International Series in Pure and Applied Mathematics

William Ted Martin and E. H. Spanier
Consulting Editors

Ahlfors Complex Analysis
Bellman Stability Theory of Differential Equations
Buck Advanced Calculus
Busacker and Saaty Finite Graphs and Networks
Cheney Introduction to Approximation Theory
Coddington and Levinson Theory of Ordinary Differential Equations
Cohn Conformal Mapping on Riemann Surfaces
Dennemeyer Introduction to Partial Differential Equations and Boundary Value Problems
Dettman Mathematical Methods in Physics and Engineering
Epstein Partial Differential Equations
Golomb and Shanks Elements of Ordinary Differential Equations
Graves The Theory of Functions of Real Variables
Greenspan Introduction to Partial Differential Equations
Griffin Elementary Theory of Numbers
Hamming Numerical Methods for Scientists and Engineers
Hildebrand Introduction to Numerical Analysis
Householder Principles of Numerical Analysis
Kalman, Falb, and Arbish Topics in Mathematical System Theory
Lass Elements of Pure and Applied Mathematics
Lass Vector and Tensor Analysis
Lepage Complex Variables and the Laplace Transform for Engineers
McCarty Topology: An Introduction with Applications to Topological Groups
Monk Introduction to Set Theory
Moore Elements of Linear Algebra and Matrix Theory
Mostow and Sampson Linear Algebra
Moursund and Duris Elementary Theory and Application of Numerical Analysis
Nef Linear Algebra
Nehari Conformal Mapping
Newell Vector Analysis
Ralston A First Course in Numerical Analysis
Ritger and Rose Differential Equations with Applications
Rosser Logic for Mathematicians
Rudin Principles of Mathematical Analysis
Saaty and Bram Nonlinear Mathematics
Simmons Introduction to Topology and Modern Analysis
Sneddon Elements of Partial Differential Equations
Sneddon Fourier Transforms
Stoll Linear Algebra and Matrix Theory
Struble Nonlinear Differential Equations
Weinstock Calculus of Variations
Weiss Algebraic Number Theory
Zemanian Distribution Theory and Transform Analysis

Introduction to Set Theory

J. Donald Monk
Professor of Mathematics
University of Colorado

McGraw-Hill Book Company
New York
St. Louis
San Francisco
London
Sydney
Toronto
Mexico
Panama

Introduction to Set Theory

Library of Congress Catalog Card Number 68-20056
42715
1234567890 MAMM 7654321069

To my mother

Preface

This book is intended to be a self-contained introduction to all the set theory needed by most mathematicians. It is self-contained in the sense that no prior knowledge of set theory is logically assumed; but the reader should have a working ability in the elements of the subject—the topics covered in Chap. 1. The usual elementary definitions and facts about sets, relations, functions, unions, and so on, are covered in Chap. 1. In Chap. 2 ordinal numbers are studied; the main topics are definitions by transfinite recursions and the rudiments of ordinal arithmetic. Chapter 3 is devoted to a brief study of the axiom of choice; several equivalent forms of this axiom are given, and some mathematical applications are discussed. The heart of the book is Chap. 4, on cardinal numbers. Besides elementary facts about addition, multiplication, and exponentiation of cardinals, some more advanced cardinal arithmetic is discussed; in particular, regular and singular cardinals are treated at some length.

The topics covered, although by no means exhausting abstract set theory, should suffice for the purposes of most working mathematicians. For more exhaustive treatments, see Bachmann 1967 or Sierpinski 1958 (see the Bibliography at the end of the book). Suggestions for further reading are given from time to time for those interested more deeply in set theory. Some sections of the book may be omitted without loss of continuity, since they cover specialized topics. This applies to Secs. 15, 23, and 24. The exercises, although not prerequisite for any textual material, are strongly advised in order to obtain a working knowledge of the material; there are not many and they are reasonably easy, so that the reader can try to work them all.

The approach to set theory here is axiomatic. Since set theory should form the logical basis for all mathematics, it seems clear to the author that in set theory, more than anywhere else in mathematics, one should strive for rigor. This accounts for the somewhat formal cast of this book. However, the author has strived not to be too pedantic. Logical symbolism is used, but only where it is essential, or where it seems to clarify a situation. Set theory should be based on formal logic, but here it is based on intuitive logic. Intuitive logic is expounded in the Introduction. For the purist who wants all the rigor attainable by present-day standards, the Appendix develops formal logic and indicates how to fit the set-theoretical development of the text into the formal framework of logic. Unlike almost all books at this level, we state all the set-theoretical axioms at the beginning, in Sec. 1. Hopefully this gives more clarity to the development than stringing the axioms throughout the book.

