it helpful to review the discussion of the two relations in Section 24, and to sketch geometrical figures which illustrate these axioms.[††]

S3. Prove:

(a) $M \)(N \leftrightarrow N \)(M$;
(b) $(K \subseteq L \wedge M \)(L) \rightarrow K \)(M$;
(c) $L \)(K \rightarrow \sim K \subseteq L$.

Hint: For (a), use Axiom III, and the proof goes the same way as the derivation of S1(b) from S1(a). For (b) and (c), use (a) with substitution, Axioms IV and V, and argue as in the proof of S2.

S4. $K \)(L \rightarrow (\sim K \subseteq L \wedge \sim L \subseteq K)$.

Hint: Use the law:

$$(p \rightarrow q) \rightarrow \{(p \rightarrow r) \rightarrow [p \rightarrow (q \wedge r)]\}.$$

Substitute so as to detach first Axiom V, and then a formula which is obtained by substitution from S3(c).

S5. (a) $K \subseteq L \rightarrow \sim K \)(L$;
(b) $\sim K \)(K$.

Hint: For (a), use Axiom V and the following variant of the law of contraposition (cf. Section 14):

$$(p \rightarrow \sim q) \rightarrow (q \rightarrow \sim p).$$

For (b), use (a) and one of the axioms.

S6. (a) $(K \subseteq L \wedge L \subseteq M) \rightarrow \sim K \)(M$;
(b) $(K \subseteq L \wedge L \)(M) \rightarrow \sim K \subseteq M$.

Hint: For (a), start with Axiom II. Make a substitution in S5(a) so that its hypothesis will be the same as the conclusion of Axiom II. Use the law of the hypothetical syllogism (in a form which differs from that of Section 13):

$$(p \rightarrow q) \rightarrow [(q \rightarrow r) \rightarrow (p \rightarrow r)].$$

Substitute for "q" the common conclusion and hypothesis, find appropriate substitutions for "p" and "r", and complete the proof by detachments. For (b), start with Axioms IV and V, and proceed as in (a).

[††] We remark that disjointness is very seldom regarded as an independent relation. In fact, one would normally write "$K \cap L = \emptyset$" rather than "$K \)(L$" (even though the latter statement is somewhat stronger). However, disjointness occurs as a relation in the categorical syllogisms of ARISTOTLE (cf. footnote 1 on p. 18), and two of the axioms are just the syllogisms of Section 24.—The five axioms are mutually independent. See exercise S8 and the additional discussion in the companion volume.—Furthermore, one generally assumes that the universe of discourse is non-empty, but such an assumption would play no role in this theory or in the theory of congruent segments.

We now introduce a new symbol through a definition. This definition will be constructed in accordance with the rules and the considerations of Sections 11 and 36, and in particular, will be given as an equivalence:

DEFINITION I. $K \between L \leftrightarrow [\sim K \,)(\, L \;\wedge (\sim K \subseteq L \wedge \sim L \subseteq K)]$.

The new symbol denotes the relation of overlapping. We note three evident properties of this relation:

S7. Prove:

(a) $K \between L \leftrightarrow L \between K$;
(b) $\sim K \between K$;
(c) $(K \between L \vee K \,)(\, L) \vee (K \subseteq L \vee L \subseteq K)$.

Hint: For (a), one should exploit equivalence relations in order to interchange "K" and "L" in the definition. First use S3(a) (with substitution) and an argument as in S2 to show:

$$K \between L \leftrightarrow [\sim L \,)(\, K \wedge (\sim K \subseteq L \wedge \sim L \subseteq K)].$$

Then use "$(p \wedge q) \leftrightarrow (q \wedge p)$" and a similar argument to show:

$$K \between L \leftrightarrow [\sim L \,)(\, K \wedge (\sim L \subseteq K \wedge \sim K \subseteq L)].$$

A substitution in the definition now gives an equivalence which relates "$L \between K$" and the right side. Complete the proof with the help of the following law:

$$(p \leftrightarrow r) \rightarrow [(q \leftrightarrow r) \rightarrow (p \leftrightarrow q)],$$

where "$K \between L$" is to be substituted for "p", and "$L \between K$" for "q". For (b), substitute "K" for "L" in the definition and observe that "$\sim K \subseteq K$" contradicts Axiom I. Construct a law of sentential calculus so as to be able to detach Axiom I and the definition after substitution, and to obtain the desired conclusion. (Verify your law.) For (c) you need only the definition and a suitable law of sentential calculus.

The above fragment of the algebra of classes can be developed further by introducing the following definition (of "*proper inclusion*", or, of "*is a proper subclass of*"):

DEFINITION II. $K \subset L \leftrightarrow (K \subseteq L \wedge \sim L \subseteq K)$.

The ambitious reader could now formulate and try to prove a few properties of proper inclusion. One immediate property is: $\sim K \subset K$.

S8. Substitute in Axioms I–V the symbol "$\subset$" of proper inclusion for "$\subseteq$". Verify (e.g. by arguments like those which are indicated in connection with exercise 13 of Chapter IV) that Axioms II–V remain valid in the given class C, but not Axiom I. What can you conclude about the derivability of Axiom I?

* * *

The third theory that we present here is the sentential calculus. This calculus will now be formulated as a deductive theory, which is fully in accord with the principles of Sections 36 and 40, and which does not depend on truth tables.—For the primitive terms we take the five usual connectives in symbolic form: "$\rightarrow$", "$\sim$", "$\leftrightarrow$", "$\vee$", and "$\wedge$". The rules of inference will be those of detachment and substitution. (Substitution here means: substituting a sentential variable by another sentential variable, or by a compound sentential function. This rule should be formulated more precisely, with parentheses taken into account. It would not be difficult to give such a formulation, but we shall not go into the details.)

We select a system of 15 axioms. The content of most of them should be familiar to the reader. They split naturally into five groups of three axioms each[‡]:

IA. $p \rightarrow (q \rightarrow p)$.
IB. $[p \rightarrow (q \rightarrow r)] \rightarrow [(p \rightarrow q) \rightarrow (p \rightarrow r)]$.
IC. $(p \rightarrow q) \rightarrow [(q \rightarrow r) \rightarrow (p \rightarrow r)]$.

IIA. $\sim\sim p \rightarrow p$.
IIB. $p \rightarrow \sim\sim p$.
IIC. $(p \rightarrow q) \rightarrow (\sim q \rightarrow \sim p)$.

IIIA. $(p \leftrightarrow q) \rightarrow (p \rightarrow q)$.
IIIB. $(p \leftrightarrow q) \rightarrow (q \rightarrow p)$.
IIIC. $(p \rightarrow q) \rightarrow [(q \rightarrow p) \rightarrow (p \leftrightarrow q)]$.

IVA. $p \rightarrow (p \vee q)$.
IVB. $q \rightarrow (p \vee q)$.
IVC. $(p \rightarrow r) \rightarrow \{(q \rightarrow r) \rightarrow [(p \vee q) \rightarrow r]\}$.

VA. $(p \wedge q) \rightarrow p$.
VB. $(p \wedge q) \rightarrow q$.
VC. $(p \rightarrow q) \rightarrow \{(p \rightarrow r) \rightarrow [p \rightarrow (q \wedge r)]\}$.

We shall now explore the consequences of the various groups of axioms in turn. We start with the properties of negation, while regarding the axioms of the first group (especially Axiom IC, the law of the hypothetical syllogism) as auxilliary.

S9. Prove: $(p \rightarrow \sim q) \rightarrow (q \rightarrow \sim p)$.

[‡]The reader may recall the remarks of Section 39 about including a superfluous axiom for expository reasons. Here Axiom IC can be derived from IA–B, and so is superfluous. We shall see that this axiom will be used extensively in the proofs which follow, while its derivation is somewhat intricate. However, if this axiom is deleted, then we obtain a system of mutually independent axioms.—The axiomatic approach to sentential calculus originates with FREGE; as indicated in footnote 4 on p. 123, one finds in his work the first systematic use of complete proofs. The present axiom system is adapted from one due to BERNAYS (see footnote 8 on p. 130).

Hint: Make a substitution in IC so that the first implication is: "$q \to \sim\sim q$", and so that "r" is changed to "$\sim p$". Substitute in IIB and detach; the result is:

$$(\sim\sim q \to \sim p) \to (q \to \sim p).$$

Next, substitute in IIC so as to obtain:

$$(p \to \sim q) \to (\sim\sim q \to \sim p).$$

We see that the conclusion agrees with the previous hypothesis. So invoke syllogism and make another application of Axiom IC, with "$p \to \sim q$" substituting "p" and "$\sim\sim q \to \sim p$" substituting "q".

S10. $p \to p$.

Hint: A simple proof depends on Axioms IIA–B and on the law of syllogism (IC). In the latter, substitute "p" for "r" and "$\sim\sim p$" for "q", and detach twice.

Note: The reader should now be able to recognize situations where the conclusion of one asserted statement (e.g. of a theorem) agrees with the hypothesis of another, so that Axiom IC could be applied. He or she will find a few more such applications of this axiom in subsequent proofs.— There is another interesting aspect of the theorem just given. Its statement does not involve the negation sign, and one might think that it could be proved by using only the axioms of the first group. In this case such a proof is indeed possible, and is not difficult. The first step is to substitute "p" for "r" in IB, and the rest is left for the reader.

We turn to the third group of axioms. We shall be concerned at first with proving the equivalence when an implication and its converse are available. Then we shall see two examples which illustrate the replacement of equivalent expressions. (Cf. the note following S2.)

S11. (a) $p \leftrightarrow \sim\sim p$;
(b) $(p \to \sim q) \leftrightarrow (q \to \sim p)$.

Hint: Use Axiom IIIC. In addition, for (a), use two other axioms. For (b) use a theorem and its converse, where the converse follows from the theorem by substitution.

S12. (a) $(p \leftrightarrow q) \to [(p \to r) \to (q \to r)]$;
(b) $(p \leftrightarrow q) \to [(\sim p \to r) \to (\sim q \to r)]$.

Hint: For (a), start with Axioms IIIB and IC, and make a substitution so as to apply syllogism (and therefore IC again). For (b), use Axioms IIIA and IIC and syllogism (instead of using IIIB). Then continue as in (a).

We come to the last two groups. The respective theorems will be analogous, and we give hints only for proving the theorems which involve the symbol "$\vee$".

S13. Prove:

(a) $(p \vee p) \to p$;
(b) $(p \vee p) \leftrightarrow p$;
(c) $(p \vee q) \to (q \vee p)$;
(d) $(p \vee q) \leftrightarrow (q \vee p)$.

Hint: For (a), use Axiom IVc and S10. For (b), derive the converse of (a) by substituting in an axiom and use Axiom IIIc. For (c) use again IVc, but with "$q \vee p$" substituted for "r". Then use IVa–b with substitutions. For (d), argue as in S11(b).

S14. Prove:

(a) $p \to [(p \vee q) \vee r]$;
(b) $q \to [(p \vee q) \vee r]$;
(c) $r \to [(p \vee q) \vee r]$;
(d) $(q \vee r) \to [(p \vee q) \vee r]$;
(e) $[p \vee (q \vee r)] \to [(p \vee q) \vee r]$.

Hint: For (a), use Axiom IVa, then IVa with substitution so as to obtain: $(p \vee q) \to [(p \vee q) \vee r]$. Then use syllogism. For (b) start with Axiom IVb and find a similar argument. For (c) you do not need syllogism. Then (d) follows from (b), (c), and Axiom IVc. Finally, for (e) use (a) and (d).

Note: The converse of S14(e) can be proved by arguments which are entirely analogous. The equivalence can therefore be established as well.

S15. Prove:

(a) $p \to (p \wedge p)$;
(b) $p \leftrightarrow (p \wedge p)$;
(c) $(p \wedge q) \to (q \wedge p)$;
(d) $(p \wedge q) \leftrightarrow (q \wedge p)$.

S16. Prove: $[p \wedge (q \wedge r)] \to [(p \wedge q) \wedge r]$, by breaking the proof into steps as in S14.

The last three theorems of this series depend on axioms of only the first group.

S17. Prove:

(a) $(q \to r) \to [(p \to q) \to (p \to r)]$;
(b) $(p \to q) \to [p \to (r \to q)]$;
(c) $(p \to q) \to \{(q \to r) \to [(r \to s) \to (p \to s)]\}$.

Hint: For (a), substitute in Ia so as to apply syllogism between the resulting formula and Ib. For (b), substitute in (a) and in Ia and detach. For (c), first deduce from Ic by substitution:

(1) $$(p \to r) \to [(r \to s) \to (p \to s)].$$

Abbreviate the conclusion by "A", so that this becomes: "$(p \to r) \to A$". Deduce by substitution in (a) and detachment:

$$(2) \qquad [(q \to r) \to (p \to r)] \to [(q \to r) \to A].$$

Invoke IC again: its conclusion agrees with the hypothesis of (2). So make one more application of IC and obtain: $(p \to q) \to [(q \to r) \to A]$.

Note: The reader should observe the similarity and the difference between S17(a) and Axiom IC.

For the reader who desires further practice with proving theorems of sentential calculus, we suggest the following: (i) Extend S17(c) to five (or to more) variables. (ii) Formulate and prove further examples which illustrate the replacement of equivalent expressions. For instance, certain formulas similar to those in S12 can be easily established with the help of S17(a). (iii) Prove another version of the law of contraposition. (For proving e.g. "$(\sim p \to q) \to (\sim q \to p)$", one can use a technique which involves Axiom IIA and S17(a), and which is like the technique indicated in connection with S9.) (iv) Prove some of the laws which involve both symbols "$\vee$" and "$\wedge$". (Cf. exercise 14 of Chapter II.) E.g. one can prove "$\sim (p \vee q) \to (\sim p \wedge \sim q)$" easily by invoking in turn Axioms IVA–B, IIC, and VC. For proving:

$$[(p \wedge q) \vee (p \wedge r)] \to [p \wedge (q \vee r)],$$

first show that each member of the first disjunction implies each member of the last conjunction.

S18. Notice which of the foregoing axioms and theorems are known already from Chapter II, and recall their names. Include in your listing also the laws which resemble those of Chapter II, but differ from the latter in form or, to some extent, in content.

S19. The above axioms for sentential calculus are all true in the sense of truth tables (as the reader can easily verify). Conclude from this observation that all theorems which are derived from these axioms by employing the rules of substitution and detachment will likewise be true when tested by the method of truth tables.

Note: It is also possible to show that, conversely, every sentence which is true in the sense of truth tables can be derived from these axioms. The two methods of constructing sentential calculus are therefore equivalent. This task, however, is a great deal more difficult.

This page intentionally left blank

Second Part

Applications of Logic and Methodology in Constructing Mathematical Theories

This page intentionally left blank

VII

Construction of a Mathematical Theory: Laws of Order for Numbers

43 The primitive terms of the theory to be constructed; the axioms which concern the fundamental relations among numbers

With a certain amount of knowledge of the fields of logic and methodology at our disposal, we shall now undertake to lay the foundations of a particular mathematical theory which, incidentally, is very simple. This will give us an excellent opportunity to assimilate better our previously acquired knowledge, and even to expand it somewhat.

The theory with which we shall be concerned constitutes a fragment of the arithmetic of real numbers. It contains fundamental theorems which govern the basic relations *less than* and *greater than* among numbers, as well as theorems pertaining to two basic operations on numbers: to those of addition and subtraction. It presupposes nothing but logic; that is, logic is the only preceding theory.

* * *

The primitive terms which we shall adopt in this theory are the following:

real number,

is less than,

is greater than,

sum.

Instead of "*real number*" we shall simply say "*number*", as before. Actually, instead of the term "*number*", it is somewhat more convenient to consider the expression "*the set of all numbers*" as a primitive term, for which we shall use the symbol "**R**"; thus, in order to express that x is a number, we write:

$$x \in \mathbf{R}.$$

As an alternative, we may stipulate that the universe of discourse of our theory consists of real numbers only, and that variables such as "x", "y",... range over the set of numbers (or, that they stand for names of numbers); in this case, the term "*real number*" would be altogether dispensable in the formulation of statements of our theory, and when the symbol "**R**" is needed, it can be replaced by "**V**" (cf. Section 23).

The expressions "*is less than*" and "*is greater than*" are to be treated as if they consisted of a single word each; they will be replaced by the usual symbols "<" and ">", respectively. Instead of "*is not less than*" and "*is not greater than*" we shall write "$\not<$" and "$\not>$" (cf. Section 28). Further, for "*the sum of the numbers (of the summands) x and y*", or, for "*the result of adding x and y*", we shall use the customary notation:

$$x + y.$$

To sum up, the symbol "**R**" designates a certain set, the symbols "<" and ">" denote certain binary relations, and finally the symbol "+", a certain binary operation.

* * *

Among the axioms of the theory in question, two groups may be distinguished. The axioms of the first group express fundamental properties of the relations *less than* and *greater than*, while those of the second concern primarily the operation of addition. For the time being we shall be dealing with the first group only; it consists, altogether, of five sentences:

AXIOM 1. *For any numbers x and y (*that is, *for arbitrary elements of the set* **R**) *we have:* $x = y$ *or* $x < y$ *or* $x > y$.

AXIOM 2. *If* $x < y$, *then* $y \not< x$.

AXIOM 3. *If* $x > y$, *then* $y \not> x$.

AXIOM 4. *If* $x < y$ *and* $y < z$, *then* $x < z$.

AXIOM 5. *If* $x > y$ *and* $y > z$, *then* $x > z$.

The axioms listed here, just like any arithmetical theorem having universal character and stating that arbitrary numbers x, y, ... have such and such a property, should really begin with the words "*for any numbers x, y,...*" or "*for any elements x, y, ... of the set* **R**" or, simply, "*for any x, y, ...*" (if we agree that the variables "x", "y", ... here range over the set of numbers). But since we want to conform to the usage which was discussed in Section 3, we shall usually omit such a phrase and only add it in our minds; this holds not only for the axioms, but also for the theorems and the definitions which will be encountered in the course of our considerations. Axiom 2, for instance, can be read in full as follows:

For any x and y (or, *for any elements x and y of the set* $\mathbf{R}$), *if $x < y$, then $y \nless x$.*

We shall refer to Axiom 1 as the WEAK LAW OF TRICHOTOMY (there is also a strong form of this law with which we shall become acquainted later). Axioms 2–5 state that the relations *less than* and *greater than* are asymmetric and transitive (cf. Section 29); accordingly, they are called the LAWS OF ASYMMETRY and the LAWS OF TRANSITIVITY for the relations *less than* and *greater than*. More generally, the axioms of the first group and the theorems following from them are called the LAWS OF ORDER FOR NUMBERS.