Precisely speaking, the axiomatic approach used is that of Kelley

and Morse, expounded in the appendix of Kelley 1955. It seems to the author that the Kelley, Morse system, or the closely related system of Gödel 1940, is used more often than any others by working mathematicians when any question of the foundation of set theory arises. It has the advantage of minimizing the necessary discussion of the symbolism of set theory. The axiom of choice is used freely in the book whenever it is needed or whenever its use shortens an exposition. We do, however, give many informal comments on important cases in which its use is essential or inessential, but no effort is made to indicate all cases where it may be eliminated. Transfinite recursion is used to establish some equivalents to the axiom of choice, shortening the usual proofs appreciably. Metamathematical results concerning the axioms are stated informally at various places in the book, but especially at the end of Sec. 1.

The book developed from several courses given by the author, and first thanks go to the students in those courses. I am also very much indebted to Stephen Comer, Gebhard Fuhrken, James S. Johnson, and Jack L. Hursch, Jr., for valuable comments on various versions of the book. In addition I also wish to express thanks to Mrs. Mae Jean Ruehlman for her expert typing. Although the material of the book is well known, I hope that the influence of my teacher, Alfred Tarski, is evident in its presentation here.

J. Donald Monk

Contents

Introduction to Set Theory

INTRODUCTION

Intuitive Logic

Before beginning the development of set theory itself, we want to clarify on an intuitive basis some words of common usage that are of a purely logical nature and are necessary in the development. Examples of such words are "and," "not," "there exists," and "equals." We should indicate why we feel that it is necessary to do this. In almost all fields of mathematics we proceed from the mathematics previously developed. Set theory, however, is intended to be at the very foundations of mathematics; aside from assuming a habituation with abstraction, it does not depend on any prior knowledge of mathematics. It does assume a substratum of ordinary logic common to a large number of people, not all of whom are mathematicians. Not much reflection is required to convince oneself that this substratum is not universal and not without its controversial points. Hence the purpose of this introduction: to try in a nonrigorous way to affix precise meanings to common logical usage. To do this in a rigorous way would require a formal, rigorous development of logic itself, prior to a discussion of set theory. Such a development

is given in the Appendix. Thus, by starting with the Appendix and then proceeding to Chap. 1, the reader can see a development of set theory that meets all modern standards of rigor. To a great extent, however, common sense subsumes logic, and its rigorous development is inappropriate to a book on set theory. We also sympathize with those who are impatient with any sign of pedanticism and wish to get right to the heart of the matter; for them, even the present introduction may be skipped (but see the Index of Notation at the end of the book).

Logic (technically, first-order logic) has two main parts, *sentential logic* and *quantifier logic*. The first is easier to explain and grasp than the second, so that we consider it first. In sentential logic we are concerned with sentences, their truth or falsity, and ways of combining or connecting sentences to produce new ones. A *sentence* is an expression about which it is reasonable to assert its truth or falsity. This "definition" looks all right but on closer analysis it not very good; in Chap. 1 and in the Appendix a better definition is given, restricted to a well-defined artificial language but correspondingly further from the intuition and to this extent unsuitable. Let us take some examples. "$2 + 2 = 4$" is a sentence, in fact a true sentence. "π is rational" is a false sentence; but it is not a task of logic to decide whether or not "π is rational" is true or false. "There are infinitely many pairs of prime numbers p, q such that $q = p + 2$" is a sentence (the twin-prime conjecture), but it is unknown at this time whether it is true or false. Now consider the expression

(1) The set of all sets not members of themselves is a member of itself.

Offhand, it seems quite reasonable to consider this expression a sentence. But if it is true, then so is the sentence

(2) The set of all sets not members of themselves is not a member of itself,

which expresses the opposite of (1); and if (2) is true, then so is its opposite (1). Thus it turns out not to be reasonable to assert the truth or falsity of (1); by our definition (1) is not to be considered a sentence. Clearly there may be cases in which it is very difficult under our definition to determine whether or not a given expression is a sentence. This is a major defect of the definition.

The paradoxical nature of the inference from (1) to (2) and from (2) to (1) is known as *Russell's paradox*. Because of its set-theoretical form it has played a large role in the historical development of set theory. In fact, it is mainly because of this paradox that it will be necessary in Chap. 1 to redefine "sentence" in order not to set up an obviously contradictory system.

At least there are some expressions, like the first three of the above examples, which are incontrovertibly sentences under our definition. So let us proceed to a description of some purely logical ways of connecting sentences so as to form new ones. The most useful connectives are the following.