* * *

The arithmetical relations $<$ and $>$, together with the logical relation of identity $=$, will be referred to in this book as the FUNDAMENTAL RELATIONS AMONG NUMBERS.

44 The laws of irreflexivity for the fundamental relations; indirect proofs

Our next task consists in deriving a number of theorems from the axioms which we have adopted. We do not aim at a systematic presentation, and our goal in this and in the following chapter is to include only those theorems, which will serve us in illustrating certain concepts and laws from the fields of logic and methodology.

THEOREM 1. *No number is smaller than itself: $x \nless x$ (for every x).*

Proof. Suppose our theorem were false. Then there would be a number x satisfying the formula:

(1) $$x < x.$$

Now, Axiom 2 refers to arbitrary numbers x and y (which need not be distinct), so that it remains valid if in place of "y" we write the variable "x", and if this latter variable then denotes the number x introduced above; we obtain in this way:

(2) $$\text{if } x < x, \text{ then } x \nless x.$$

But from (1) and (2) it follows immediately that

$$x \nless x;$$

this consequence, however, forms an obvious contradiction to formula (1). We must, therefore, reject the original assumption and accept the theorem as proved.†

We shall now show how to transform this argument into a complete proof, which we shall express for clarity in logical symbolism (cf. Sections 13 and 15). For this purpose we resort to the so-called LAW OF REDUCTIO AD ABSURDUM of sentential calculus[1]:

(I) $$(p \to \sim p) \to \sim p.$$

Moreover, we use Axiom 2 in the following symbolic form:

(II) $$(x < y) \to \sim (y < x).$$

Our proof is based exclusively on sentences (I) and (II). First we apply the rule of substitution to (I), replacing "p" in it throughout by "$(x < x)$":

(III) $$\{(x < x) \to \sim (x < x)\} \to \sim (x < x).$$

We next apply the rule of substitution to (II), replacing "y" by "x":

(IV) $$(x < x) \to \sim (x < x).$$

Finally we observe that sentence (IV) is the hypothesis of the conditional sentence (III), so that the rule of detachment may be applied. We are thus led to the formula:

(V) $$\sim (x < x),$$

which is the symbolic form of the theorem to be proved.

* * *

†It is instructive to compare the first (informal) proof with the complete proof that follows. On the one hand, the beginning step of the first proof, "Suppose our theorem were false...", is not an admissible consideration in a complete proof. On the other hand, the reader may have detected a certain lack of precision in the first proof, when a number x is introduced so as to satisfy a particular condition, and then the discussion continues with mention of arbitrary numbers x and y and the associated variables. This defect could be easily avoided by specifying instead that, say, the number u satisfies the condition in question, and by adjusting the rest of the proof accordingly. Such a device, however, could be confusing to the inexperienced. In the second proof this complication is bypassed by referring only to the variables "x" and "y", without mentioning numbers. (Note also that the universal quantifiers are disregarded in the second proof; the analogous practice was followed in the *Supplementary exercises* of Chapter VI.)

[1]This law, together with a related one of the same name:

$$(\sim p \to p) \to p,$$

has been used in many intricate and historically important arguments in logic and mathematics. The Italian logician and mathematician G. VAILATI (1863–1909) devoted a special monograph to its history.

The proof of Theorem 1 provides an example of what is called an INDIRECT PROOF, also known as a PROOF BY REDUCTIO AD ABSURDUM. Proofs of this kind may be characterized quite generally as follows: in order to prove a theorem, we assume the theorem to be false, and then derive certain consequences which compel us to reject the original assumption. Indirect proofs are very common in mathematics. They do not all follow the pattern of the proof of Theorem 1; on the contrary, the latter represents a comparatively rare form of such a proof, and we shall encounter more typical examples later on.

* * *

The axiom system adopted by us is perfectly symmetric with respect to the two symbols "$<$" and "$>$". To every theorem concerning the relation *less than*, we therefore automatically obtain the corresponding theorem concerning the relation *greater than*, the proofs being entirely analogous, so that the proof of the second theorem may be omitted altogether. In particular, corresponding to Theorem 1 we have:

THEOREM 2. *No number is greater than itself:* $x \not> x$.

The relation of identity $=$ is reflexive, as we know from logic, but Theorems 1 and 2 show that the other two fundamental relations among numbers, $<$ and $>$, are irreflexive; these theorems are therefore called the LAWS OF IRREFLEXIVITY (for the relations *less than* and *greater than*).

45 Further theorems on the fundamental relations

We shall next prove the following theorem:

THEOREM 3. $x > y$ *if, and only if,* $y < x$.

Proof. One has to show that the formulas:

$$x > y \quad \text{and} \quad y < x$$

are equivalent, that is to say, that the first implies the second, and vice versa (cf. Section 10).

Suppose, first, that

(1) $$y < x.$$

By Axiom 1 we must have at least one of the three cases:

(2) $$x = y, \quad x < y, \quad \text{or} \quad x > y.$$

If we had $x = y$, then by virtue of LEIBNIZ's law of the theory of identity (or, more directly, by the rule of replacement of equals; cf. Section 17), we could replace the variable "x" by "y" in formula (1); however, the resulting formula:

$$y < y$$

constitutes an obvious contradiction to Theorem 1. Hence we have:

(3) $$x \neq y.$$

But we also have:

(4) $$x \nless y$$

since, by Axiom 2, the formulas:

$$x < y \quad \text{and} \quad y < x$$

cannot hold simultaneously. On account of (2), (3), and (4), we conclude that the third case must apply:

(5) $$x > y.$$

We have now shown that formula (5) is implied by (1); the converse implication, in the opposite direction, can be established by an analogous procedure. The two formulas are, therefore, indeed equivalent, *q.e.d.*[2]

We may recall the terminology of the algebra of relations (cf. Section 28) and say that, according to Theorem 3, each of the relations $<$ and $>$ is the converse of the other.

THEOREM 4. *If* $x \neq y$, *then* $x < y$ *or* $y < x$.

Proof. Since

$$x \neq y,$$

we have, by Axiom 1:

$$x < y \text{ or } x > y;$$

the second member of this disjunction, by Theorem 3, is equivalent to:

$$y < x.$$

Hence we have[†]:

$$x < y \text{ or } y < x,$$

q.e.d.

In the analogous way we can prove:

---

[2]The letters "*q.e.d.*" are the customary abbreviation of the expression "*quod erat demonstrandum*", meaning "*which was to be proved*" and marking the end of a proof.

[†]This informal proof illustrates the important RULE OF REPLACEMENT OF EQUIVALENT SENTENTIAL FUNCTIONS (including, of course, sentences). This rule, which was considered (but not named) in exercise 16 of Chapter II, provides in particular the basis for handling defined terms; in this connection see also Sections 11 and 46.

THEOREM 5. *If* $x \neq y$, *then* $x > y$ *or* $y > x$.

Theorems 4 and 5 state that the relations $<$ and $>$ are connected; accordingly, these theorems are known as the LAWS OF CONNECTIVITY (for the relations *less than* and *greater than*). Axioms 2–5, together with Theorems 4 and 5, show that the set of numbers **R** is ordered by either of the two relations $<$ and $>$.

THEOREM 6. *Any numbers* x *and* y *satisfy one, and only one, of the three formulas:* $x = y$, $x < y$, *and* $x > y$.

Proof. It follows from Axiom 1 that at least one of the given formulas must be satisfied. In order to prove that the formulas:

$$x = y \quad \text{and} \quad x < y$$

exclude each other, we proceed as in the proof of Theorem 3: we replace in the second of these "x" by "y" and arrive at a contradiction to Theorem 1. Similarly it can be shown that the formulas:

$$x = y \quad \text{and} \quad x > y$$

exclude each other. And finally, the two formulas:

$$x < y \quad \text{and} \quad x > y$$

cannot hold simultaneously, because we would then have (by Theorem 3):

$$x < y \quad \text{and} \quad y < x,$$

in contradiction to Axiom 2. Hence, any numbers x and y satisfy one and no more of the three formulas in question, *q.e.d.*

We shall call Theorem 6 the STRONG LAW OF TRICHOTOMY, or simply, the LAW OF TRICHOTOMY; according to this law, one and only one of the three fundamental relations holds between any two given numbers. Using the phrase "*either... or...*" (with extension to three members) in the exclusive meaning proposed in Section 7, we can restate Theorem 6 in a more succinct manner:

For any numbers x *and* y *we have either* $x = y$ *or* $x < y$ *or* $x > y$.

46 Other relations among numbers

Apart from the fundamental relations, three other relations play an important role in arithmetic. One of these is the logical relation of diversity $\neq$, which we already know; the other two are the arithmetical relations $\leq$ and $\geq$, which will now be discussed.

The meaning of the symbol "$\leq$" is explained in the following definition:

DEFINITION 1. *We say that* $x \leq y$ *if, and only if,* $x = y$ *or* $x < y$.

The formula:

$$x \leq y$$

is to be read: "*x is less than or equal to y*" or "*x is at most equal to y*".

* * *

Although the content of the definition as stated appears to be clear, experience shows that in practical applications one sometimes comes upon certain misunderstandings. Some people, who believe they understand the meaning of the symbol "$\leq$" perfectly well, protest nevertheless against its application to definite numbers. Not only do they reject a sentence like:

$$1 \leq 0$$

as obviously false—and this rightly so—, but they also consider as meaningless or even false such formulas as:

$$0 \leq 0 \quad \text{or} \quad 0 \leq 1;$$

for they maintain that there is no sense in saying that $0 \leq 0$ or that $0 \leq 1$, since it is known that $0 = 0$ and $0 < 1$. In other words, it is not possible to exhibit a single pair of numbers which, in their opinion, satisfies the formula:

$$x \leq y.$$

This view is palpably mistaken. Just because "$0 < 1$" holds, it follows that the sentence:

$$0 = 1 \quad \textit{or} \quad 0 < 1$$

is true, for the disjunction of two sentences is certainly true provided one of them is true (cf. Section 7); but according to Definition 1, this disjunction is equivalent to the sentence:

$$0 \leq 1.$$

For an analogous reason the formula:

$$0 \leq 0$$

is also true.

The source of these misunderstandings lies, presumably, in certain habits of everyday life (*which we brought to the attention of the reader already at the end of Section 7*). In ordinary language it is customary to assert the disjunction of two sentences only if we know that one of the sentences is true, without knowing which. It would not be natural for us to say that $0 = 1$ *or* $0 < 1$, though this is undoubtedly true, since we can say something that is simpler and at the same time logically stronger, namely that $0 < 1$.

In mathematical considerations, however, it is not always advantageous to state everything that we know in its strongest possible form. For example, we sometimes assert about a quadrangle that it is a parallelogram, although we know more, e.g. that it is a square, and we would do this in order to apply a general theorem concerning arbitrary parallelograms. For similar reasons, one might know a certain number x to be less than 1 (for instance, this applies to the number 0), and yet one may assert only that $x \leq 1$, in other words, that either $x = 1$ or $x < 1$.

* * *

We shall now state two theorems which concern the relation $\leq$.

THEOREM 7. $x \leq y$ *if, and only if,* $x \not> y$.

Proof. This theorem is an immediate consequence of the law of trichotomy (Theorem 6). In fact, if

$$(1) \qquad x \leq y$$

and hence, by Definition 1,

$$(2) \qquad x = y \text{ or } x < y,$$

it is impossible for the formula:

$$x > y$$

to hold. Conversely, if

$$(3) \qquad x \not> y,$$

we must have (2) and hence, again by Definition 1, formula (1) has to hold. Formulas (1) and (3) are thus equivalent, *q.e.d.*

In the terminology of Section 28, Theorem 7 states that the relation $\leq$ is the negation of the relation $>$.

On account of its structure, Theorem 7 might be looked upon as an alternative definition of the symbol "$\leq$"; it would be different from the one adopted here but equivalent to it. The statement of this theorem may also help to dispel any last doubts about the usage of the symbol "$\leq$"; for nobody will hesitate to accept formulas like the following as true:

$$0 \leq 0 \text{ and } 0 \leq 1,$$

after having realized that these are equivalent† to the respective formulas:

$$0 \not> 0 \text{ and } 0 \not> 1.$$

†The reader should be aware that here the meaning of the word "*equivalent*" differs from the technical meaning in sentential calculus (cf. Section 10). It is not a question of two formulas or, say, definitions having the same truth value, but rather, of their

If we wished, we could have avoided the symbol "$\leq$" completely, by always employing "$\not>$" instead.

THEOREM 8. $x < y$ *if, and only if,* $x \leq y$ *and* $x \neq y$.

Proof. If

(1) $$x < y,$$

then, by Definition 1,

(2) $$x \leq y,$$

while, by the law of trichotomy, the formula:

$$x = y$$

cannot hold. Conversely, if (2) holds, then by Definition 1 we obtain:

(3) $$x < y \;\; or \;\; x = y;$$

but if, at the same time, we have:

$$x \neq y,$$

we have to accept the first part of the disjunction in (3), that is, formula (1). The implication therefore holds in both directions, *q.e.d.*

We shall pass over a number of other properties of the relation $\leq$; there are, in particular, theorems to the effect that this relation is reflexive and transitive. The proofs of these theorems do not present any difficulties.

* * *

The definition of the symbol "$\geq$" is entirely analogous to Definition 1; and from theorems involving the relation $\leq$ we automatically obtain the corresponding theorems for the relation $\geq$, simply by replacing the symbols "$\leq$", "$<$", and "$>$" throughout by the symbols "$\geq$", "$>$", and "$<$".

* * *

Formulas of the form:

$$x = y,$$

---

having (nearly) the same sense and scope. Something like the following is meant here: Two formulas (or methods, etc.) have about the same content, and for various purposes either one can be replaced by the other. We do not attempt to explain the crucial expressions "*content*" and "*various purposes*"; it would be quite impossible to do so in a general way. The resulting usage of "*equivalent*" should therefore be regarded as having primarily a heuristic value. (In various particular cases, however, such usage can easily be made precise.)—See further examples in Section 22 and at the end of *Supplementary exercises* of Chapter VI.

in which the places of "x" and "y" may be taken by constants, variables, or compound expressions denoting numbers, are usually called EQUATIONS. Similar formulas of the form:

$$x < y \text{ or } x > y$$

are called INEQUALITIES (IN THE NARROWER SENSE); among the INEQUALITIES IN THE WIDER SENSE we have, in addition, formulas of the form:

$$x \neq y, \quad x \leq y, \quad \text{or} \quad x \geq y.$$

The expressions which appear on the left and on the right side of the symbols "=", "<", and so on, in such a formula are referred to as the LEFT and the RIGHT SIDE OF THE EQUATION or OF THE INEQUALITY.

Exercises

1. Consider two relations among people: that of being of a smaller stature, and that of being of a larger stature. What condition has to be satisfied by an arbitrary set of people, so that this set together with those two relations should form a model of the first group of axioms (cf. Section 37)?

2. Let the formula:

$$x \olessthan y$$

mean that the numbers x and y satisfy one of the following conditions: (i) the number x has a smaller absolute value than the number y, or (ii) if the absolute values of x and y are the same, then x is negative and y is positive.† Moreover, let the formula:

$$x \ogreaterthan y$$

have the same meaning as:

$$y \olessthan x.$$

Show, on the basis of familiar laws of arithmetic, that the set of all numbers and the relations $\olessthan$ and $\ogreaterthan$ just defined constitute a model of the first group of axioms.

Give other examples of models (or of interpretations) of these axioms within arithmetic and within geometry.

---

†Recall: The absolute value of 0 is 0, and if $x \neq 0$, then the absolute value of x is x or $-x$, whichever is positive. For instance, the absolute value of both -2 and 2 is 2.

3. From Theorem 1 derive the following theorem:

if $x < y$, *then* $x \neq y$.

Conversely, derive Theorem 1 from the theorem just stated, without making use of any other arithmetical statements. Are your two proofs indirect, and do they fall under the pattern of the proof of Theorem 1 of Section 44?

4. Generalize the proof of Theorem 1 of Section 44, and in this way establish the following general law of the theory of relations (cf. remarks made in Section 37):

every relation R which is asymmetric in a given class K is also irreflexive in that class.

5. Show that, if Theorem 1 is adopted as a new axiom, then the old Axiom 2 can be derived as a consequence of this axiom and Axiom 4.

As a generalization of this argument, prove the following general law of the theory of relations:

every relation R which is irreflexive and transitive in a given class K is also asymmetric in that class.

***6.** At the end of Section 44 we tried to explain that the proof of Theorem 2 is analogous in view of overall symmetry, and so may be omitted. An alternate justification for omitting the proof depends upon applying certain general considerations of Chapter VI. Explain this alternate justification in detail, and in particular, specify the considerations upon which it depends.

7. Translate the following theorems into ordinary language, and derive them from the first group of axioms:

(a) $\forall_{x,y}\{x = y \leftrightarrow [\sim(x < y) \wedge \sim(y < x)]\}$;
(b) $\forall_{x,y}[x < y \rightarrow \forall_z(x < z \vee z < y)]$.

8. Derive the following theorems from Axiom 4 and Definition 1, and moreover, express them in logical symbolism:

(a) *if* $x < y$ *and* $y \leq z$, *then* $x < z$;
(b) *if* $x \leq y$ *and* $y < z$, *then* $x < z$;
(c) *if* $x \leq y$ *and* $y < z$ *and* $z \leq t$, *then* $x < t$.

9. Show that the relations $\leq$ and $\geq$ are reflexive, transitive, and connected. Are these relations symmetric or asymmetric?

10. Show that, between any two numbers, exactly three of the following six relations hold: $=$, $<$, $>$, $\neq$, $\leq$, and $\geq$.