OR

Writing "or" between two sentences, we form a sentence which is true if at least one of the two sentences is true and false if both are false. This usage deviates somewhat from the ordinary use of "or." We do not mean "either . . . or . . . ," but "either . . . or . . . or both." Furthermore, by our agreement the truth or falsity of the component sentences is all that matters in determining the truth or falsity of the compound sentence; there does not have to be any connection between the components. For example, the following sentences must be taken as true, odd as they may look:

$1 > 0$ or π is not a real number.
$4^2 = 16$ or there is an x such that $x^2 = -1$.
$\int_0^1 \sin x\, dx = 7$ or $\sqrt{2}$ is irrational.
$2 = 2$ or $2 \neq 2$.

Of course if both components are false, the compound sentence is false, for example:

$2 \neq 2$ or $2 \neq 2$.
$4^2 = 15$ or π is rational.

We use the abbreviation $\vee$ for "or."

In order to prove a sentence of the form $\varphi \vee \psi$, where φ and ψ are component sentences, we may assume that φ is false and give an argument that ψ is true. For if φ is true, then $\varphi \vee \psi$ is true, and if φ is false, the argument shows that ψ is true and hence $\varphi \vee \psi$ is true. Symmetrically we may assume that ψ is false and prove that φ is true.

Example If p is a prime and p divides $a \cdot b$, then p divides a or p divides b. To prove this, first assume that p is a prime and p divides $a \cdot b$. Further, assume that p does not divide a. Then $(p,a) = 1$, so that there exist integers s and t such that $1 = sp + ta$. [Here we use (p,a) for the *greatest common divisor* of p and a.] Multiplying by b, we get $b = spb + tab$. Now p divides spb and p divides tab, so that p divides b. The proof is complete.

AND

Writing "and" between two sentences makes a new sentence that is true if both sentences are true and is false otherwise. Only the truth values

matter, nothing else. Thus

$$1 > 0 \text{ and } 2 + 2 = 4$$

is true, but the following sentences are all false:

$0 > 1$ and $\sqrt{3}$ is irrational.
$i^2 = -1$ and π is rational.
$0 > 1$ and $\sqrt{3}$ is rational.

We use the abbreviation $\wedge$ for "and."

IMPLIES

A sentence φ implies a sentence ψ if either φ is false or ψ is true. Thus, again, we are interested only in truth values. In mathematics this use of the word "implies" has become more or less standard, although outside of mathematics other meanings are probably more prevalent. The main difference occurs in sentences like

$2 + 2 = 5$ implies π is rational,

or

$2 + 2 = 5$ implies π is irrational.

Both of these are true under our specification, since in both cases the sentence to the left of "implies," namely, "$2 + 2 = 5$," is false. In such cases we say that the implication holds *vacuously*.

Some further examples of our use of "implies" are

$i^2 = -1$ implies $2 = 2$.	true
$i^2 = -1$ implies $2 = 3$.	false
$i^2 = 0$ implies $2 = 2$.	true
$i^2 = 0$ implies $2 = 3$.	true

We use $\Rightarrow$ as an abbreviation for "implies." In a sentence $\varphi \Rightarrow \psi$ we call φ the *hypothesis* of the implication and ψ the *conclusion*. In order to prove a sentence of the form $\varphi \Rightarrow \psi$, we frequently take φ as an additional assumption (along with whatever other mathematical assumptions we are using at the time), argue awhile, come up with the conclusion ψ, and then "discharge" the assumption φ, that is, state that $\varphi \Rightarrow \psi$ has been proved, from which point we no longer have φ as an assumption.

Example $a > 0$ implies that a has a square root. To prove this, assume that $a > 0$. Let b be the largest real number such that $b^2 \leq a$. By a supplementary argument, it is seen that $b^2 = a$; that is, a has a square root. Thus we have shown that $a > 0$ implies that a has a square root.

NOT

This "connective" converts one sentence into another, namely, a true sentence into a false one and a false one into a true one. Thus "every positive integer is a sum of four squares" is true, and so "not (every positive integer is a sum of four squares)" is false, and "not (not (every positive integer is a sum of four squares))" is true. In these examples we transgress the common rules of grammar, but the meaning should be clear. In order to simplify the rules for combining sentences, it is frequently desirable to abuse grammar in this way, and this practice is continued throughout the book.

We use $\neg$ to abbreviate "not"; the sentence $\neg\varphi$ is called the *negation* of φ.

Arguments involving "not" are frequently of an indirect nature, in which we do not go directly from assumptions to a conclusion but infer that the assumptions imply the conclusion by means of a logical trick. We now describe two important kinds of indirect arguments.

CONTRAPOSITION

If we want to prove a sentence of the form $\varphi \Rightarrow \psi$, we may, instead of the direct method described above in the discussion of "implies," prove the sentence $\neg\psi \Rightarrow \neg\varphi$. For if we have done this and if φ is true, then $\neg\varphi$ is false, and hence, $\neg\psi \Rightarrow \neg\varphi$ being true, it cannot be the case that $\neg\psi$ is true; that is, ψ is true. The sentence $\neg\psi \Rightarrow \neg\varphi$ is called the *contrapositive* of the sentence $\varphi \Rightarrow \psi$.