11. Both the converse and the negation of each of the relations listed in the preceding exercise are again among these six relations. Show in detail that this is the case.

***12.** Between which of the relations given in exercise 10 does the relation of inclusion hold? What will be the sum, the product, and the relative product of any pair among these relations?

Hint: Recall the terms; they are explained in Section 28. Do not forget to consider pairs consisting of two equal relations, and remember that the relative product may depend on the order of the factors (cf. exercise 4 of Chapter V). Altogether 36 pairs of relations should be examined.

This page intentionally left blank

VIII

Construction of a Mathematical Theory: Laws of Addition and Subtraction

47 The axioms concerning addition; general properties of operations, the concept of a group and the concept of an Abelian group

We now turn to the second group of axioms, which consists of the following six sentences†:

AXIOM 6. *For any numbers x and y there exists a number z such that* $z = x + y$; in other words: *if* $x \in \mathbf{R}$ *and* $y \in \mathbf{R}$, *then there exists z such that* $z \in \mathbf{R}$ *and* $z = x + y$.

AXIOM 7. $x + y = y + x$.

AXIOM 8. $x + (y + z) = (x + y) + z$.

AXIOM 9. *For any numbers x and y there exists a number z such that* $x = y + z$.

AXIOM 10. *If* $y < z$, *then* $x + y < x + z$.

AXIOM 11. *If* $y > z$, *then* $x + y > x + z$.

For the moment let us concentrate on the first four sentences of this second group, that is, on Axioms 6-9. They ascribe to the operation of addition a number of simple properties, which are also frequently encountered in connection with other operations in various branches of logic and mathematics.

---

†Axiom 6 can be formulated more concisely as follows: *If* $x \in \mathbf{R}$ *and* $y \in \mathbf{R}$, *then* $x + y \in \mathbf{R}$. This form of the axiom is convenient in applications.

Special terms have been introduced to designate these properties. To start with, we say that the operation O is PERFORMABLE IN THE CLASS K, or, that the class K is CLOSED UNDER THE OPERATION O, if performing the operation O on any two elements of the class K results again in an element of this class; in other words, if, for any two elements x and y of the class K, there exists an element z of K such that

$$z = xOy.$$

The operation O is called COMMUTATIVE IN THE CLASS K, if the result of this operation is independent of the order of the elements of the class K on which it is carried out, or, more concisely, if for any two elements x and y of this class we have:

$$xOy = yOx.$$

The operation O is ASSOCIATIVE IN THE CLASS K, if the result is independent of the way in which the elements are grouped together, or, to be more precise, if for any three elements x, y, and z of this class, the condition:

$$xO(yOz) = (xOy)Oz$$

is satisfied. The operation O is said to be RIGHT-INVERTIBLE, or respectively, LEFT-INVERTIBLE IN THE CLASS K, if for any two elements x and y (of the class K), there always exists an element z of K such that

$$x = yOz \quad \text{or} \quad x = zOy$$

holds. An operation O which is both right- and left-invertible is simply called INVERTIBLE IN THE CLASS K. It follows at once that a commutative operation which is right- or left-invertible must be invertible. We shall now say that a class K is a GROUP WITH RESPECT TO THE OPERATION O, if this operation is performable, associative, and invertible in K; if, moreover, the operation O is commutative, then the class K is called an ABELIAN GROUP WITH RESPECT TO THE OPERATION O. Such groups, including Abelian groups as a special case, form the subject of an independent mathematical discipline which is known as the THEORY OF GROUPS; this discipline was mentioned already in Chapter V.[1]

We remark that in case the class K is the universal class (or the universe of discourse of the theory under consideration—cf. Section 23), we

[1]The concept of a group was introduced into mathematics by the French mathematician E. GALOIS (1811–1832). The term "*Abelian group*" was chosen in honor of the Norwegian mathematician N. H. ABEL (1802–1829), whose research had a great influence upon the development of higher algebra. The far-reaching importance of the concept of a group for mathematics was recognized particularly clearly after the appearance of the works of another Norwegian mathematician, S. LIE (1842–1899), and of the German mathematician F. KLEIN (1849–1925).

usually omit the reference to this class when employing such terms as "*performable*", "*commutative*", and so on.

In accordance with the terminology which has just been introduced, Axioms 6–9 are referred to as the LAW OF PERFORMABILITY, the COMMUTATIVE LAW, the ASSOCIATIVE LAW, and the LAW OF RIGHT INVERTIBILITY for the operation of addition, respectively; together they state that the set of all numbers constitutes an Abelian group with respect to addition.

48 Commutative and associative laws for larger numbers of summands

Axiom 7, the commutative law, and Axiom 8, the associative law, in the form in which they are stated here, refer to two and three summands, respectively.† But there are infinitely many other commutative and associative laws, which involve larger numbers of summands. The formula:

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

for instance, provides an example of a commutative law for three summands, and the formula:

$$x + [y + (z + u)] = [(x + y) + z] + u$$

is one of the associative laws for four summands. Furthermore, there are theorems of a mixed character which assert, generally speaking, that arbitrary changes in either the order or the grouping of the summands have no effect upon the result of the addition. By way of an example, we may state the following theorem.

THEOREM 9. $x + (y + z) = (x + z) + y$.

Proof. By suitable substitutions we obtain from Axioms 7 and 8:

(1) $z + y = y + z$,

(2) $x + (z + y) = (x + z) + y$.

In view of (1), and in accordance with LEIBNIZ's law, we may replace "$z+y$" in (2) by "$y + z$"; the result is the desired formula:

$$x + (y + z) = (x + z) + y.$$

In a similar manner we can derive all commutative and associative laws involving arbitrary numbers of summands from Axioms 7 and 8, perhaps

†Note that Axioms 7 and 8 also have two and three distinct variables, respectively. The concept of a summand, however, is basically different from that of a variable.

together with Axiom 6. Such theorems are often used in practice for the transformation of algebraic expressions. By a transformation of an expression denoting a number we mean here, as usual, an alteration which leads to an expression denoting the same number, so that the new expression may be joined with the original one by the identity sign; the expressions which are most frequently subjected to transformations of this kind are those which contain variables, and which therefore are designatory functions. We can say in other words: on the basis of the commutative and the associative law we can transform any expression having a form such as:

$$x + 2, \quad x + (y + z), \quad x + [(y + x) + (z + u)], \ldots,$$

that is, expression consisting of numerical constants and variables which are separated by addition signs and parentheses; the transformation depends on interchanging at will both the numerical symbols and the parentheses (provided only the resulting expression has not become meaningless on account of transposition of parentheses).

49 The laws of monotonicity for addition and their converses

Axioms 10 and 11, to which we now turn, are the so-called LAWS OF MONOTONICITY for addition with respect to the relations *less than* and *greater than*. We say, in a general way, that the binary operation O is MONOTONIC IN THE class K WITH RESPECT TO THE BINARY RELATION R, if for any elements x, y, z of the class K, the formula:

$$yRz$$

implies:

$$(xOy)R(xOz);$$

the latter formula means that the result of performing the operation O on the objects x and y has the relation R to the result of performing the operation O on x and z. (In the case of a non-commutative operation, one should, strictly speaking, distinguish between right and left monotonicity, the property just defined being that of right monotonicity.)

The operation of addition is monotonic not only with respect to the relations *less than* and *greater than*—as a consequence of Axioms 10 and 11—but also with respect to the other relations among numbers which were discussed in Section 46. The proof for the relation of identity is especially easy:

THEOREM 10. *If* $y = z$, *then* $x + y = x + z$.

Proof. The sum $x + y$, whose existence is guaranteed by Axiom 6, is equal to itself (by Law II of Section 17):

$$x + y = x + y.$$

In view of the hypothesis of the theorem, the variable "y" on the right side of this equation may be replaced by the variable "z", and we obtain the desired formula:

$$x + y = x + z.$$

The converse of Theorem 10 is also true:

THEOREM 11. *If* $x + y = x + z$, *then* $y = z$.

We shall sketch here two proofs of this theorem. The first, based upon the law of trichotomy and Axioms 6, 10, and 11, is comparatively simple. For our later aims, however, we require another proof which is considerably more involved, but does not make use of anything except Axioms 7–9.[†]

First proof. Suppose the theorem in question were false. Then there would be numbers x, y, and z such that

(1) $$x + y = x + z$$

and yet

(2) $$y \neq z.$$

Since $x + y$ and $x + z$ are numbers (according to Axiom 6), it follows by the law of trichotomy that they satisfy only one of the formulas:

$$x + y = x + z, \quad x + y < x + z, \quad \text{and} \quad x + y > x + z.$$

Since the first holds, by (1), the others are automatically eliminated. We therefore have:

(3) $$x + y \not< x + z \quad \textit{and} \quad x + y \not> x + z.$$

[†]The reader might be puzzled by the claim that Axiom 6 is not used. After all, if the combination "$x + y$" occurs in the statement of an axiom or of a theorem, does this not presuppose that $x + y$ is a number? This question becomes particularly relevant when alternate interpretations of primitive terms are envisaged. In this connection, let us review the options which were considered in Section 47. One option refers to the class $\mathbf{R}$ of numbers, and the other, to a general (unspecified) class K. In the latter case, if Axiom 6 is not assumed, then Axioms 7 and 8 require some additional discussion.—It may be helpful to the reader if we comment briefly on the situation where we have a class K which is a subclass of $\mathbf{R}$, but is otherwise unspecified. If now $x \in K$ and $y \in K$, then $x + y$ will still be a number, even though it might perhaps not be in K. We therefore have the possibility that Axiom 6 does not hold when adapted to K; nonetheless, in this case all sums are meaningful, and Axioms 7 and 8 remain valid.—However, theories where Axiom 6 does not hold are not of direct interest to us. Indeed, if we assume Axiom 9 in addition to Axioms 7 and 8 (which now have to be suitably interpreted), as in the second proof of Theorem 11, then Axiom 6 can be derived; see Section 55. All the sums in the second proof therefore are well defined, and in fact, belong to the assumed underlying class.

On the other hand, by applying the law of trichotomy once more, we can infer from inequality (2) that

$$y < z \quad or \quad y > z.$$

Hence, by Axioms 10 and 11,

(4) $$x + y < x + z \quad or \quad x + y > x + z,$$

which represents an obvious contradiction to (3).[††] The supposition is thus refuted, and the theorem must be considered proved.

**Second proof.*[‡] Apply Axiom 9, with "x" and "z" replaced by "y" and "u", respectively. It follows that there exists a number u which fulfills (or, satisfies) the formula:

$$y = y + u.$$

But by Axiom 7,

$$y + u = u + y,$$

and in view of the transitivity of the relation of identity (cf. Law IV of Section 17):

(1) $$y = u + y.$$

[††] The contradiction between (3) and (4) can be expressed in an explicit way by invoking one of DE MORGAN's laws. Cf. exercise 14 of Chapter II.—There is another interesting aspect of the assertion in (4), which regards its derivation from the preceding statement. Here one starts with a disjunction whose members (or, DISJUNCTS) are the hypotheses of valid implications, and one then asserts the disjunction of the respective conclusions. Such an argument may seem unconvincing to some readers. One way to confirm its correctness depends on the following law of sentential calculus:

$$(p \to q) \to \{(r \to s) \to [(p \lor r) \to (q \lor s)]\}.$$

Indeed, by starting with this law and by making appropriate substitutions and three detachments, one can exhibit the derivation under discussion in the form of a complete proof.

[‡] The second proof goes through a number of equations which are elementary (and linear), but they all look rather similar. Their enumeration therefore does not reveal the overall structure, or idea, of the argument. Such an underlying idea, however, should not be disregarded, and it is essential to understand it if one wants to be able to reconstruct the proof, or to adapt it to a different situation.—If we are faced with the equation: "$x + y = x + z$", the obvious thought is to consider the number $-x$ and to transform the equation by means of addition. The axioms tell us nothing about inverse quantities, however, and we have to build up to $-x$, so to say. First a number u is introduced, and it has the interpretation of zero. Next v is introduced. This is an auxilliary which is used to show that u retains the interpretation as zero also when combined with a quantity other than y, in particular with z. Then comes w, of which we can think as $-x$, and which enables us to realize the initial idea.—Note also that in the second proof the author chose to apply Laws IV and V of the theory of identity, while a direct application of LEIBNIZ's law (or of the rule of replacement of equals) might seem more natural.

Now apply Axiom 9 again, with "x" and "z" replaced by "z" and "v", respectively; we thereby obtain a number v satisfying the equation:

$$(2) \qquad z = y + v.$$

On account of (1), we may replace here the variable "y" by the expression "$u + y$":

$$z = (u + y) + v.$$

Next, by the associative law, Axiom 8, we have:

$$u + (y + v) = (u + y) + v,$$

and by applying Law V of Section 17 we arrive at:

$$z = u + (y + v).$$

On account of (2) we may replace here "$y + v$" by "z" (using LEIBNIZ's law), so that we finally obtain:

$$(3) \qquad z = u + z.$$

Applying Axiom 9 for the third time, this time with "x", "y", and "z" replaced by "u", "x", and "w", respectively, we obtain a number w such that

$$u = x + w,$$

and since

$$x + w = w + x,$$

we have:

$$(4) \qquad u = w + x.$$

From (4) and (1) we obtain the following:

$$y = (w + x) + y;$$

but by the associative law we have:

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

and the last two equations yield:

$$(5) \qquad y = w + (x + y).$$

In view of the hypothesis of the theorem which we are proving, we may replace "$x + y$" in (5) by "$x + z$", and this brings us to:

$$(6) \qquad y = w + (x + z).$$

Applying again the associative law, we have:

$$w + (x + z) = (w + x) + z,$$

so that (6) becomes:

$$y = (w + x) + z.$$

On account of (4), we may here replace "$w + x$" by "u". In this way we obtain:

$$(7) \qquad y = u + z.$$

But from equations (7) and (3) it follows that

$$y = z,$$

*q.e.d.**

* * *

Let us insert here a few remarks which concern the first proof of Theorem 11. Like the proof of Theorem 1, it constitutes an example of an indirect inference. The schema[‡‡] of this proof may be described as follows. In order to prove a certain sentence, say "p", we suppose the sentence to be false, that is, we assume the sentence "*not p*". From this assumption a consequence "q" is derived; that is to say, we demonstrate the implication:

if not p, then q

(in the case under consideration, the consequence "q" is the conjunction of conditions (3) and (4) which appear in the proof). On the other hand, by invoking some general laws of logic (as in the case under consideration) or some theorems previously proved within the mathematical discipline in which all these arguments are carried out, we are able to show that the consequence obtained is false, that is, that "*not q*" holds; thereby we are compelled to give up the original assumption, and thus to accept the sentence "p" as true. If this argument were set down in the form of a complete proof, an essential role would be played by a logical law, which is a variant of the law of contraposition known to us from Section 14; this law reads as follows:

From: if not p, then q, it follows that: if not q, then p.

The schema under consideration differs somewhat from that of Theorem 1. There, having assumed that the theorem is false, we inferred that the theorem is true, that is, we derived a consequence directly contradicting our assumption; here, however, we derived from an analogous assumption a consequence which we knew from other sources to be false. But this difference is not an essential one; it can easily be shown with the help of

[‡‡] A mathematical proof is often arranged according to a certain familiar outline, and the word "*schema*" here refers to an outline of this kind. (The more commonplace word "*pattern*" was used in the previous chapter in the same way.)

logical laws that the proof of Theorem 1—like any other indirect mode of inference—can be brought under the schema which we have just sketched.

* * *

Besides Theorem 10, the other laws of monotonicity, Axioms 10 and 11, admit a conversion:

THEOREM 12. *If* $x + y < x + z$, *then* $y < z$.

THEOREM 13. *If* $x + y > x + z$, *then* $y > z$.

The proof of these theorems can be obtained without difficulty along the line of the first proof of Theorem 11.

50 Closed systems of sentences

There exists a general law, the knowledge of which can considerably simplify the proofs of the last three theorems (11, 12, and 13). This law, which sometimes is called the LAW OF CLOSED SYSTEMS or HAUBER'S LAW,[2] applies to certain situations where we have succeeded in proving several conditional sentences, and it allows us to infer that the corresponding converse sentences must also be true.

Suppose we are given a number of implications, say three, to which we shall assign the following schematic form:

if p_1, *then* q_1;
if p_2, *then* q_2;
if p_3, *then* q_3.

These three sentences are said to form a CLOSED SYSTEM, if their antecedents are of such a kind as to exhaust all possible cases, that is, if it is true that

p_1 *or* p_2 *or* p_3,

and if, at the same time, their consequents exclude one another:

if q_1, *then not* q_2; *if* q_1, *then not* q_3; *if* q_2, *then not* q_3.

The law of closed systems asserts that if several conditional sentences are true and form a closed system, then the corresponding converse sentences are also true.[†]

[2] After the name of the German mathematician K. F. HAUBER (1775–1851).

[†] The reader should observe that the law of closed systems is a methodological law, not belonging to the domain of logic proper.

The simplest example of a closed system is given by a system of two sentences, consisting of an implication:

if p, then q,

and of its inverse sentence:

if not p, then not q.

In order to demonstrate the two converse sentences in this case, it is not even necessary to resort to the law of closed systems; it suffices to apply the laws of contraposition.

Theorem 10 and Axioms 10 and 11 form a closed system of three sentences. This is a consequence of the law of trichotomy; since between any two numbers we have exactly one of the relations $=$, $<$, and $>$, we see that the hypotheses of these three sentences, that is, the formulas:

$$y = z, \quad y < z, \quad y > z,$$

exhaust all possible cases, while their conclusions:

$$x + y = x + z, \quad x + y < x + z, \quad x + y > x + z,$$

exclude one another. (The law of trichotomy implies even more, which however is irrelevant for our purpose, namely: the first three formulas do not only exhaust all possible cases but also exclude one another, and the last three formulas do not only exclude one another but also exhaust all possible cases.) Just because the three statements form a closed system, the converse theorems 11–13 must hold.

Numerous examples of closed systems can be found in elementary geometry; for instance, if we should want to analyze certain properties which involve the relative position of two circles, having radii of equal length, we may have to deal with a closed system consisting of four sentences.

In conclusion it may be remarked, that if someone does not want to refer to the law of closed systems but wants to prove the converses of statements forming a system of this kind, he or she may simply apply the same mode of inference which we employed in the first proof of Theorem 11.

51 A few consequences of the laws of monotonicity

Theorems 10 and 11 may be combined into a single sentence:

$$y = z \text{ \textit{if, and only if,} } x + y = x + z.$$

Similarly it is possible to combine Axioms 10 and 11 with Theorems 12 and 13. The theorems which are so obtained may be designated as the LAWS OF

TRANSFORMATION OF EQUATIONS AND INEQUALITIES INTO EQUIVALENT ONES by means of addition. The content of these theorems is sometimes described as follows: if the same number is added to each side of an equation or inequality, without changing the equality or inequality sign, then the resulting equation or inequality is equivalent to the original one (this formulation, of course, is not entirely correct, since the sides of an equation or inequality are not numbers but expressions, to which it is not possible to add any numbers). The theorems which we have just mentioned play a basic role in the solution of equations and of inequalities.

* * *

We shall derive here one more consequence from the laws of monotonicity, a consequence which will be needed later:

THEOREM 14. *If* $x + z < y + t$, *then* $x < y$ *or* $z < t$.

Proof. Suppose the conclusion of the theorem is false; in other words, neither is x smaller than y nor is z smaller than t. From this it follows by the law of trichotomy that one of the two formulas:

$$x = y \text{ or } x > y$$

and also one of the two formulas:

$$z = t \text{ or } z > t$$

must hold. We therefore have to discuss the following four possibilities:

(1) $x = y$ *and* $z = t$,

(2) $x = y$ *and* $z > t$,

(3) $x > y$ *and* $z = t$,

(4) $x > y$ *and* $z > t$.

Let us begin by considering the first case. If the two equations in (1) are valid, we obtain from the first:

$$z + x = z + y$$

by Theorem 10; and since, according to Axiom 7,

$$x + z = z + x \text{ and } z + y = y + z,$$

we may infer, by a twofold application of the law of transitivity for the relation of identity:

(5) $x + z = y + z$.

If now we apply Theorem 10 to the second equation in (1), we obtain:

(6) $y + z = y + t$,

which, together with (5), yields:

(7) $x + z = y + t$.

By an entirely analogous inference—depending in addition on Axioms 5, 6, and 11—each of the three remaining cases, that is, (2), (3), and (4), leads to the inequality:

(8) $x + z > y + t$.

One of these two formulas, (7) or (8), must therefore hold in any case. But since $x + z$ and $y + t$ are numbers (Axiom 6), it follows by the law of trichotomy that the formula:

$$x + z < y + t$$

cannot hold.†

---

†This proof, like the first proof of Theorem 11, hinges on the laws of monotonicity and trichotomy. The steps of the present proof, when examined in detail, illustrate an interesting interplay of various laws of sentential calculus, especially of those involving the symbol "$\vee$" (and of course "$\rightarrow$"). It should be instructive to scrutinize some of these steps, and to present them as fragments of a complete proof. This will give us, in particular, an opportunity for reviewing some of those laws (cf. also the example in footnote "††" on p. 164).—Let us disregard the quantifiers, which are implied, and let us abbreviate the hypothesis and the conclusion of the theorem by "H" and "C" and the four statements in (1)–(4) by "S", "T", "U", "V", respectively. For an indirect proof we consider "$\sim C$", and an important preliminary assertion is:

$$\sim C \rightarrow [(S \vee T) \vee (U \vee V)]$$

(where the disjuncts can also be grouped in other ways). A proof of this assertion would involve especially the law of trichotomy, as well as one of DE MORGAN's laws and a distributive law (cf. exercise 14 of Chapter II); we leave it to the reader to work out the details.—Next, let us abbreviate the statements in (7) and (8) by "Y" and "Z". Arguments which are given (or indicated) in the text and depend on the laws of monotonicity show, in effect, that

$$S \rightarrow Y, \quad T \rightarrow Z, \quad U \rightarrow Z, \quad V \rightarrow Z.$$

With the help of the law of logical addition (we refer to the form as in Section 12 and to its obvious variant) and of the law of the hypothetical syllogism, cf. Axioms IVA–B and also IC of *Supplementary exercises* of Chapter VI, we now derive:

$$S \rightarrow (Y \vee Z), \ldots, V \rightarrow (Y \vee Z).$$

Then, by a threefold application of the law given as Axiom IVC, cf. *loc. cit.*, we find:

$$[(S \vee T) \vee (U \vee V)] \rightarrow (Y \vee Z).$$

From this and the first assertion, by syllogism: $\sim C \rightarrow (Y \vee Z)$. Next one can easily deduce, using trichotomy: $\sim C \rightarrow \sim H$, and now we refer the reader to the discussion in the text.

Thus, by assuming the conclusion to be false, we have arrived at a direct contradiction to the hypothesis of the theorem. The assumption is therefore refuted, and we see that the conclusion does indeed follow from the hypothesis.

* * *

The argument just carried out is regarded as an indirect proof; apart from an inessential modification, it could be brought under the schema which is sketched in Section 49 in connection with the first proof of Theorem 11. Formally considered, however, the course of the argument differs a little from those followed in the proofs of Theorems 1 and 11 . The inference proceeds in the following way. In order to prove a sentence having the form of an implication, that is, having the schematic form:

if p, then q,

we assume the conclusion of the sentence, that is "q", to be false (and not the whole sentence); from this assumption, that is, from "*not q*", one infers that the hypothesis is false, or, that "*not p*" holds. In other words, instead of demonstrating the sentence in question, a proof of the corresponding contrapositive sentence:

if not q, then not p

is given, and from this the validity of the original sentence is inferred. This last inference depends on a variant of the law of contraposition, according to which the truth of the contrapositive sentence implies that of the original sentence (cf. Section 14).

Inferences of this form are very common in all mathematical disciplines; they constitute the most usual type of indirect proof.

52 The definition of subtraction; inverse operations

Our next task is to show how the notion of subtraction can be introduced into our considerations. With this goal in mind, we first prove the following theorem:

THEOREM 15. *For any two numbers y and z there is exactly one number x such that $y = z + x$.*

Proof. If arbitrary numbers y and z are given, then Axiom 9 guarantees the existence of at least one number x satisfying the formula:

$$y = z + x.$$

We have to show that there is no more than one such number, in other words, that any two numbers u and v satisfying this formula are identical. Let us therefore suppose that

$$y = z + u \text{ and } y = z + v.$$

This implies at once (by the laws of symmetry and transitivity for the relation $=$):

$$z + u = z + v,$$

from which we obtain, by Theorem 11:

$$u = v.$$

Thus there is exactly one number x (cf. Section 20) for which

$$y = z + x,$$

q.e.d.

This unique number x, which is specified in the above theorem, is designated by the symbol:

$$y - z;$$

we read it, as usual, "*the difference of the numbers y and z*" or, better, "*the result of subtracting the number z from the number y*". The precise definition of the concept of difference is as follows:

DEFINITION 2. *We say that* $x = y - z$ *if, and only if,* $y = z + x$.

More generally, an operation I is called a RIGHT INVERSE OF THE OPERATION O IN THE CLASS K if these two operations O and I fulfill the following condition:

for any elements x, y, and z of the class K, we have:
$x = yIz$ *if, and only if,* $y = zOx$.

The analogous concept of a LEFT INVERSE OF THE OPERATION O is defined similarly. If the operation O is commutative in the class K, then its two inverses—the right and the left—coincide, and we can speak simply of the INVERSE OF THE OPERATION O (or else, of the INVERSE OPERATION OF O). In accordance with this terminology, Definition 2 expresses the fact that subtraction is the right inverse (or simply, the inverse) of addition.

53 Definitions whose definiendum contains the identity sign

*Definition 2 exemplifies a kind of definition which occurs very frequently in mathematics. Such a definition stipulates the meaning of a symbol designating either a single object or an operation on a certain number of objects (the latter, we recall, has the same content as a function with a certain number of arguments). In every definition of this kind, the definiendum has the form of an equation:

$$x = \dots;$$

on the right side of this equation we have the symbol itself which was to be defined, or else a designatory function constructed out of the symbol to be defined and certain variables "y", "z", ..., depending on whether the symbol in question designates a single object or an operation on objects. The definiens may be a sentential function of any form, which contains the same free variables as the definiendum, and which states that the object x—together possibly with the objects y, z, ...—satisfies such and such a condition.—Definition 2 establishes the meaning of a symbol which denotes an operation on two numbers. To give a different example of this type of definition, let us state the definition of the symbol "0", which designates a single number:

We say that $x = 0$ *if, and only if, for any number* y,
the formula: $y + x = y$ *holds.*

* * *

A certain danger is connected with definitions of this type; for if one does not proceed with sufficient caution in constructing such definitions, one can easily become confronted with a contradiction. A concrete example will make this clear.

Let us leave for the moment our present investigations, and let us assume that in arithmetic we already have the symbol of multiplication at our disposal, and that, with its help, we want to define the symbol of division. It is then natural to consider the following provisional definition, which has precisely the same form as Definition 2:

we say that $x = y \div z$ *if, and only if,* $y = z \cdot x$.

If now, in this definition, we replace both "y" and "z" by "0", and "x" first by "1" and then by "2", and if we observe that the following formulas are true:

$$0 = 0 \cdot 1 \quad \text{and} \quad 0 = 0 \cdot 2,$$

we obtain at once:

$$1 = 0 \div 0 \quad \text{and} \quad 2 = 0 \div 0.$$

But since two objects which are equal to the same object are equal to each other, we arrive at:

$$1 = 2,$$

which is absurd.

It is not hard to find the reason for this phenomenon. Both in Definition 2 and in the provisional definition of the quotient, the definiens has the form of a sentential function with three free variables, "x", "y", and "z".

To each such sentential function there corresponds a three-place relation which holds among the numbers x, y, and z if, and only if, these numbers satisfy that sentential function (cf. Sections 27 and 34); and one can state a definition which would introduce a symbol designating this relation. But if one gives the definiendum the form:

$$x = y - z \text{ or } x = y \div z,$$

one assumes in advance that this relation is functional (and hence an operation, or, a function, cf. Section 34), and that therefore, to any two numbers y and z, there is at most one number x standing to them in the relation in question. The condition that the relation be functional, however, is not automatically fulfilled, and its validity must first be established. This we did in the case of Definition 2; but we failed to do so in the case of the above definition of the quotient, and we would indeed have been unable to do so, simply because the relation in question ceases to be functional in a certain exceptional case: for, if

$$y = 0 \text{ and } z = 0,$$

then there exist infinitely many numbers x for which

$$y = z \cdot x.$$

Therefore, if one wants to formulate the definition of the quotient in the above form without introducing a contradiction, one has to exclude the case where both numbers y and z are 0,—for instance, by inserting an additional condition in the definiens.

These considerations lead us to the following conclusion. Every definition of the type of Definition 2 should be preceded by a theorem which would be analogous to Theorem 15, that is to say, by a theorem to the effect that there is but one number x which satisfies the definiens. (The question arises whether it is essential to have exactly one number x, or whether it suffices that there be at most one such number. A discussion of this rather difficult problem will be omitted.)*

54 Theorems on subtraction

On the basis of Definition 2 and the laws of addition we can prove without difficulty the fundamental theorems relating to subtraction, such as the law of performability, the laws of monotonicity, and the laws of transformation of equations and inequalities into equivalent ones by means of subtraction. One also has here those theorems which govern the transformation of so-called algebraic sums, that is, of expressions consisting of numerical constants and variables, where these symbols are separated by the signs

"+" and "−" as well as by parentheses (the latter often being omitted in accordance with special conventions). The following theorem may serve as an example of the last-named category:

THEOREM 16. $x + (y - z) = (x + y) - z$.

Proof. To arbitrary numbers y and z, according to Axiom 9, there corresponds a number u such that

(1) $y = z + u$;

this implies, by Definition 2,

(2) $u = y - z$.

The commutative law states that:

$$x + y = y + x.$$

On account of (1), "y" may here be replaced by "$z + u$" on the right side, so that we obtain:

(3) $x + y = (z + u) + x$.

From Theorem 9, on the other hand, it follows that:

(4) $z + (x + u) = (z + u) + x$.

But since two numbers which are equal to the same number are equal to each other, we can infer from (3) and (4):

(5) $x + y = z + (x + u)$.

Now, since $x + u$ and $x + y$ are numbers (by Axiom 6), we may substitute "$x + u$" and "$x + y$" for "x" and "y" in Definition 2. (5) shows that the definiens is then satisfied, and hence the definiendum must also hold:

$$x + u = (x + y) - z.$$

If, in view of (2), we replace "u" by "$y - z$" in the last equation, we finally arrive at:

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

q.e.d.

Having reached this point, we terminate the construction of our fragment of arithmetic.

Exercises

1. Consider the following three systems, each consisting of a certain set, two relations, and one operation:

(a) the set of all numbers, the relations $\leq$ and $\geq$, the operation of addition;
(b) the set of all numbers, the relations $<$ and $>$, the operation of multiplication;
(c) the set of all positive numbers, the relations $<$ and $>$, the operation of multiplication.

Determine (on the basis of familiar laws of arithmetic) which of these systems are models of the system of Axioms 1–11 (cf. Section 37).

2. Consider an arbitrary straight line, to which we shall refer as the number line; let the points on this line be denoted by the letters "X", "Y", "Z", ... On the number line we choose a fixed initial point O and a unit point U distinct from O. Now let X and Y be any two distinct points on our line. We consider the two half-lines, one beginning at O and going through U, the other beginning at X and going through Y. We shall say that the point X precedes the point Y, in symbols:

$$X \overset{\circ}{<} Y,$$

if, and only if, either the two half-lines are identical, or one of them—no matter which—is a part of the other. Under the same circumstances we shall also say that the point Y succeeds the point X, and we shall write:

$$Y \overset{\circ}{>} X,$$

The point Z is called the sum of the points X and Y if it fulfills the following conditions: (i) the segment OX is congruent to the segment YZ; (ii) if $O \overset{\circ}{<} X$, then $Y \overset{\circ}{<} Z$, but if $O \overset{\circ}{>} X$, then $Y \overset{\circ}{>} Z$. The sum of the points X and Y is denoted by:

$$X \overset{\circ}{+} Y.$$

Show by employing the theorems of geometry, that the set of all points of the number line (that is, the number line itself), the relations $\overset{\circ}{<}$ and $\overset{\circ}{>}$, and the operation $\overset{\circ}{+}$ together form a model of the axiom system adopted by us, and that this system therefore has an interpretation within geometry.

3. We consider four operations A, B, G, and L which—like addition—correlate a third number with any two numbers. For the result of the operation A on the numbers x and y we always take the number x, and for the result of the operation B, the number y:

$$xAy = x, \quad xBy = y.$$

By the symbols "xGy", or "xLy", we denote that of the two numbers x and y which is not less than, or not greater than, the other, respectively; to say in other words,

$$xGy = x \text{ and } xLy = y \text{ in case that } x \geq y;$$
$$xGy = y \text{ and } xLy = x \text{ in case that } x \leq y.$$

Which of the properties discussed in Section 47 pertain to each of these four operations? Is the set of all numbers a group, and in particular, an Abelian group, with respect to any of these operations?

4. Let C be the class of all point sets (or, of all geometrical configurations). Are the addition and the multiplication of sets (as defined in Section 25) performable, commutative, associative, and invertible in the class C? Is the class C therefore a group, and in particular, an Abelian group, with respect to either of these operations?

5. Show (on the basis of familiar laws of arithmetic) that the set of all numbers is not an Abelian group with respect to multiplication, but that each of the following sets is an Abelian group with respect to that operation:

(a) the set of all numbers different from 0;
(b) the set of all positive numbers;
(c) the set consisting of the two numbers 1 and -1.

6. Consider the set S consisting of the two numbers 0 and 1, and let the operation $\oplus$ on the elements of this set be defined by the following formulas:

$$0 \oplus 0 = 1 \oplus 1 = 0,$$
$$0 \oplus 1 = 1 \oplus 0 = 1.$$

Determine whether the set S is an Abelian group with respect to the operation $\oplus$.

7. Consider the set S consisting of the three numbers 0, 1, and 2. Define an operation $\oplus$ on the elements of this set, so that the set S will be an Abelian group with respect to this operation.

8. Prove that no set consisting of two or three different numbers can be an Abelian group with respect to addition. Is there a set which consists of a single number and forms an Abelian group with respect to addition?

9. Derive the following theorems from Axioms 6–8:

(a) $x + (y + z) = (z + x) + y$;
(b) $x + [y + (z + t)] = (t + y) + (x + z)$.

10. How many expressions can be obtained from each of these expressions:

$$x + (y + z), \quad x + [y + (z + t)], \quad x + \{y + [z + (t + u)]\}$$

if they are transformed solely on the basis of Axioms 6–8?

11. Formulate the general definition of left monotonicity of a binary operation O with respect to a (binary) relation R.

12. On the basis of the axioms adopted by us and the theorems derived from them, prove that addition is a monotonic operation with respect to each of the relations $\neq$, $\leq$, and $\geq$.

13. Determine (on the basis of familiar laws of arithmetic) whether multiplication is a monotonic operation with respect to each of the two relations $<$ and $>$ in the following sets:

(a) in the set of all numbers,
(b) in the set of all positive numbers,
(c) in the set of all negative numbers.

14. Which of the operations defined in exercise 3 are monotonic with respect to each of the relations $=$, $<$, $>$, $\neq$, $\leq$, and $\geq$?

15. Are the addition and the multiplication of classes monotonic with respect to the relation of inclusion? Or with respect to any of the other relations among classes which are discussed in Section 24?

16. Derive from our axioms the following theorem:

$$\textit{if}\ \ x < y\ \ \textit{and}\ \ z < t,\ \ \textit{then}\ \ x + z < y + t.$$

Replace in this sentence the symbol "$<$" in turn by "$>$", "$=$", "$\neq$", "$\leq$", and "$\geq$", and examine which of the sentences obtained in this way are true.