Example Let I be the ring of integers, and for any n in I, $n > 1$, let (n) be the principal ideal generated by n. For any a in I, let $[a]$ be the equivalence class of a with respect to the ideal (n). Then $I/(n)$ is an integral domain implies n is a prime. To prove this, assume that n is not a prime. Then there exist a, b in I such that $0 < a < n$, $0 < b < n$, and $n = a \cdot b$. Thus $[a] \neq 0$, $[b] \neq 0$, but $[a]\cdot[b] = 0$. Hence $I/(n)$ is not an integral domain. Therefore, $I/(n)$ is an integral domain implies n is a prime.

REDUCTIO AD ABSURDUM, OR ARGUMENT BY CONTRADICTION

To prove a sentence φ, it is enough to assume $\neg\varphi$ and then prove some statement $\neg\psi$, where ψ is known to be true, which amounts to proving a contradiction $\psi \wedge \neg\psi$. For then the argument shows that φ could not be false after all.

Example If A is a set of positive integers such that 1 is in A, and $x + 1$ is in A whenever x is in A, then every positive integer is in A. To prove this statement, we assume that every nonempty set of positive integers has a least element. To argue by contradiction, we assume that there is a positive integer not in A. Then the set B of all positive integers not in A has a least element m. Because 1 is in A, we have $m > 1$. But then $m - 1$ is a positive integer, so that, by the minimality property of m, $m - 1$ is in A. But by the assump-

tion on A, this implies that $m = (m - 1) + 1$ is in A. Having already noted that m is not in A, this is a contradiction. Hence, after all, every positive integer is in A.

IF AND ONLY IF

We say that a sentence of the form "φ if and only if ψ" is true only in case both φ and ψ are true, or both φ and ψ are false. Thus the following two sentences are true:

$1 > 0$ if and only if $\pi^2 < 20$.
$1 = 0$ if and only if $\pi^2 < 0$.

On the other hand, these sentences are false:

$3^2 + 4^2 = 5^2$ if and only if π is rational.
$4^2 + 5^2 = 6^2$ if and only if π is irrational.

We use the symbol iff or $\Leftrightarrow$ in place of "if and only if." To prove a statement $\varphi \Leftrightarrow \psi$, we usually prove $\varphi \Rightarrow \psi$ and then prove $\psi \Rightarrow \varphi$. For, having done this, the first proof excludes the possibility that φ is true and ψ false, and the second excludes the possibility that ψ is true and φ false. Hence either both are true or both are false; that is, $\varphi \Leftrightarrow \psi$ is true.

Example $I/(n)$ is an integral domain iff n is a prime. For, as in the example above in the discussion of contraposition, we have ($I/(n)$ an integral domain) $\Rightarrow$ (n is a prime). Second, suppose that n is a prime. If $[k]\cdot[l] = 0$, then n divides $k\cdot l$ and hence n divides either k or l, from which it follows that $[k] = 0$ or $[l] = 0$. Thus $I/(n)$ is an integral domain. Hence (n is a prime) $\Rightarrow$ ($I/(n)$ is an integral domain), and the proof is complete.

Another method for proving $\varphi \Leftrightarrow \psi$ is to prove that $\varphi \Rightarrow \psi$ and $\neg\varphi \Rightarrow \neg\psi$; this amounts to the same method just described, because $\neg\varphi \Rightarrow \neg\psi$ yields $\psi \Rightarrow \varphi$, by contraposition.

These five connectives—"or," "and," "implies," "not," and "if and only if"—are sufficient to express conveniently any sentential combinations that we consider in this book. The rules for the truth values of compound sentences formed by using these connectives, which we again emphasize depend only on the truth value of the components, are summarized in the following tables:

φ	$\neg\varphi$
True	False
False	True

φ	ψ	$\varphi \vee \psi$	$\varphi \wedge \psi$	$\varphi \Rightarrow \psi$	$\varphi \Leftrightarrow \psi$
True	True	True	True	True	True
True	False	True	False	False	False
False	True	True	False	True	False
False	False	False	False	True	True

There are many English phrases that we consider synonomous with these five connectives, and there are certain grammatical variations of the connection process that are applied. Thus "φ is necessary for ψ" is considered synonymous with $\psi \Rightarrow \varphi$; "φ is sufficient for ψ" is synonymous with $\varphi \Rightarrow \psi$; "φ is necessary and sufficient for ψ" means $\varphi \Leftrightarrow \psi$; "$\varphi$ is equivalent to ψ" means $\varphi \Leftrightarrow \psi$; "$\varphi$ if ψ" means $\psi \Rightarrow \varphi$; "φ only if ψ" means $\varphi \Rightarrow \psi$; "φ whenever ψ" means $\psi \Rightarrow \varphi$; "if φ then ψ" means $\varphi \Rightarrow \psi$; "φ just in case ψ" means $\varphi \Leftrightarrow \psi$. There are, of course, many other phrases that may be used in place of the five we singled out. Furthermore, we may, for the sake of grammatical usage, modify the formal use of the connectives. For example, it is nicer to write

π is not a rational number

than to write

not (π is a rational number)

However, as we indicated in the discussion of "not" above, we write ungrammatical sentences where this would seem to aid clarity.

To conclude the discussion of sentential logic, we mention that there are some sentences that are true on logical grounds only. Rather than trying to explain what we mean by this assertion, we give a number of examples. The following sentences are true no matter what truth values the component sentences φ, ψ, χ have:

$\varphi \vee \neg\varphi$. law of excluded middle

$\neg(\varphi \wedge \neg\varphi)$.

$(\varphi \Rightarrow \psi) \Leftrightarrow (\neg\psi \Rightarrow \neg\varphi)$.

$\neg(\varphi \wedge \psi) \Leftrightarrow (\neg\varphi \vee \neg\psi)$.
$\neg(\varphi \vee \psi) \Leftrightarrow (\neg\varphi \wedge \neg\psi)$. } De Morgan's laws

$[\varphi \vee (\psi \wedge \chi)] \Leftrightarrow [(\varphi \vee \psi) \wedge (\varphi \vee \chi)]$.
$[\varphi \wedge (\psi \vee \chi)] \Leftrightarrow [(\varphi \wedge \psi) \vee (\varphi \wedge \chi)]$. } distributive laws

We now turn to the discussion of quantifier logic, which revolves around the phrases "there is" and "for every." In mathematics these phrases are almost universally used in conjunction with variables. A *variable* is simply a letter of our alphabet (or perhaps of some other alphabet, like the Greek or German alphabet) used together with these phrases. Let us first consider the phrase "for every," which is called the *universal quantifier*. As an example, consider the sentence

(3) For every x, $x < 0$ or $x = 0$ or $0 < x$.

All the occurrences of x in this sentence are bound up with the phrase "for every." Every variable has a range of values, which in this case consists of (say) all integers. Having asserted the sentence (3), we may

on purely logical grounds, also assert the sentence obtained from that following the comma in (3) by replacing all three occurrences of x by the name of some integer. Thus the following sentences follow logically from (3):

$$\begin{array}{rcrcrcl} -5 < 0 & \text{or} & -5 = 0 & \text{or} & 0 < -5. \\ 0 < 0 & \text{or} & 0 = 0 & \text{or} & 0 < 0. \\ 1 < 0 & \text{or} & 1 = 0 & \text{or} & 0 < 1. \end{array}$$

As another example take the sentence

(4) For every positive real number a, a has a square root.

Here the range of values of a consists of all positive real numbers. As special cases of assertion (4) we have

1 has a square root.
7 has a square root.
$\sqrt{2}$ has a square root.

Analogously with the case of sentential connectives there are other phrases that in mathematics are taken to be synonymous with "for every," for example, the phrases "for any," "for all," and "for each." We also allow grammatical variations in the process of applying "for every." Thus instead of (4) we might write

Every positive real number has a square root,

in which the variable is understood but not used.

Thus a variable has no meaning standing alone, but only in connection with the phrases "for every" or "there is," or synonymous phrases. In mathematics constants are also used profusely. A *constant* is simply a proper name of some kind. We may distinguish between *permanent* and *temporary constants*. An example of a permanent constant is ω, which is used throughout this book to denote the set of all nonnegative integers. Other examples are π, 0, $\sqrt{2}$. Temporary constants are constants used only for a short discussion, as in the proof of a theorem. Frequently the same letters are used for temporary constants in one place as for variables in a different context.

We use the abbreviation $\forall$ for "for every." To illustrate the use of variables and constants, consider the task of proving a sentence of the form $\forall x \varphi(x)$, where $\varphi(x)$ is some expression; we write $\varphi(x)$ to emphasize that this expression is likely to involve x, although it is not necessary that it do so. If the range of x is a set X, we frequently prove $\forall x \varphi(x)$ by taking an arbitrary element a in X (thus using a as a temporary constant) and then proving $\varphi(a)$. Nothing is wrong with saying "Let x be an

arbitrary element of X" and then proving $\varphi(x)$, thus using x in place of a as a temporary constant. The two uses of x, as a variable in $\forall x\varphi(x)$, and as a constant in the argument, should be kept distinct.