17. Give examples of closed systems of sentences within arithmetic and within geometry.

18. Derive the following theorems from our axioms:

(a) *if* $x + x = y + y$, *then* $x = y$;
(b) *if* $x + x < y + y$, *then* $x < y$;
(c) *if* $x + x > y + y$, *then* $x > y$.

Hint: Prove the converse sentences first (using the results of exercise 16), and show that they form a closed system.

***19.** Any theorem which is derivable from Axioms 6–9 alone can, evidently, be extended to arbitrary Abelian groups; for, a given class K which forms an Abelian group with respect to an operation O also constitutes, together with this operation, a model of Axioms 6–9 (cf. Sections 37 and 38).† This

† Axioms 6–9 of course fully characterize Abelian groups, and one can therefore sharpen the statement in the exercise by referring to the following (methodological) definition: *We say that a set K is* (or, *forms) an Abelian group with respect to an operation O if, and only if, K and O jointly constitute a model of Axioms 6–9.* In fact, several exercises of this chapter depend on this definition.

applies, in particular, to Theorem 11 (in view of the second proof of this theorem), and we have the following general group-theoretical result:

every class K which is an Abelian group with respect to the operation O satisfies the following condition:
if $x \in K,\ y \in K,\ z \in K$, and $xOy = xOz$, then $y = z$.

Give a careful proof of this theorem.

Show, on the other hand, that Theorem (a) of exercise 18 cannot be extended to arbitrary Abelian groups, by exhibiting an example of a class K and an operation O with the following properties: (i) the class K is an Abelian group with respect to the operation O, and (ii) there exist two distinct elements x and y of the class K for which $xOx = yOy$ (cf. exercise 6). Consequently, is it possible to derive Theorem (a) from Axioms 6–9 alone?

20. Transform the proof of Theorem 14 in such a manner, that it conforms to the schema which is sketched in Section 49 in connection with the first proof of Theorem 11.

21. May one say that the operation of division is the inverse of multiplication in the set of all numbers?

22. Do the operations mentioned in exercises 3 and 4 possess inverses (in the set of all numbers, or in the class of all geometrical configurations, respectively)?

23. What operations are the left and the right inverse of subtraction (in the set of all numbers)?

***24.** In Section 53 the definition of the symbol "0" was stated as an example. In order to assure that this definition does not lead to a contradiction, one should first establish the following theorem:

There exists exactly one number x such that, for any number y, we have: $y + x = y$.

Prove this theorem on the basis of Axioms 6–9 alone.

25. Formulate the sentences which assert that subtraction is performable, commutative, associative, right- and left-invertible, and right- and left-monotonic with respect to the relation *less than*. Which of these sentences are true? Prove those for which this is the case, using our axioms and Definition 2 of Section 52.

26. Derive the following theorems from our axioms and Definition 2:

(a) $x - (y + z) = (x - y) - z$,
(b) $x - (y - z) = (x - y) + z$,
(c) $x + y = x - [(x - y) - x]$.

***27.** Using the law of performability for subtraction and Theorem (c) of the preceding exercise, prove the following theorem:

> *For a set K of numbers to be an Abelian group with respect to addition, it is necessary and sufficient that the difference of any two numbers of the set K also belong to the set K (*that is, *that the formulas $x \in K$ and $y \in K$ always imply $x - y \in K$).*

Use this theorem in order to find examples of sets of numbers that are Abelian groups with respect to addition.

28. Write in the symbolism of logic all the axioms, definitions, and theorems which are given in the last two chapters.

Hint: Before formulating Theorem 15 in symbols, put it into an equivalent form, such that the numerical quantifiers have been eliminated in accordance with the explanations given in Section 20.

***29.** Write in logical symbolism the formulas which express the condition that a given class K is an Abelian group with respect to the operation O (according to the definition given in Section 47). Furthermore, consider the following three formulas:

(a) $\forall_{x,y}[(x \in K \wedge y \in K) \rightarrow xOy \in K]$.
(b) $\forall_{x,y,z}[(x \in K \wedge y \in K \wedge z \in K) \rightarrow (xOy)Oz = xO(yOz)]$.
(c) $\forall_{x,y}[(x \in K \wedge y \in K) \rightarrow \exists_z(z \in K \wedge x = yOz \wedge x = zOy)]$.

Try to show that these three formulas provide an equivalent definition of the concept of an Abelian group; in other words, try to show that they (taken together) constitute a necessary and sufficient condition that

> *the class K is an Abelian group with respect to the operation O.*

IX

Methodological Considerations on the Constructed Theory

55 Elimination of superfluous axioms from the original axiom system

In the two preceding chapters we sketched the foundations of an elementary mathematical theory, which constitutes a fragment of arithmetic. In the present chapter we shall continue with a methodological discussion, and we shall be concerned at first with the system of axioms and primitive terms upon which that theory is based.

We begin with concrete examples which illustrate the considerations of Section 39; these examples relate to such problems as arbitrariness in the selection of axioms and primitive terms, the possibility of omitting superfluous axioms, and so on.

* * *

Let us start with the question whether our system of Axioms 1–11—we shall refer to it briefly as SYSTEM $\mathcal{A}$—perhaps contains any superfluous axioms, that is, whether any of its axioms can be derived from the remaining axioms of the system. We shall see immediately that it is easy to answer this question, and moreover, in the affirmative. In fact, we have:

> *Three of the axioms of System $\mathcal{A}$, namely, one of Axioms 4–5, Axiom 6, and one of Axioms 10–11, can be derived from the remaining axioms.*

Proof. We show first that

(I) *each of Axioms 4 and 5 can be derived from the other with the help of Axioms 1–3.*

Let us observe that the proof of Theorem 3 was based exclusively—whether directly or indirectly—upon Axioms 1–3. However, once we have Theorem 3 at our disposal, we may derive Axiom 5 from Axiom 4 (or vice versa) by reasoning of the following kind:

If

$$x > y \textit{ and } y > z,$$

then, by Theorem 3,

$$y < x \textit{ and } z < y;$$

hence, by applying Axiom 4 (with "x" having been replaced by "z", and "z" by "x"), we obtain:

$$z < x,$$

which implies, again by Theorem 3:

$$x > z,$$

and this is the conclusion of Axiom 5.

In a similar way it can be shown that:

(II) *each of Axioms 10 and 11 can be derived from the other with the help of Axioms 1–3 and 6.*

Finally, we have:

(III) *Axiom 6 can be derived from Axioms 7–9.*

*The proof of the last assertion is not very simple and resembles the second proof of Theorem 11. Two arbitrary numbers x and y are given; by a fourfold application of Axiom 9, we introduce four new numbers u, w, z, and v one by one, so as to satisfy the following formulas:

(1) $y = y + u$,

(2) $u = x + w$,

(3) $y = w + z$,

(4) $z = y + v$.

From (1) we have, by the commutative law,

$$y = u + y;$$

combining this equation with (4) as in the proof of Theorem 11, and using the associative law, we obtain,

(5) $$z = u + z.$$

From this, taking (2) into account,

$$z = (x + w) + z,$$

and hence, again by the associative law,

$$z = x + (w + z),$$

which, in view of (3), yields:

$$(6) \qquad z = x + y.$$

We have thus shown that, for any two numbers x and y, there exists a number z for which (6) holds; and this is just the content of Axiom 6.†

It might be added that the reasoning which we sketched above applies not only to addition, but—in accordance with the general remarks of Sections 37 and 38—also to any other operation; an operation O which is commutative, associative, and right-invertible in a class K is also performable in that class, and the class K therefore forms an Abelian group with respect to the operation O (cf. Section 47).*

* * *

We have shown that System $\mathcal{A}$ contains at least three axioms which are superfluous and may therefore be omitted. Consequently, System $\mathcal{A}$ may be replaced by the system consisting of the following eight axioms:

AXIOM $1^{(1)}$. *For any numbers x and y we have: $x = y$ or $x < y$ or $x > y$.*

AXIOM $2^{(1)}$. *If $x < y$, then $y \not< x$.*

AXIOM $3^{(1)}$. *If $x > y$, then $y \not> x$.*

AXIOM $4^{(1)}$. *If $x < y$ and $y < z$, then $x < z$.*

AXIOM $5^{(1)}$. $x + y = y + x$.

AXIOM $6^{(1)}$. $x + (y + z) = (x + y) + z$.

AXIOM $7^{(1)}$. *For any numbers x and y there exists a number z such that $x = y + z$.*

AXIOM $8^{(1)}$. *If $y < z$, then $x + y < x + z$.*

We shall refer to this axiom system as SYSTEM $\mathcal{A}^{(1)}$, and we can now

†The careful reader will note that for the above derivation of Axiom 6, the commutative law should be interpreted as follows: *If $x + y \in K$, where K is the universe of discourse, then also $y + x \in K$, and the two elements are equal.* A similar statement would apply to the associative law. (Cf. also footnote on p. 163.)—We might add that the possibility of deriving Axiom 6 is not surprising. Indeed, a derivation was indicated already in exercises 26 and *27 of Chapter VIII.

claim the following result[††]:

Systems $\mathcal{A}$ and $\mathcal{A}^{(1)}$ are equipollent.

In comparison with the original system, the new reduced system has certain shortcomings, both from the esthetic and from the didactic point of view. It is no longer symmetric with respect to the two primitive symbols "$<$" and "$>$", certain properties of the relation $<$ being accepted without proof, while quite analogous properties of the relation $>$ have first to be demonstrated. Also missing in the new system is Axiom 6, which is of a very elementary and intuitively evident character, while its derivation from the axioms of System $\mathcal{A}^{(1)}$ might give the reader some difficulties.

56 Independence of the axioms of the reduced system

The question now arises whether there are any other superfluous axioms contained in System $\mathcal{A}^{(1)}$. It turns out that this is not the case:

$\mathcal{A}^{(1)}$ is a system of mutually independent axioms.

In order to establish the methodological statement just formulated, we can employ the method of proof by interpretation, which is analogous to the method of proof by exhibiting a model (see Sections 37 and 38).

We have to show that none of the axioms of System $\mathcal{A}^{(1)}$ is derivable from the remaining axioms of this system. We shall confine ourselves to Axiom $2^{(1)}$, as example. Suppose we replace the symbol "$<$" in the axioms of System $\mathcal{A}^{(1)}$ throughout by "$\leq$", without altering the axioms in any other way. In the course of this transformation every axiom, with the exception of Axiom $2^{(1)}$, retains its validity; in fact, Axioms $3^{(1)}$, $5^{(1)}$, $6^{(1)}$, and $7^{(1)}$ are left unaltered since they do not contain the symbol "$<$", and Axioms $1^{(1)}$, $4^{(1)}$, and $8^{(1)}$ go over into certain arithmetical theorems, whose proofs on the basis of System $\mathcal{A}$ or of $\mathcal{A}^{(1)}$ and the definition of the symbol "$\leq$" (cf. Section 46) do not present any difficulties. We may now conclude that the set $\mathbf{R}$ of all numbers, the relations $\leq$ and $>$, and the operation of addition form a model of Axioms $1^{(1)}$ and $3^{(1)}$–$8^{(1)}$; our discussion, moreover, shows directly that the system of these seven axioms has found a new interpretation within arithmetic. On the other hand, the sentence into which Axiom $2^{(1)}$ is transformed by the above replacement is false, for its negation can easily be established in arithmetic; the formula:

$$x \leq y$$

[††]We recall from Section 39 two kinds of equipollence: *equipollence in means of proof* and *equipollence in means of expression*. In this chapter, "*equipollence*" will ordinarily refer to means of proof (i.e. except in a passage of Section 58 and in exercise 16), and as before, the qualifying phrases will be left out.

does not always exclude:

$$y \leq x,$$

for there are numbers x and y which simultaneously satisfy the two inequalities:

$$x \leq y \text{ and } y \leq x$$

(of course, this is the case if, and only if, x and y are equal).† Therefore, if one believes in the consistency of arithmetic (cf. Section 41), one has to accept the fact that the sentence which is obtained from Axiom $2^{(1)}$ is not a theorem of this discipline. And from this it follows that Axiom $2^{(1)}$ is not derivable from the remaining axioms of System $\mathcal{A}^{(1)}$, as otherwise this axiom would necessarily be valid in any interpretation in which the other axioms hold (cf. the analogous considerations in Section 37).

By using such a method of reasoning but by applying other suitable interpretations, or other models, we can obtain the corresponding result for each of the other axioms.††

* * *

*In general, the method of proof by interpretation can be described as follows. It is a question of showing that a particular sentence A does not follow from a certain system $\mathcal{S}$ of statements (which may be axioms or not) of a given deductive theory. For this purpose, we consider an arbitrary deductive theory $\mathcal{T}$ (which we assume to be consistent; $\mathcal{T}$ may, furthermore, be the same theory to which the statements of the system $\mathcal{S}$ belong). We then try to find an interpretation of the system $\mathcal{S}$ within this theory, of such a kind that the negation of the sentence A becomes a theorem (or

†Let us look more closely at the argument just given. It involves the transformed Axiom $2^{(1)}$, which we can express in the schematic form: *for any x and y,* (...). This transformed axiom was negated by showing, in effect, that: *there exist x and y such that* [∼ (...)]. The assertion that these two statements contradict each other certainly agrees with our intuition, and it is implied by the following equivalence, which we state using symbols:

$$\sim \forall_{x,y}(\ldots) \leftrightarrow \exists_{x,y}[\sim (\ldots)].$$

In fact, we have here an example of a law of *predicate logic*. This branch of logic, we recall, was mentioned briefly in Section 20. (Note also that analogous equivalences were implicitly assumed in the indirect proofs of Chapters VII and VIII.)

††We should like to comment on the relation of the proof just described to a proof by exhibiting a model. In the above example we seem to have all the ingredients that are needed for a proof of the latter kind. There is a complication, however which is caused by the fact that the set of real numbers is infinite, and in case of infinite models one can encounter some rather intricate questions of consistency. An alternative is to proceed by way of an underlying theory, and this option takes us from a model to an interpretation. In such an approach the role of consistency appears to be conceptually simpler; cf. the discussion which follows. (Of course, the resolution of an ensuing problem can still be highly non-trivial.)—However, some of the independence proofs which are outlined in the exercises depend on simple (and finite) models, for which the interpretations, like those within arithmetic, can be disregarded.

possibly an axiom) of the theory $\mathcal{T}$, but not this sentence itself. If we are successful in doing so, we may invoke the law of deduction. As we know from Section 38, this law implies that if the sentence A could be derived from the statements of the system $\mathcal{S}$, it would then remain valid for any interpretation of this system. Consequently, the very existence of such an interpretation of $\mathcal{S}$, where A is not valid, would constitute a proof that this sentence cannot be derived from the system $\mathcal{S}$.—Let us now return to the condition of consistency, and let us state our conclusion with greater precision:

if the theory $\mathcal{T}$ is consistent, then the sentence A cannot be derived from the statements of the system $\mathcal{S}$.

It is easy to see why we must include the hypothesis, that the theory $\mathcal{T}$ is consistent. For otherwise the theory $\mathcal{T}$ could contain two contradictory sentences among its axioms and theorems, and from the fact that $\mathcal{T}$ contains the negation of A we could not conclude that $\mathcal{T}$ does not contain A itself (or rather, the interpretation of A); thus our argument would no longer be valid.

In order to prove the independence of a given axiom system in such a way, the above method has to be applied as many times as there are axioms in the system in question; each axiom in turn is taken as the sentence A, while $\mathcal{S}$ consists of the remaining axioms of the system.*

57 Elimination of a superfluous primitive term and the subsequent simplification of the axiom system; the concept of an ordered Abelian group

We return once more to the axiom system $\mathcal{A}^{(1)}$. Since this system is independent, it cannot be simplified any further by finding and omitting a superfluous axiom. Nevertheless, a simplification can be achieved in a different way. For it turns out that the primitive terms of System $\mathcal{A}^{(1)}$ are not mutually independent. In fact, either one of the two symbols "<" and ">" may be stricken from the list of primitive terms, and then it can be defined in terms of the other. This is easily seen from Theorem 3; on account of its form, this theorem may be considered as a definition of the symbol ">" by means of the symbol "<", and if in this theorem we exchange the two sides of the equivalence, we may look upon the resulting statement as a definition of the symbol "<" by means of the symbol ">". (If such an interpretation is adopted, it is desirable to have the phrase "*we say that*" precede the theorem; cf. Section 11.) From the heuristic point of view, this reduction of the primitive terms might provoke certain objections; for the terms "<" and ">" are equally clear in their meaning, and the relations denoted by them possess properties which are entirely analogous, so that it would appear a little artificial to consider one of these terms as immediately comprehensible, while the other has to be defined with the help of the first. But these objections cannot be regarded as convincing.

* * *

If we now disregard the above objections, and resolve to eliminate one of the symbols in question from the list of primitive terms, we then have the task of casting our axiom system into such a form that no defined terms occur in it. (This is a methodological postulate which, by the way, in practice is frequently ignored; in geometry, especially, the axioms are usually formulated with the help of defined terms in order to enhance their simplicity and intuitive appeal.) This task does not present any difficulties; we simply replace in the axiom system $\mathcal{A}^{(1)}$ every formula such as:

$$x > y$$

by the corresponding formula:

$$y < x$$

which, by Theorem 3, is equivalent to it. It is easy to see that after this transformation, Axiom 1 (or, $1^{(1)}$) may be replaced by the law of connectivity, Theorem 4, since each sentence follows from the other on the basis of general laws of logic (in fact, one needs here only certain laws of sentential calculus, if the universal quantifiers are neglected); moreover, Axiom 3 becomes a simple substitution of Axiom 2, and so may be omitted altogether. In this way we arrive at the system consisting of the following seven axioms:

AXIOM $1^{(2)}$. *If* $x \neq y$, *then* $x < y$ *or* $y < x$.

AXIOM $2^{(2)}$. *If* $x < y$, *then* $y \not< x$.

AXIOM $3^{(2)}$. *If* $x < y$ *and* $y < z$, *then* $x < z$.

AXIOM $4^{(2)}$. $x + y = y + x$.

AXIOM $5^{(2)}$. $x + (y + z) = (x + y) + z$.