Example For every rational number x, $x^2 \neq 2$. To prove this, let a be an arbitrary rational number, say $a = r/s$, where $(r,s) = 1$. [Again we use (r,s) for the greatest common divisor of r and s.] Suppose, to argue by contradiction, that $a^2 = 2$. Thus $r^2 = 2s^2$. Let $r = 2^t r'$ with $(2,r') = 1$ and let $s = 2^u s'$ with $(2,s') = 1$. Then we easily infer that $2t = 2u + 1$, which is impossible. Hence $a^2 \neq 2$ after all. Because a is arbitrary we have shown that for every rational number x, $x^2 \neq 2$.

As in the case of implications, a sentence of the form $\forall x\varphi(x)$ may be true *vacuously*. This is the case with the sentence

Every rational square root of 2 is negative,

which may be reformulated as "$\forall x(x$ is a rational square root of $2 \Rightarrow x < 0)$"; the hypothesis of the implication is false, and hence the implication itself is true, for any possible value of x.

In stating theorems or axioms, frequently the universal quantifier is omitted but is understood to be present. Thus in writing

$$x + y = y + x$$

as a theorem, the sentence

$$\forall x\ \forall y(x + y = y + x)$$

is understood.

Now let us discuss the *existential quantifier* "there is," for which we use $\exists$ as an abbreviation. Synonymous phrases or grammatical variants are "there are," "there exist," "there exists," "for some." We use the existential quantifier in the sense "there is at least one." In order to prove a sentence of the form $\exists x\varphi(x)$, one usually constructs, in some sense, an object a such that $\varphi(a)$. It is possible, however to give a nonconstructive proof of such a sentence, say by assuming $\neg\exists x\varphi(x)$ and deriving a contradiction. We give an example of each procedure.

Example There is an x such that $x^3 - 7x^2 + 11x - 5 = 0$. Indeed, $5^3 - 7\cdot 5^2 + 11\cdot 5 - 5 = 125 - 175 + 55 - 5 = 0$. Thus we may take $x = 5$.

Example Every nonempty set of positive integers has a least element (we assume the complete induction principle for positive integers); compare with the example illustrating argument by contradiction, page 5. Let A be a nonempty set of positive integers. To argue by contradiction, suppose that A does not have a least element. Let B be the set of all positive integers not in A. Then 1 is in B, for if 1 were in A, it would be the least element of A. If

$\forall x(x < y \Rightarrow x$ in $B)$, then y is in B, for if y were in A, it would be the least element of A. Hence, by the principle of complete induction, every positive integer is in B, and so A is empty, which contradicts the assumption above. Hence, after all, A has a least element that is, $\exists x(x$ is in A and $x \leq y$ for all y in $A)$.

In arguments where a sentence of the form $\exists x\varphi(x)$ occurs as an assumption, one frequently introduces a temporary constant, like a, with the added assumption that $\varphi(a)$ holds. After drawing a conclusion that a sentence ψ, not involving a, then holds, one may logically drop the added assumption $\varphi(a)$ and state that ψ has been derived from the original assumptions [which, as we said, include $\exists x\varphi(x)$]. In introducing the assumption $\varphi(a)$, we may use terminology such as "choose a such that $\varphi(a)$" or "let a be such that $\varphi(a)$."

Example Let G be a group with identity e. Suppose that H is a nonempty subset of G such that $x \cdot y^{-1}$ is in H for all x, y in H. Then e is in H. To prove this, we assume that $\exists x(x$ is in $H)$, and we choose a in H. Then, by the assumption of the theorem, $a \cdot a^{-1}$ is in H. But $a \cdot a^{-1} = e$, so that e is in H.

Having the quantifiers available, we can add to the list of sentences that are true on logical grounds only. All these sentences are valid no matter what the expressions $\varphi(x)$, $\psi(x)$, $\chi(x,y)$ are:

$$\forall x\varphi(x) \Rightarrow \exists x\varphi(x).$$
$$\exists x\, \forall y\chi(x,y) \Rightarrow \forall y\, \exists x\chi(x,y).$$
$$\forall x[\varphi(x) \wedge \psi(x)] \Leftrightarrow \forall x\varphi(x) \wedge \forall x\psi(x).$$
$$\exists x[\varphi(x) \vee \psi(x)] \Leftrightarrow \exists x\varphi(x) \vee \exists x\psi(x).$$
$$\forall x\varphi(x) \Leftrightarrow \neg\exists x \neg\varphi(x).$$
$$\exists x\varphi(x) \Leftrightarrow \neg\forall x \neg\varphi(x).$$
$$\neg\forall x\varphi(x) \Leftrightarrow \exists x \neg\varphi(x).$$
$$\neg\exists x\varphi(x) \Leftrightarrow \forall x \neg\varphi(x).$$

Note, with regard to the first of these sentences, that a variable is always assumed to have a nonempty range of values. Hence if the sentence $\forall x\varphi(x)$ holds, then for a particular element a of the range of values of x, $\varphi(a)$ holds, and hence $\exists x\varphi(x)$ holds.

We conclude this introduction with a brief mention of one last logical notion, that of equality. The sentence $a = b$ expresses the assertion that a denotes the same thing as b. From this "definition" we see that equality possesses the following properties:

$$a = a.$$
$$a = b \Rightarrow b = a.$$
$$a = b \wedge b = c \Rightarrow a = c.$$

Equals may be substituted for equals.

Some authors assume only these properties of equality (that is, that = is merely a relation, between possibly distinct objects, satisfying these conditions), but our "definition" seems intuitively more satisfactory.

Remark We wish to reemphasize that the treatment of logic in this introduction has not been rigorous. For the purposes of exposition we have been more dogmatic than is justifiable. For good elementary treatments of logic, in which the pitfalls of the intuitive notions are contrasted to the rigorous approach, see Tarski 1965 and Mates 1965.[1] As previously mentioned, a rigorous development of logic can be found in the Appendix. A more advanced treatment of logic is found in Mendelson 1964.

All these treatments have to do with classical two-valued logic. The reader should be aware that, even with regard to basic logical facts, there are some alternative approaches. For example, the logical truth $\varphi \vee \neg\varphi$ has been questioned as a reasonable principle to use in mathematics; the *intuitionism* of Brouwer holds this and other similar principles in doubt, basically because in intuitionism *constructions* and not *proofs* are taken as fundamental. For an account of intuitionism see Heyting 1966. Many-valued logic is discussed on a technical level in Rosser, Turquette 1952.

[1] See the Bibliography at the end of the book.

1
Elementary Set Theory

In this chapter we give the axioms for set theory and develop its most elementary part. Our treatment is definitely not exhaustive; only the facts commonly used by working mathematicians are given, and most of these facts are so simple that in later chapters of this book we generally will use them without reference to this chapter.

We will prove only representative parts of the results we state; the reader should check for himself the validity of the other parts.

Since the discovery of various paradoxes at the turn of this century it has been recognized that some kind of axiomatic approach to set theory is necessary. Russell's paradox—the set of all sets not members of themselves both is and is not a member of itself—shows that we cannot define sets precisely as we wish. Another famous paradox in set theory is that of Burali-Forti: The set A of all ordinal numbers is a well-ordered set, whose order type should be an ordinal number, hence a member of A; but then the order type of A is smaller than the order type of A. Other paradoxes are easily derived, using irrefutable set-theoretical arguments, from the assumptions that the set of all sets exists; the set of all cardinal

numbers exists; the set of all supersets of a given set exists, etc. All of these paradoxes have a common feature: the existence is asserted of sets that are very "big." To precisely proscribe consideration of such "big" sets seems to require the formation of an axiomatic system.

Zermelo first gave a workable set of axioms for the theory of sets, and Fraenkel added a further axiom (the axiom of substitution) which made the axioms strong enough for almost all mathematical purposes. The system was simplified by Skolem, and variations were made by von Neumann, Bernays, Gödel, and A. P. Morse, leading to the system developed in this book.

1 THE AXIOMS

In this section we give all the axioms. Defined notions are introduced only to the extent that axioms may be conveniently formulated; in later sections these and other defined notions are discussed fully. Although the presentation of all of the axioms at once may make the whole subject seems rather formidable, after working with the notions in the succeeding sections we hope that they will seem natural.

In any axiomatic development one starts from undefined notions and axioms, although they may not be so called and the development may be within the scope of a larger development. Geometry is frequently developed explicitly in terms of undefined notions and axioms. However, group theory, for example, is usually developed within a larger framework, in which there are sets, functions, etc., but it may still be considered an axiomatic development, with the group elements and group operations as undefined notions, and as axioms the usual group axioms. In set theory we take the explicit approach: Set theory is not a part of a larger development, although it is based upon elementary logic.

In addition to undefined notions and axioms, it is convenient and, as a practical matter, essential to introduce various definitions in developing set theory. In definitions we introduce new symbols that can in principle always be eliminated in favor of the undefined symbols (see the Appendix).

We begin by giving the undefined notions.

Definition 1.1 *The primitive notions are those of a* ***class*** *and of* ***membership.*** *Capital italic letters* A, B, . . . , X, Y, Z *stand for classes. The membership relation is denoted by* $\in$ *and may or may not hold between two classes.* $\notin$ *stands for the negation of the membership relation; thus* $A \notin B$ *means* $\neg(A \in B)$.