AXIOM $6^{(2)}$. *For any numbers* x *and* y *there exists a number* z *such that* $x = y + z$.

AXIOM $7^{(2)}$. *If* $y < z$, *then* $x + y < x + z$.

This axiom system, to be called SYSTEM $\mathcal{A}^{(2)}$, thus appears to be equipollent to each of the two former systems, $\mathcal{A}$ and $\mathcal{A}^{(1)}$. However, in saying this we commit one inexactitude; for it is impossible to derive from the axioms of System $\mathcal{A}^{(2)}$ those sentences of System $\mathcal{A}$ or of $\mathcal{A}^{(1)}$ which contain the symbol ">", unless System $\mathcal{A}^{(2)}$ is enlarged by adjoining the definition of this symbol. As we know, we may give this definition the following form:

DEFINITION $1^{(2)}$. *We say that* $x > y$ *if, and only if,* $y < x$.

We also know that this last sentence can be proved on the basis of System $\mathcal{A}$ or of $\mathcal{A}^{(1)}$ provided it is treated not as a definition, but as an ordinary theorem (so that one omits the initial phrase "*We say that*"). The fact that these two systems and the enlarged system are mutually equipollent can now be expressed as follows:

System $\mathcal{A}^{(2)}$ together with Definition $1^{(2)}$ is equipollent to each of Systems $\mathcal{A}$ and $\mathcal{A}^{(1)}$.

Such a cautious mode of expression should be retained whenever we compare two axiom systems which might be regarded as equipollent, but which contain even in part different primitive terms.

* * *

An attractive feature of the axiom system $\mathcal{A}^{(2)}$ is the striking simplicity of its structure. The first three axioms concern the relation *less than*, and together they assert that the set **R** is ordered by this relation; the next three axioms are concerned with addition, and they assert that the set **R** is an Abelian group with respect to addition; the last axiom finally—the law of monotonicity—imposes a certain interrelation between the relation *less than* and the operation of addition. We can describe this structure in a more general way: a class K is said to be an ORDERED ABELIAN GROUP WITH RESPECT TO THE RELATION R AND THE OPERATION O if (i) the class K is ordered by the relation R, (ii) the class K is an Abelian group with respect to the operation O, and (iii) the operation O is monotonic in the class K with respect to the relation R. In accordance with this terminology, we can now say that *the set of numbers is characterized by the axiom system $\mathcal{A}^{(2)}$ as an ordered Abelian group with respect to the relation* less than *and the operation of addition.*

* * *

The following facts concerning System $\mathcal{A}^{(2)}$ can be established[†]:

*System $\mathcal{A}^{(2)}$ is an independent axiom system, and moreover its three primitive terms, namely "**R**", "<", and "+", are mutually independent.*

We omit the proof of this statement. We remark only that, in order to establish the mutual independence of the primitive terms, one can again apply the method of proof by interpretation; in this case the method assumes a more involved form, and lack of space prevents us from going into the modifications which would be required for such a purpose.

[†]Some of the questions which are raised in this chapter will be taken up again in the companion volume. In particular, elementary proofs of the independence of the three primitive terms and the consistency of ordered Abelian groups will be outlined.

58 Further simplification of the axiom system; possible transformations of the system of primitive terms

One can obviously find many axiom systems which are equipollent to System $\mathcal{A}^{(2)}$. We shall give here a particularly simple example of such a system, which will be called SYSTEM $\mathcal{A}^{(3)}$, and which contains the same primitive terms as $\mathcal{A}^{(2)}$. It consists of only five sentences:

AXIOM $1^{(3)}$. *If* $x \neq y$, *then* $x < y$ *or* $y < x$.

AXIOM $2^{(3)}$. *If* $x < y$, *then* $y \not< x$.

AXIOM $3^{(3)}$. $x + (y + z) = (x + z) + y$.

AXIOM $4^{(3)}$. *For any numbers* x *and* y *there exists a number* z *such that* $x = y + z$.

AXIOM $5^{(3)}$. *If* $x + z < y + t$, *then* $x < y$ *or* $z < t$.

We shall show that

Systems $\mathcal{A}^{(2)}$ *and* $\mathcal{A}^{(3)}$ *are equipollent.*

Proof. We observe, first of all, that all the axioms of System $\mathcal{A}^{(3)}$ are either contained in System $\mathcal{A}$ (thus, Axiom $2^{(3)}$ coincides with Axiom 2, and Axiom $4^{(3)}$ with Axiom 9), or else have been proved on its basis (Axioms $1^{(3)}$, $3^{(3)}$, and $5^{(3)}$ have been established as Theorems 4, 9, and 14, respectively). But since the axiom systems $\mathcal{A}$ and $\mathcal{A}^{(2)}$ are equipollent, as we know from Section 57 (Definition $1^{(2)}$, after all, can always be added to System $\mathcal{A}^{(2)}$), we may conclude that all the sentences of System $\mathcal{A}^{(3)}$ can be proved on the basis of System $\mathcal{A}^{(2)}$.† It remains to derive those sentences of System $\mathcal{A}^{(2)}$ from the axioms of System $\mathcal{A}^{(3)}$ which are absent in $\mathcal{A}^{(3)}$, namely Axioms $3^{(2)}$, $4^{(2)}$, $5^{(2)}$, and $7^{(2)}$. This task is not quite so simple.

*We begin with Axioms $4^{(2)}$ and $5^{(2)}$.

(I) *Axiom* $4^{(2)}$ *can be derived from the axioms of System* $\mathcal{A}^{(3)}$.

† A more precise way of describing the situation is as follows. System $\mathcal{A}^{(2)}$, when enlarged by adjoining Definition $1^{(2)}$, is equipollent to System $\mathcal{A}$. However, one can show (by techniques belonging to the domain of *metamathematics*) that any consequence of the enlarged system which does not involve the symbol ">" can also be derived from System $\mathcal{A}^{(2)}$ alone, without the use of Definition $1^{(2)}$. Since all the sentences of System $\mathcal{A}^{(3)}$ have been deduced from System $\mathcal{A}$, we conclude (by equipollence) that they can also be deduced from the enlarged system, and therefore (by the above remarks), from System $\mathcal{A}^{(2)}$.— A reader, who wishes to confirm the latter conclusion by elementary means, can examine the requisite derivations one by one and verify that the definition can be avoided in each case.

For, given two numbers x and y, we can apply Axiom $4^{(3)}$ (with "x" put in place of "y", and vice versa); therefore there is a number z such that

(1) $$y = x + z.$$

If in Axiom $3^{(3)}$ we now replace "y" by "x", we obtain:

(2) $$x + (x + z) = (x + z) + x.$$

In view of (1) "$x + z$" may be replaced here by "y" on both sides, and we arrive at Axiom $4^{(2)}$:

$$x + y = y + x.$$

(II) *Axiom $5^{(2)}$ can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

In fact, we obtain from Axiom $3^{(3)}$ (by substituting "y" for "z", and vice versa):

$$x + (z + y) = (x + y) + z;$$

on account of the commutative law, which has already been derived by (I), we may replace here "$z + y$" by "$y + z$", and we obtain Axiom $5^{(2)}$:

$$x + (y + z) = (x + y) + z.$$

In order to facilitate the derivation of Axioms $3^{(2)}$ and $7^{(2)}$, we shall first show how some of the axioms and theorems which were stated in the preceding chapters may be proved on the basis of System $\mathcal{A}^{(3)}$.

(III) *Axiom 6 can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

In fact, we saw in Section 55 that Axiom 6 can be deduced from Axioms 7, 8, and 9. Axioms 7 and 8 are the same as Axioms $4^{(2)}$ and $5^{(2)}$, and therefore, by (I) and (II), can be derived from the axioms of System $\mathcal{A}^{(3)}$. Axiom 9, on the other hand, occurs as Axiom $4^{(3)}$ in System $\mathcal{A}^{(3)}$. Hence, Axiom 6 is a consequence of the axioms of $\mathcal{A}^{(3)}$.

(IV) *Theorem 11 can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

In the second proof of Theorem 11, as given in Section 49, only Axioms 7, 8, and 9 were used. Theorem 11 is therefore derivable from the axioms of System $\mathcal{A}^{(3)}$ for the same reason as Axiom 6; see (III).

(V) *Theorem 1 can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

We simply observe that the proof of Theorem 1 as given in Section 44 is based exclusively upon Axiom 2, which in turn coincides with Axiom $2^{(3)}$ of System $\mathcal{A}^{(3)}$.

(VI) *Theorem 12 can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

For, assume the hypothesis of Theorem 12:

$$x + y < x + z;$$

we apply Axiom $5^{(3)}$, having replaced "z", "y", and "t" by "y", "x", and "z", respectively. It follows that one of the formulas:

$$x < x \text{ or } y < z$$

must hold; the first possibility has to be rejected because it contradicts Theorem 1, which has already been shown to be derivable from System $\mathcal{A}^{(3)}$,— cf. (V). Hence the conclusion of Theorem 12 must hold:

$$y < z.$$

(VII) *Axiom $3^{(2)}$ can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

Let us assume the hypothesis of Axiom $3^{(2)}$, that is, the formulas:

(1) $$x < y$$

and

(2) $$y < z.$$

By Axiom 6, which has been derived, cf. (III), there are the two numbers $y + x$ and $y + z$; if we had:

$$y + x = y + z,$$

it would then follow from Theorem 11, already derived by (IV), that

$$x = z.$$

Then in (1) "x" could be replaced by "z", which would lead to:

$$z < y.$$

This inequality would contradict (2) by virtue of Axiom $2^{(3)}$, and must therefore be rejected. We thus have:

(3) $$y + x \neq y + z.$$

Since $y + x$ and $y + z$ are numbers, we may infer from (3) by Axiom $1^{(3)}$ that one of these two cases must hold:

(4) $$y + x < y + z \text{ or } y + z < y + x.$$

Let us consider the second of the formulas in (4), where we may replace "$y+x$" by "$x+y$" by virtue of Axiom $4^{(2)}$, which has already been derived; we now arrive at:

$$y + z < x + y.$$

To this formula we apply Axiom $5^{(3)}$, where we replace "x" and "t" by "y", and "y" by "x". We obtain in this way the following consequence:

$$y < x \ \ or \ \ z < y.$$

But this has to be rejected, since (in view of Axiom $2^{(3)}$) it contradicts (1) and (2), which constitute the hypothesis of Axiom $3^{(2)}$. We therefore return to the first of the formulas in (4), and we apply Theorem 12, which we know from (VI) to be derivable (and in which we replace "x" by "y" and vice versa); thus we obtain:

$$x < z,$$

and this is the conclusion of Axiom $3^{(2)}$.

(VIII) *Axiom $7^{(2)}$ can be derived from the axioms of System $\mathcal{A}^{(3)}$.*

Here the procedure is similar to the one just applied, but undoubtedly simpler. We assume the hypothesis of Axiom $7^{(2)}$:

(1) $$y < z.$$

If now we had:

$$x + y = x + z,$$

it would follow by Theorem 11 that

$$y = z;$$

we could therefore replace "y" by "z" in (1) and arrive at a contradiction to Theorem 1, derived above by (V). Hence we must have:

$$x + y \neq x + z,$$

from which it follows by Axiom $1^{(3)}$ that

(2) $$x + y < x + z \ \ or \ \ x + z < x + y.$$

So far, this derivation resembles closely the preceding one; now we apply Theorem 12 to the second of these inequalities, obtaining:

$$z < y,$$

but this contradicts our hypothesis (1) by virtue of Axiom $2^{(3)}$. Consequently we have to accept the first of the inequalities in (2):

$$x + y < x + z,$$

and this is the conclusion of Axiom $7^{(2)}$.*

We have therefore seen that all the sentences of System $\mathcal{A}^{(2)}$ are consequences of System $\mathcal{A}^{(3)}$, and conversely; the two axiom systems $\mathcal{A}^{(2)}$ and $\mathcal{A}^{(3)}$ are accordingly established as equipollent.

* * *

System $\mathcal{A}^{(3)}$, no doubt, is simpler than System $\mathcal{A}^{(2)}$, and hence still simpler than each of Systems $\mathcal{A}$ and $\mathcal{A}^{(1)}$. Particularly instructive is a comparison of the two systems $\mathcal{A}$ and $\mathcal{A}^{(3)}$; as a result of the successive reductions that were carried out, the original number of axioms has diminished by more than one half. On the other hand, it should be noted that some of the sentences of System $\mathcal{A}^{(3)}$ (namely, Axioms $3^{(3)}$ and $5^{(3)}$) are less natural and less simple than the axioms of the other systems, and also that the proofs of several theorems, even of very elementary ones, are rather more difficult and more involved than the proofs on the basis of those other systems.

* * *

Just as in the case of axioms, one can find systems of primitive terms which are equipollent to a given system. This applies, in particular, to the system of the three terms "**R**", "$<$", and "$+$", which occur as the only primitive terms in the axiom system that was considered last. If, for instance, in this system we replace the symbol "$<$" by the symbol "$\leq$", we obtain an equipollent system; for the second of these symbols was defined in terms of the first, and Theorem 8 tells us how the first may be defined by means of the second. But such a transformation of the system of primitive terms would not be in any way advantageous, and in particular, it would contribute nothing to a simplification of the axioms; moreover, if the reader should be more familiar with the symbol "$<$" than with "$\leq$", it might even appear rather artificial. Another equipollent system can be obtained by replacing in the original system the symbol "$+$" by "$-$"; but, again, this transformation would not be at all expedient.—In conclusion let us mention the fact, that systems of primitive terms are known which are equipollent to the system in question and consist of but two terms.

59 The problem of consistency of the constructed theory

We shall now touch briefly upon two fundamental problems which concern the fragment of arithmetic considered above and its methodology; these are the problems of consistency and completeness (cf. Section 41). Because of equipollence our remarks apply equally well to each of the above axiom systems, and we shall now refer only to System $\mathcal{A}$.

If we believe in the consistency of the full arithmetic of real numbers (this assumption was made previously and will be made again in our subsequent considerations), then we must all the more accept the fact that:

The mathematical theory which is based on System $\mathcal{A}$ is consistent.

But while every attempt to give a strict proof of consistency of the full arithmetic necessarily leads to essential difficulties (cf. Section 41), a proof of this kind for System $\mathcal{A}$ is not only possible but even comparatively simple. One reason for this is the fact, that the variety of theorems which can be derived from the axiom system $\mathcal{A}$ is very small indeed; on its basis, for instance, it is not possible to answer the very elementary question, whether any distinct numbers exist at all. This circumstance facilitates greatly the proof, that the part of arithmetic under consideration does not contain a single pair of contradictory theorems. With the means which are at our disposal, however, it would be a hopeless undertaking to sketch the proof of consistency, or even to try to acquaint the reader with its fundamental idea; this would require a much deeper knowledge of logic, and an essential preliminary task would be the reconstruction of the part of arithmetic in question as a formalized deductive theory (cf. Section 40). It may be added, that if System $\mathcal{A}$ is enlarged by a single sentence to the effect that at least two distinct numbers exist, then any attempt to prove the consistency of the axiom system which is extended in this way will meet with serious complications; they will be of comparable magnitude as those which are encountered in case of various more comprehensive axiom systems for real numbers.[†]

[†] The reader might not have much appreciation at this stage for the differences among the various arithmetical theories. The over-all situation is perhaps somewhat intricate, and it may be helpful to rephrase and to supplement the above résumé. A proof of consistency of an elementary theory of real numbers (see Section 41, and also below) usually depends on the solution of serious technical problems; such problems, however, can be resolved. These remarks apply in particular to System $\mathcal{A}$ when it is enlarged as described. (Such problems, furthermore, can be avoided in special cases, like that of System $\mathcal{A}$ without the enlargement.) On the other hand, when one deals with the arithmetic of natural numbers or of all integers, and includes the four basic operations on the respective numbers, then a proof of consistency has to involve difficulties of a fundamental nature.—This essential difference extends also to the question of completeness. The intrinsic incompleteness, which was described in *loc. cit.*, applies in particular to the latter theories. For contrast, one can construct an elementary theory of real numbers which is complete, by choosing the primitive terms and the axioms in an appropriate way.—The distinction between elementary and non-elementary theories was brought out already in *loc. cit.* We repeat here: In elementary theories the concept of a set does not appear (whether one would consider a set of

60 The problem of completeness of the constructed theory

In comparison with the question of consistency, that of completeness of System $\mathcal{A}$ can be settled much more readily.

There are numerous problems which can be formulated by using exclusively logical terms and the primitive terms of System $\mathcal{A}$, but which cannot be resolved on the basis of this system. One such problem was mentioned already in the preceding section. Another example is given by the sentence stating that for any number x, there exists a number y such that

$$x = y + y.$$

On the basis of the axioms of System $\mathcal{A}$ alone, it is impossible either to prove or to disprove this sentence. This can be seen from the following consideration. We have been denoting by the symbol "**R**" the set of all real numbers; in other words, the set **R** contains as elements the integers and the proper fractions, the rational numbers as well as the irrational ones. But one can see at once that all the axioms would retain their validity, and hence, so would all theorems which follow from them, if we were to denote by the symbol "**R**" either the set of all integers (the positive and the negative ones as well as the number 0), or the set of all rational numbers; that is to say, all these statements would remain valid if the word "*number*" meant either "*integer*" or "*rational number*". In the first case, the above sentence, which states that for any given number there is another number half as large, would be false; in the second case it would be true. If, therefore, we succeeded in proving this sentence on the basis of System $\mathcal{A}$, we would arrive at a falsehood when referring to the set of integers; on the other hand, if we were able to disprove it, we would find ourselves involved in a falsehood in the case of the set of rational numbers.

The argument which has just been sketched falls under the category of proofs by exhibiting a model, and it can be reformulated as a proof by interpretation, as follows (cf. Sections 37 and 56). Let "**Z**" denote the set of all integers, and "**Q**", the set of all rational numbers. We shall now give two interpretations of System $\mathcal{A}$ within arithmetic. The symbols "$<$", "$>$", and "$+$" remain unchanged in both interpretations, while the symbol

---

numbers, of points, etc.). This restriction means, in particular, that in the elementary theories of real numbers one cannot introduce the set of natural numbers nor the set of integers, and for this reason one can bypass the aforementioned fundamental difficulties. With regard to non-elementary theories, we have as example the full arithmetic of real numbers, which will be discussed in the next chapter.—In view of the above considerations, it may be appropriate to return here to the stipulation which was made in the beginning of Chapter VII: "... *logic is the only preceding theory.*" We see that this statement has to be qualified; one should specify which branches of logic are assumed.

"**R**", which occurs explicitly or implicitly in each of the axioms, is to be replaced by "**Z**" in the first and by "**Q**" in the second interpretation.† (We disregard here the remarks which were made in Section 43 about the possible elimination of the symbol "**R**", since this would complicate our reasoning slightly.) All the axioms of System $\mathcal{A}$ retain their validity in both interpretations; the sentence:

for every number x there exists a number y such that
$$x = y + y,$$

however, is valid only in the case of the second interpretation, while in the case of the first interpretation its negation holds:

not for every number x is there a number y such that $x = y + y$.

Having assumed the consistency of arithmetic, we conclude from the first interpretation that the sentence in question cannot be proved on the basis of System $\mathcal{A}$, and from the second interpretation we conclude that it also cannot be disproved.

We have thus shown that there exist two contradictory sentences, formulated exclusively in logical terms and in the primitive terms of the mathematical theory which we have been considering, and such that neither of them can be derived from the axioms of that theory. Consequently we have the conclusion††:

The mathematical theory based on System $\mathcal{A}$ is incomplete.

†The reader may recall the previous footnote and ask: Which of the arithmetical theories should we assume here? Clearly, it suffices to choose a theory in which the five symbols "**Z**", ..., "+" occur, or can be defined (but not necessarily "**R**"). The full arithmetic of real numbers, as in the next chapter, could be used, and one can also find simpler fragments which are suitable.

††With the matters of consistency and completeness settled, let us recall that there is still a third related problem for a given deductive theory, namely the decision problem (see Section 41). It turns out that the theory of ordered Abelian groups is decidable. That is to say, there is an algorithm which would allow one to determine whether a given sentence can be derived from the axioms or not.—Such an algorithm necessarily has to exploit the form of the axioms and of the theorems, and therefore, an essential first step for investigating this problem would be to formulate the theory in question as a formalized theory. Furthermore, the algorithm applies only when this theory is regarded as an elementary theory (as in the last three chapters), i.e., concepts which are specific to the theory of sets are to be avoided. One therefore has to allow in the sentences only the following expressions: variables ranging over **R** and the associated quantifiers, the symbols "<", "+", and "=", the sentential connectives, and of course parentheses. (Cf. also Section 59, and in particular footnote on pp. 194–195.)

Exercises

1. Let us agree that the formula:

$$x \otimes y$$

means the same as:

$$x + 1 < y.$$

In the axioms of System $\mathcal{A}^{(2)}$ of Section 57 now replace the symbol "$<$" throughout by "$\otimes$", determine which of the axioms retain their validity and which do not, and show in this way that Axiom $1^{(2)}$ cannot be derived from the remaining axioms. What is the name of the method which is applied here?

2. By adapting the independence proof which was sketched in Section 56 for Axiom $2^{(1)}$, show that Axiom $2^{(2)}$ cannot be derived from the remaining axioms of System $\mathcal{A}^{(2)}$.

3. Let the symbol "$\overset{\circ}{\mathbf{R}}$" designate the set which consists of the three numbers 0, 1, and 2. For the elements of this set we define a relation $\overset{\circ}{<}$, stipulating that it should hold only in these three cases:

$$0 \overset{\circ}{<} 1, \quad 1 \overset{\circ}{<} 2, \quad 2 \overset{\circ}{<} 0.$$

Moreover, we define the operation $\overset{\circ}{+}$ on the elements of the set $\overset{\circ}{\mathbf{R}}$ by the following formulas:

$$0 \overset{\circ}{+} 0 = 1 \overset{\circ}{+} 2 = 2 \overset{\circ}{+} 1 = 0,$$

$$0 \overset{\circ}{+} 1 = 1 \overset{\circ}{+} 0 = 2 \overset{\circ}{+} 2 = 1,$$

$$0 \overset{\circ}{+} 2 = 1 \overset{\circ}{+} 1 = 2 \overset{\circ}{+} 0 = 2.$$

In the axioms of System $\mathcal{A}^{(2)}$ now replace the primitive terms of that system by "$\overset{\circ}{\mathbf{R}}$", "$\overset{\circ}{<}$", and "$\overset{\circ}{+}$" respectively (and let the word "*number*" now mean "*one of the three numbers 0, 1, and 2*"); show in this way that Axiom $3^{(2)}$ cannot be derived from the remaining axioms.

4. In order to show by means of a proof by interpretation that Axiom $4^{(2)}$ cannot be derived from the remaining axioms of System $\mathcal{A}^{(2)}$, it suffices to replace the symbol of addition in all the axioms by the symbol of one of the four operations mentioned in exercise 3 of Chapter VIII. Which is the operation that has to be used?

5. Consider the operation $\oplus$ satisfying the following formula:

$$x \oplus y = 2 \cdot (x + y).$$

Show with the help of this operation, that Axiom $5^{(2)}$ cannot be deduced from the other axioms of System $\mathcal{A}^{(2)}$.

6. Construct a set of numbers of such a kind that, together with the relation $<$ and the operation $+$, it fails to satisfy Axiom $6^{(2)}$ but forms a model of the remaining axioms of System $\mathcal{A}^{(2)}$. What conclusion may be drawn from such a construction, regarding the possibility of deriving Axiom $6^{(2)}$?

7. In order to show that Axiom $7^{(2)}$ is not a consequence of the other axioms of System $\mathcal{A}^{(2)}$, one can proceed by replacing (in all the axioms) two of the primitive terms of the system by the symbols which were introduced in exercise 3 (and which correspond to these primitive terms), while leaving the third primitive term unchanged. Determine which term should be left unchanged.

8. The results obtained in exercises 1–7 show that none of the axioms of System $\mathcal{A}^{(2)}$ can be derived from the remaining axioms of this system. Carry out analogous independence proofs for the axiom systems $\mathcal{A}^{(1)}$ of Section 55 and $\mathcal{A}^{(3)}$ of Section 58 (using, in part, the interpretations which were applied in the preceding exercises).

9. Show, by referring to the axiom system $\mathcal{A}^{(2)}$, that if a set of numbers is an Abelian group with respect to the operation of addition, then it is at the same time an ordered Abelian group with respect to the relation *less than* and the operation of addition. Give examples of sets of numbers of this kind.

10. In exercise 5 of Chapter VIII several sets of numbers were given which form Abelian groups with respect to multiplication. Which of these sets are ordered Abelian groups with respect to the relation *less than* and the operation of multiplication, and which are not?

11. With the help of the conclusions which were reached in exercise 10, construct a new proof of the independence of Axiom $7^{(2)}$ from the remaining axioms of System $\mathcal{A}^{(2)}$ (cf. exercise 7).

***12.** Derive the following theorem from the axiom system $\mathcal{A}^{(2)}$:

if there are at least two different numbers, then, for any number x, there is a number y such that $x < y$.

As a generalization of this result, prove the following general group-theoretical theorem:

> *If the class K is an ordered Abelian group with respect to the relation R and the operation O, and if K has at least two elements, then, for any element x of K, there exists an element y of K such that xRy.*

Show with the help of this theorem, that no class which is an ordered Abelian group can consist of exactly two, or three, and so on, elements. Can it consist of just one element? (Cf. exercise 8 of Chapter VIII.)

***13.** Show that the system of Axioms $1^{(2)}$–$3^{(2)}$ (of Section 57) is equipollent to the system consisting of Axiom $1^{(2)}$ and the following sentence:

if $x < y,\ y < z,\ z < t,\ t < u,$ *and* $u < v,$ *then* $v \not< x$.

As a generalization of this result, establish the following general law of the theory of relations:

For the class K to be ordered by the relation R, it is necessary and sufficient that R be connected in K and that it satisfy the following condition:

if x, y, z, t, u, and v are any elements of K, and if xRy, yRz, zRt, tRu, and uRv, then it is not the case that vRx.

***14.** Use the considerations of Sections 48, 55, and 58 to show that the following three systems of sentences are equipollent:

(a) the system of Axioms 6–9 of Section 47;
(b) the system of Axioms $4^{(2)}$–$6^{(2)}$ of Section 57;
(c) the system of Axioms $3^{(3)}$ and $4^{(3)}$ of Section 58.

Generalize this result by formulating new definitions of the expression:

the class K is an Abelian group with respect to the operation O;

these should be equivalent to the definition given in Section 47, but simpler; write the relevant axioms in symbols (cf. exercise *29 of Chapter VIII).

***15.** Consider System $\mathcal{A}^{(4)}$, consisting of the following five axioms:

AXIOM $1^{(4)}$. *If* $x \neq y,$ *then* $x < y$ *or* $y < x$.
AXIOM $2^{(4)}$. *If* $x < y,\ y < z,\ z < t,\ t < u,$ *and* $u < v,$ *then* $v \not< x$.
AXIOM $3^{(4)}$. $x + (y + z) = (x + z) + y$.
AXIOM $4^{(4)}$. *For any numbers* x *and* y *there exists a number* z *such that* $x = y + z$.
AXIOM $5^{(4)}$. *If* $y < z,$ *then* $x + y < x + z$.

Use the results of exercises *13 and *14 to show that System $\mathcal{A}^{(4)}$ is equipollent to each of Systems $\mathcal{A}^{(2)}$ and $\mathcal{A}^{(3)}$.

16. In Section 58 it was pointed out that the system of the three primitive terms "**R**", "<", and "+" is equipollent to the system of the terms "**R**", "≤", and "+"; to this assertion it should really have been added, that these systems are equipollent with respect to a certain system of sentences, for instance, with respect to System $\mathcal{A}^{(3)}$ of Section 58 and Definition 1 of Section 46. Explain why such a qualification is indispensable. In general, why is it always necessary to assume a particular system of sentences, when intending to establish the equipollence of two systems of terms (in the sense of Section 39)?

***17.** Consider System $\mathcal{A}^{(5)}$ which consists of the following seven sentences:

AXIOM $1^{(5)}$. *For any numbers* x *and* y *we have:* $x \leq y$ *or* $y \leq x$.
AXIOM $2^{(5)}$. *If* $x \leq y$ *and* $y \leq x$, *then* $x = y$.
AXIOM $3^{(5)}$. *If* $x \leq y$ *and* $y \leq z$, *then* $x \leq z$.
AXIOM $4^{(5)}$. $x + y = y + x$.
AXIOM $5^{(5)}$. $x + (y + z) = (x + y) + z$.
AXIOM $6^{(5)}$. *For any numbers* x *and* y *there exists a number* z *such that* $x = y + z$.
AXIOM $7^{(5)}$. *If* $y \leq z$, *then* $x + y \leq x + z$.

Show that the axiom systems $\mathcal{A}^{(2)}$ (of Section 57) and $\mathcal{A}^{(5)}$ become equipollent systems of sentences, if Definition 1 of Section 46 is added to the first, and Theorem 8 of Section 46 to the second, considering the latter theorem as a definition of the symbol "$<$". Why may we not simply say that Systems $\mathcal{A}^{(2)}$ and $\mathcal{A}^{(5)}$ are equipollent?

18. By adapting the line of argument taken in Section 60, show that the following sentence can be neither proved nor disproved on the basis of System $\mathcal{A}$:

if $x < z$, *then there exists a number* y *such that* $x < y$ *and* $y < z$.

***19.** Show that the following sentence can be neither proved nor disproved, if we base our arguments on System $\mathcal{A}$:

$$\forall_x \exists_{y,z}[(y < x) \land (x < z)].$$

***20.** In the present chapter we used the method of proof by interpretation in order to establish the independence or the incompleteness of a given axiom system. This method can also be employed in investigations of its consistency. In fact, we have the following methodological law at our disposal, which expresses a consequence of the law of deduction:

If the deductive theory $\mathcal{S}$ *has an interpretation in the deductive theory* $\mathcal{T}$ *and the theory* $\mathcal{T}$ *is consistent, then the theory* $\mathcal{S}$ *is also consistent.*

Justify this statement. In Section 38 some remarks were made concerning certain interpretations of arithmetic and geometry; applying the law just given, draw from these remarks appropriate conclusions regarding the consistency of arithmetic and geometry and its connection with the consistency of logic.†

† The reader should remember that the theory of sets plays a key role in these interpretations. The reader, moreover, should recall the discussion of consistency proofs in Section 41, and realize that such interpretations allow only limited conclusions regarding the consistency of theories. (One uses the term RELATIVE CONSISTENCY under such circumstances.)

X

Extension of the Constructed Theory: Foundations of Arithmetic of Real Numbers

61 The first axiom system for the arithmetic of real numbers

The axiom system $\mathcal{A}$ does not provide an adequate foundation for the whole of the arithmetic of real numbers, because—as we saw in Section 60—numerous theorems of this discipline cannot be deduced from the axioms of this system, and also for another reason, which is analogous and not less important: a number of concepts belonging to the field of arithmetic are not definable if we utilize only the primitive terms which occur in System $\mathcal{A}$. For example, on its basis we shall never be able to define the symbols of multiplication or division, or such symbols as "1", "2", and so on.

The following question then arises in a natural way: how can we transform or supplement our system of axioms and primitive terms in order to arrive at a satisfactory basis for constructing the full arithmetic of real numbers? This problem can be solved in a variety of ways. Two quite different methods of solution will be sketched here.[1]

[1]The first axiom system for the full arithmetic of real numbers was published by HILBERT in 1900; this system is related to System $\mathcal{A}''$ with which we shall soon become acquainted. Before the year 1900, axiom systems for certain less comprehensive parts of arithmetic had been known; one system of this kind concerned the arithmetic of natural numbers and was given in 1889 by PEANO (cf. footnote 2 on p. 111). Several axiom systems for arithmetic and for various parts of it—and, in particular, the first axiom system for the arithmetic of complex numbers—were published by the American mathematician E. V. HUNTINGTON (1874–1952), who made many important contributions to the axiomatic foundations of logical and of mathematical theories.

* * *

In the case of the first method, we choose as our point of departure System $\mathcal{A}^{(3)}$ (cf. Section 58); to the primitive terms appearing in that system we add the word "*one*" which, as usual, will be replaced by the symbol "1", and the axioms of the system are supplemented by four new sentences. In this way the new SYSTEM $\mathcal{A}'$ is obtained, containing the four primitive terms "**R**", "<", "+", and "1", and consisting of nine axioms which we write out explicitly below:

AXIOM 1′. *If* $x \neq y$, *then* $x < y$ *or* $y < x$.

AXIOM 2′. *If* $x < y$, *then* $y \not< x$.

AXIOM 3′. *If* $x < z$, *then there exists a number* y *such that* $x < y$ *and* $y < z$.

AXIOM 4′. *If* S *and* T *are any sets of numbers (*that is, $S \subseteq \mathbf{R}$ *and* $T \subseteq \mathbf{R}$*) satisfying the condition:*

for any x belonging to S and any y belonging to T,
we have: $x < y$,

then there exists a number z for which the following condition holds:

if x is any element of S and y any element of T, and
if $x \neq z$ *and* $y \neq z$, *then* $x < z$ *and* $z < y$.

AXIOM 5′. $x + (y + z) = (x + z) + y$.

AXIOM 6′. *For any numbers x and y there exists a number z such that* $x = y + z$.

AXIOM 7′. *If* $x + z < y + t$, *then* $x < y$ *or* $z < t$.

AXIOM 8′. $1 \in \mathbf{R}$.

AXIOM 9′. $1 < 1 + 1$.

62 Closer characterization of the first axiom system; its methodological advantages and didactic disadvantages

The axioms that are listed in the preceding section fall into three groups. In the first group, consisting of Axioms 1′–4′, only the two primitive terms "**R**" and "<" occur; in the second group, to which Axioms 5′–7′ belong, we have moreover the symbol "+"; finally, in the third group, which includes Axioms 8′ and 9′, the new symbol "1" appears.

Among the axioms of the first group there are two which we have not encountered before, namely Axioms 3′ and 4′. Axiom 3′ is called the LAW

OF DENSITY for the relation *less than*—it expresses the fact that this relation is dense in the set of all numbers. In general, we say that the relation R is DENSE IN THE CLASS K if, for any two elements x and y of this class, the formula:

$$xRy$$

always entails the existence of an element z of the class K for which

$$xRz \text{ and } zRy$$

hold. Axiom 4′ is known as the LAW OF CONTINUITY for the relation *less than*, or, as the AXIOM OF CONTINUITY, or, as DEDEKIND'S AXIOM[2]; in order to state a general condition for the relation R to be CONTINUOUS IN THE CLASS K, it suffices to replace in Axiom 4′ "**R**" by "K" (and, in connection with this, the word "*number*" by the expression "*element of the class* K"), and further, "<" by "R". If, in particular, the class K is ordered by the relation R, and if in addition R is dense in K or continuous in K, then K is said to be DENSELY ORDERED or CONTINUOUSLY ORDERED, respectively.

Axiom 4′ is intuitively less evident and more complicated than the remaining axioms; for one thing, it differs from the other axioms inasmuch as it deals not with individual numbers, but with sets of numbers. In order to give this axiom a simpler and more comprehensible form, it is convenient to introduce first the following definitions:

We say that the set of numbers S PRECEDES *the set of numbers T if, and only if, every number of S is less than every number of T.*

We say that the number z SEPARATES *the set of numbers S from the set T if, and only if, for any two elements x of S and y of T, both distinct from z, we have: $x < z$ and $z < y$.*

By referring to these definitions, we can give the axiom of continuity the following very simple formulation:

If one set of numbers precedes another, then there exists at least one number separating the first set from the second.

All the axioms of the second group are already known to us from earlier considerations. The axioms of the third group, though new, have so simple and obvious a content that they hardly require any comment. We might only make the remark, that if Axiom 9′ is preceded by definitions of the

[2]This axiom—in a slightly more complicated formulation—was proposed by the German mathematician R. DEDEKIND (1831–1916), whose researches contributed a great deal to the foundations of arithmetic, and especially, to the theory of irrational numbers.

symbol "0" and the expression "*positive number*", then it may be replaced either by the formula:

$$0 < 1$$

or else by the sentence:

1 is a positive number.

Let us now observe that Axioms 1′, 2′, 5′, 6′, and 7′ form just what we called System $\mathcal{A}^{(3)}$, which—like the equipollent system $\mathcal{A}^{(2)}$—characterizes the set of all numbers as an ordered Abelian group (cf. Section 57). Considering the content of the newly added axioms 3′, 4′, 8′, and 9′, we may describe the whole system as follows:

> *System $\mathcal{A}'$ expresses the fact that the set of all numbers is a densely and continuously ordered Abelian group with respect to the relation $<$ and the operation of addition, and it singles out a certain positive element 1 in that set.*

* * *

From the methodological point of view, System $\mathcal{A}'$ possesses several advantages. Formally considered, it appears to be the simplest of all known axiom systems upon which the theory of full arithmetic can be founded. With the exception of Axiom 1′, which—though not very easily—can be derived from the remaining axioms, all the other axioms of the system as well as the primitive terms occurring in these axioms are mutually independent. On the other hand, the didactic value of System $\mathcal{A}'$ as well as its appeal to intuition are far smaller; a basic drawback is that the simplicity of the foundations causes considerable complications in the subsequent development. Even constructing the definition of multiplication and deriving the basic laws for this operation are not easy tasks to carry through. Almost from the beginning the arguments have to involve the axiom of continuity in an essential way, and inferences which are based on this axiom are usually rather difficult for the beginner (for instance, without its help it would not be possible to prove the existence of the number $\frac{1}{2}$, that is, of a number y such that $y + y = 1$).

63 The second axiom system for the arithmetic of real numbers

For the reasons which were noted above, it is worthwhile to search for a different system of axioms upon which to base the construction of arithmetic. We can obtain such a system in the following way. As our point of departure we take System $\mathcal{A}^{(2)}$. Three additional primitive terms will be adopted: "*zero*", "*one*", and "*product*"; the first two will be replaced by the symbols "0" and "1", as usual, while instead of the expression "*the product of the numbers* (or, *of the factors*) x *and* y" (and instead of: "*the result of multiplying* x *and* y") we shall use the customary notation "$x \cdot y$". Moreover, we enlarge the system by thirteen new axioms; two of these are already known to us, namely the axiom of continuity and the law of performability for addition. We arrive in this way at SYSTEM $\mathcal{A}''$, containing the six primitive terms "**R**", "$<$", "$+$", "0", "$\cdot$", and "1", and consisting of the following twenty sentences:

AXIOM 1''. *If* $x \neq y$, *then* $x < y$ *or* $y < x$.

AXIOM 2''. *If* $x < y$, *then* $y \nless x$.

AXIOM 3''. *If* $x < y$ *and* $y < z$, *then* $x < z$.

AXIOM 4''. *If* S *and* T *are any sets of numbers satisfying the condition:*

> *for any* x *belonging to* S *and any* y *belonging to* T, *we have:* $x < y$,

then there exists a number z *for which the following condition holds:*

> *if* x *is any element of* S *and* y *any element of* T, *and if* $x \neq z$ *and* $y \neq z$, *then* $x < z$ *and* $z < y$.

AXIOM 5''. *For any numbers* x *and* y *there exists a number* z *such that* $z = x + y$ (in other words: *if* $x \in \mathbf{R}$ *and* $y \in \mathbf{R}$, *then* $x + y \in \mathbf{R}$).

AXIOM 6''. $x + y = y + x$.

AXIOM 7''. $x + (y + z) = (x + y) + z$.

AXIOM 8''. *For any numbers* x *and* y *there exists a number* z *such that* $x = y + z$.

AXIOM 9''. *If* $y < z$, *then* $x + y < x + z$.

AXIOM 10''. $0 \in \mathbf{R}$.

AXIOM 11''. $x + 0 = x$.

AXIOM 12''. *For any numbers* x *and* y *there exists a number* z *such that* $z = x \cdot y$ (in other words: *if* $x \in \mathbf{R}$ *and* $y \in \mathbf{R}$, *then* $x \cdot y \in \mathbf{R}$).

AXIOM 13''. $x \cdot y = y \cdot x$.

AXIOM 14''. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

AXIOM 15''. *For any numbers* x *and* y, *if* $y \neq 0$, *then there exists a number* z *such that* $x = y \cdot z$.

AXIOM 16''. *If* $0 < x$ *and* $y < z$, *then* $x \cdot y < x \cdot z$.

AXIOM 17''. $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

AXIOM 18″. $1 \in \mathbf{R}$.

AXIOM 19″. $x \cdot 1 = x$.

AXIOM 20″. $0 \neq 1$.

64 Closer characterization of the second axiom system; the concept of a field and that of an ordered field

In System $\mathcal{A}''$, as in System $\mathcal{A}'$, three groups of axioms may be distinguished. In Axioms 1″–4″, which form the first group, we have only the two primitive terms "**R**" and "$<$"; the second group, consisting of Axioms 5″–11″, contains two further symbols: the addition sign "+" and the symbol "0"; finally, the third group, which is made up of Axioms 12″–20″, involves primarily the multiplication sign "·" and the symbol "1".

All the axioms of the first two groups, with the exception of Axioms 10″ and 11″, are already known to us. Axioms 10″ and 11″ together state that 0 is a (right-hand) identity element for the operation of addition. In general, a given element i is said to be a RIGHT-HAND or a LEFT-HAND IDENTITY ELEMENT FOR THE OPERATION O IN THE CLASS K, if i belongs to K and if every element x of K satisfies the respective formula:

$$xOi = x \quad \text{or} \quad iOx = x.$$

If i is both a right-hand and a left-hand identity element, it is called simply an IDENTITY ELEMENT FOR THE OPERATION O IN THE CLASS K; evidently, in the case of a commutative operation O, every object which is a right-hand or a left-hand identity element is also an identity element.

* * *

The first three axioms of the third group, Axioms 12″–14″, we recognize as the LAW OF PERFORMABILITY and the COMMUTATIVE and the ASSOCIATIVE LAW for multiplication; they correspond precisely to Axioms 5″–7″. Axioms 15″ and 16″ are called the LAW OF RIGHT INVERTIBILITY for multiplication and the LAW OF MONOTONICITY for multiplication with respect to the relation *less than.* These axioms correspond to the laws of invertibility and monotonicity for addition, but not exactly. The difference lies in the fact that their hypotheses contain the restrictive conditions "$y \neq 0$" and "$0 < x$"; in disregard of their names, therefore, they do not permit us to assert simply that multiplication is invertible, or that it is monotonic with respect to the relation $<$ (in the sense of Sections 47 and 49).

Axiom 17″ establishes a fundamental connection between addition and multiplication; it is the so-called DISTRIBUTIVE LAW (strictly speaking, the LAW OF RIGHT DISTRIBUTIVITY) for multiplication with respect to addition. In general, the operation P is said to be RIGHT- or LEFT-DISTRIBUTIVE WITH RESPECT TO THE OPERATION O IN THE CLASS K if any three elements x, y, and z of this class satisfy the formula:

$$xP(yOz) = (xPy)O(xPz)$$
$$\text{or} \quad (xOy)Pz = (xPz)O(yPz),$$

respectively. If the operation P is commutative, the notions of right and left distributivity coincide, and we simply say that the operation P is DISTRIBUTIVE WITH RESPECT TO THE OPERATION O IN THE CLASS K.

The last three axioms concern the number 1. Axioms 18″ and 19″ together state that 1 is a right-hand identity element for the operation of multiplication. The content of Axiom 20″ does not call for any explanation; the role which is played by this axiom in the construction of arithmetic is greater than might be supposed at first, for without its help one cannot show that the set of all numbers is infinite.

* * *

In order to describe briefly the totality of properties which are attributed to addition and multiplication in Axioms 5″–8″, 12″–15″, and 17″, we say that these axioms characterize the set **R** as a FIELD (or, more precisely, as a COMMUTATIVE FIELD) WITH RESPECT TO THE OPERATIONS OF ADDITION AND MULTIPLICATION. If, in addition, the axioms of order 1″–3″ and the axioms of monotonicity 9″ and 16″ are taken into account, the set **R** is characterized as an ORDERED FIELD WITH RESPECT TO THE RELATION $<$ AND THE OPERATIONS OF ADDITION AND MULTIPLICATION. The reader will easily guess how the concept of a field and that of an ordered field are to be extended to arbitrary classes, operations, and relations.—If, finally, the continuity axiom 4″ and the axioms concerning the numbers 0 and 1 (Axioms 10″, 11″, 18″–20″) are taken into consideration, then the content of the whole axiom system $\mathcal{A}''$ may be described as follows:

> *System $\mathcal{A}''$ expresses the fact that the set of all numbers is a continuously ordered field with respect to the relation $<$ and the operations of addition and multiplication, and singles out two distinct elements 0 and 1 in that set, of which the first is the identity element for addition, and the second, the identity element for multiplication.*

65 Equipollence of the two axiom systems; methodological disadvantages and didactic advantages of the second system

The axiom systems $\mathcal{A}'$ and $\mathcal{A}''$ are equipollent (or rather, they become equipollent as soon as the first system is supplemented by the definitions of the symbol "0" and the multiplication sign "·", which can be formulated with the help of its primitive terms). However, the proof of this equipollence is not easy. It is true that the derivation of the axioms of the first system from those of the second is not especially difficult; but as far as the opposite task is concerned, we already indicated that both the definition of multiplication on the basis of the first system, and the subsequent proofs of the basic laws governing this operation, present considerable difficulties (while these laws appear as axioms in the second system).

In methodological respects System $\mathcal{A}'$ surpasses System $\mathcal{A}''$ considerably. The number of axioms in $\mathcal{A}''$ is more than twice as large. The axioms are not mutually independent; for instance, Axioms 5″ and 12″, that is, the laws of performability for addition and multiplication, are derivable from the remaining axioms, or, if these two axioms are retained, certain others such as Axioms 6″, 11″, and 14″ may be eliminated. Moreover, the primitive terms are not independent, for three of them can be defined in terms of the others, namely "<", "0", and "1" (*a possible definition of the symbol "0" was given in Section 53*); consequently the number of axioms can be reduced still further.

We see therefore that System $\mathcal{A}''$ admits important simplifications of various kinds; but as a result of such simplifications, the didactic advantages of the system (like its intuitive values) would be considerably diminished.† And these advantages are indeed great. On the basis of System $\mathcal{A}''$ one can develop without any difficulty the most important parts of the arithmetic of real numbers,—such as the theory of the fundamental relations among numbers, the theory of the four familiar arithmetical operations of addition, subtraction, multiplication, and division, the theory of linear equations, inequalities, and functions. The methods of inference which are needed for this purpose are of a very natural and quite elementary character, and in particular, the axiom of continuity does not enter at all at this stage; it plays an essential role only when we go over to the "higher" arithmetical operations of raising to a power, extracting roots, and taking logarithms,

†We may recall that several times the author referred to the circle of ideas, which can be described by words and phrases like "*expository*", "*didactic*" (or "*pedagogical*"), "*appealing to intuition*", "*heuristic*", even "*esthetic*". (See in particular the discussion of Section 39 regarding the choice of an axiom system, and also Sections 55, 57, and 62.) Indeed, attempts to refine the expository form were to him always an important part of scientific research.

and it is indispensable for the proof of existence of irrational numbers. No system of axioms and primitive terms appears to be known that might furnish a more advantageous basis for an elementary and, at the same time, strictly deductive construction of the arithmetic of real numbers.

Exercises

1. Show that the set of all positive numbers, the relation $<$, the operation of multiplication, and the number 2 form a model of System $\mathcal{A}'$, and that this system therefore possesses at least two different interpretations within arithmetic.

2. Which of the relations listed in exercise 6 of Chapter V are dense?

***3.** How can we express—in the symbolism of the algebra of relations—the fact that a given relation R is dense (in the universal class)? How can we express the fact that a certain relation S is both transitive and dense by means of one equation from the algebra of relations? (Cf. exercise *17 of Chapter V.)

4. Determine which of the following sets of numbers are densely ordered by the relation $<$:

(a) the set of all natural numbers,
(b) the set of all integers,
(c) the set of all rational numbers,
(d) the set of all positive numbers,
(e) the set of all numbers different from 0.

***5.** We assume System $\mathcal{A}'$ as the basis for this exercise; in order to prove the existence of the number $\frac{1}{2}$, that is, of a number z such that:

$$z + z = 1,$$

we can proceed as follows. Let K be the set of all numbers x such that:

$$x + x < 1,$$

and similarly, let L be the set of all numbers y such that:

$$1 < y + y.$$

We show first the set K precedes the set L. Applying now the axiom of continuity, we obtain a number z which separates the set K from the set L. Next it can be shown that the number z can belong neither to K (otherwise a number x in K greater than z would exist) nor to L. From this we may conclude that z is the number looked for, in other words, that

$$z + z = 1$$

holds.—Carry out in detail the proof which has just been sketched.

***6.** Generalizing the procedure of the preceding exercise, prove the following Theorem T on the basis of System $\mathcal{A}'$:

T. *For any number x there exists a number y such that*
$$x = y + y.$$

Compare the result obtained in this way with the discussion in Section 60.

***7.** In System $\mathcal{A}'$ replace Axiom 3′ by Theorem T of the previous exercise. Show that the system of sentences which is obtained in this way is equipollent to System $\mathcal{A}'$.

Hint: In order to derive Axiom 3′ from the modified system, substitute "$x+z$" for "x" in T; in view of the hypothesis of Axiom 3′, it can be shown easily that the resulting number y validates the conclusion of this axiom.

***8.** Use the method of proof by interpretation to show that, if Axiom 1′ is omitted, System $\mathcal{A}'$ becomes a system of mutually independent axioms.

***9.** Give a geometrical interpretation of the axiom systems $\mathcal{A}'$ and $\mathcal{A}''$, by way of an extension of exercise 2 of Chapter VIII.

***10.** Write all the axioms of Systems $\mathcal{A}'$ and $\mathcal{A}''$ in logical symbolism.

11. Do the operations of subtraction and division and the operations mentioned in exercise 3 of Chapter VIII have right-hand or left-hand identity elements, or identity elements, in the set of all numbers? Do the operations of addition and multiplication of sets of points possess identity elements in the class of all sets of points?

***12.** Show that any operation which is commutative in a class possesses at most one identity element in that class. Furthermore, as a generalization of the result which was obtained in exercise *24 of Chapter VIII, prove the following group-theoretical theorem:

If the class K is an Abelian group with respect to the operation O, then this operation has exactly one identity element in the class K.

13. Consider the five arithmetical operations of addition, subtraction, multiplication, division, and raising to a power. Formulate the sentences which assert that one of these operations is right-distributive, and sentences asserting that it is left-distributive, with respect to another operation (there are altogether 40 such sentences); determine which of them are true.

14. Solve the same problem as in the preceding exercise for the four operations A, B, G, and L which were introduced in exercise 3 of Chapter VIII. Moreover, show the following: if a given operation is performable in a certain set of numbers, then it is right- and left-distributive in that set with respect to the operations A and B.

15. Is the addition of classes distributive with respect to multiplication, and vice versa? (Cf. exercise 13 of Chapter IV.)

16. Which of the sets of numbers listed in exercise 4 are fields with respect to addition and multiplication, or ordered fields with respect to these operations and the relation $<$?

17. Show that the set consisting of the numbers 0 and 1 is a field with respect to the operation $\oplus$ defined in exercise 6 of Chapter VIII and multiplication.

18. Find two operations on the numbers 0, 1, and 2 of such a kind, that the set of these three numbers forms a field with respect to those two operations.

19. How can one define the symbol "1" with the help of the symbol of multiplication?

20. The following theorem can be derived from the axioms of System $\mathcal{A}''$:

$$\forall_x[0 < x \rightarrow \exists_y(x = y \cdot y)].$$

Supposing this theorem to have been already proved, derive from the axioms of System $\mathcal{A}''$ and with its help the following theorem:

$$\forall_{x,y}[x < y \leftrightarrow (x \neq y \wedge \exists_z[x + (z \cdot z) = y])].$$

Does the latter theorem justify the remark made in Section 65 concerning a possible reduction in the number of primitive terms of System $\mathcal{A}''$?

***21.** Prove Theorem T of exercise *6 on the basis of System $\mathcal{A}''$. Compare this proof with the one suggested in exercise *6 with reference to System $\mathcal{A}'$; which of the two proofs is more difficult and requires a greater knowledge of logical concepts?

Hint: In order to derive Theorem T from System $\mathcal{A}''$, apply Axiom 15″, with "y" and "z" replaced by "$1 + 1$" and "y" respectively (one first has to show, however, that $1 + 1$ is distinct from 0); in this way a number y is obtained, and one can show with the help of Axioms 13″, 17″, and 19″ that it satisfies the formula which is given in Theorem T.

***22.** Derive all the axioms of System $\mathcal{A}'$ from the axioms of System $\mathcal{A}''$.

Hint: In order to deduce Axiom 3′, assume Theorem T of exercise *6, which can be proved on the basis of System $\mathcal{A}''$ (cf. the previous exercise); from there on proceed in the same way as in exercise *7.

This page intentionally left blank

Index

The index contains the more important terms which occur in this book in a technical meaning, the expressions which take on a particular meaning or significance, a few non-technical entries, and the names of philosophers and scientists who are mentioned in the text (excepting the prefaces). The square brackets following a term (or an expression) enclose a term (or an expression) which is used in this book as a synonym of the former.

The index shows primarily the pages where the given term occurs, or where comments on the given expression can be found. On occasion the reference is to a passage which relates to the term or to the expression otherwise. E.g., one may have there a synonym of the term, or the corresponding symbol, or an illustrative example, or the occurrence of an expression but without a comment.

In case of a term which occurs extensively in the book, usually only selected pages are listed, and they are then shown in italics. On the other hand, if an entry is in italics, this indicates foreign words, or an antiquated term, or an expression which is used with a particular meaning, as remarked above.