As synonymous with the word "class" we will take "collection," "family"

(especially when thinking of a class of classes rather than a class of objects), "aggregate," but not "set," for which we reserve a special meaning described in Definition 1.3. $A \in B$ is read "*A is a member of B*" or "*A is an element of B*," but not "*A* contained in *B*" (see Definition 1.10). Since we always assume that variables range over some nonvoid range, we also tacitly assume that classes exist.

Axiom 1.2 (*Extensionality axiom*) $\forall A \forall B[\forall C(C \in A \Leftrightarrow C \in B) \Rightarrow A = B]$.

This axiom expresses the assertion that two classes with the same members are equal. The classes A and B may be defined in entirely different ways, for example,

A = set of all nonnegative integers,
B = set of all integers that can be written as a sum of four squares,

but if they have the same members, they are the same class (in this example $A = B$ by a well-known theorem of Lagrange). Note that the axiom of extensionality implies that there is at most one class with no elements. Since we are going to allow variables to range over classes only, we thus rule out of consideration "objects" or "Urelemente." The reasons for doing this are that it is sufficient in mathematics to consider everything a class and that it complicates the development considerably to admit objects (see Suppes 1960). The members of classes must also, then, be classes; the members of the members are classes, and so forth.

Examining Russell's paradox, we see that the class considered there, the class of all classes not members of themselves, is very big. The standard way of getting around this paradox in axiomatic set theory is to refuse to admit big classes, or at least not to allow big classes the same privileges as small ones. In many developments of set theory, for example Halmos 1960, the first alternative is followed. Big classes are not admitted. We elect the second alternative, which is more often used in research articles in mathematics, and we will distinguish between two kinds of classes. We define a set as a little class and a proper class as a big class.

Definition 1.3 *A is a* ***set*** *iff there is a B such that* $A \in B$. *A is a* ***proper class*** *iff A is not a set. Lowercase italic letters a, b, c, . . . , x, y, z are used for sets unless otherwise stated.*

Thus a set is a class small enough to be a member of some other class. Proper classes are too big for this. Directly from Definition 1.3 we have the following.

Corollary 1.4 $\exists B(a \in B)$.

Note that here we tacitly assume a universal quantifier $\forall a$. Corollary 1.4 is read "for every set a, there is a class B such that $a \in B$" (see the Introduction).

As a consequence of the axiom of extensionality we have the following.

Corollary 1.5 $\forall x(x \in A \Leftrightarrow x \in B) \Rightarrow A = B$.

Proof Assume $\forall x(x \in A \Leftrightarrow x \in B)$. Let C be an arbitrary class. If $C \in A$, then C is a set, and hence by the assumption, $C \in B$. Similarly $C \in B \Rightarrow C \in A$. Since C is arbitrary, it follows that $\forall C(C \in A \Leftrightarrow C \in B)$. Thus by the axiom of extensionality, $A = B$.

We now wish to describe an axiom that allows us to define a class of objects having a specified property. We wish to exclude Russell's paradox and thus the class of all X such that $X \notin X$. There is a natural way to make such an exclusion: We have defined a set as a class capable of membership in another class, and we admit only classes whose members are already known to be sets. Thus we may consider the class A of all sets X such that $X \notin X$. Then the argument giving Russell's paradox yields only "A is a proper class" (see the argument following Eq. (1) below), which is certainly not a contradiction.

To give the axiom in a rigorous form, we have to define what we mean by "specified property"; we prefer the terminology "set-theoretical formula."

Definition 1.6 *The expressions*

$$A = A,\ A = B,\ A = C,\ \ldots,\ B = A,\ B = B,\ B = C,\ \ldots,$$
$$C = A,\ C = B,\ C = C,\ \ldots$$

are all ***set-theoretical formulas,*** *as are*

$$A \in A,\ A \in B,\ A \in C,\ \ldots,\ B \in A,\ B \in B,\ B \in C,\ \ldots,$$
$$C \in A,\ C \in B,\ C \in C,\ \ldots.$$

If φ and ψ are set-theoretical formulas, so are $\neg\varphi$, $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, $(\varphi \Rightarrow \psi)$, $(\varphi \Leftrightarrow \psi)$, $\exists A\varphi$, $\exists B\varphi$, $\ldots$, $\forall A\varphi$, $\forall B\varphi$, $\ldots$. Set-theoretical formulas can only be obtained by (finitely many) applications of the processes just mentioned.

The following, then, are examples of set-theoretical formulas:

$$\neg X \in X$$
$$\exists X(X \in Y \wedge X \in Z)$$
$$\exists X X \in Y \wedge \exists Y Y \in Z$$

Note that the first part of 1.6 can be stated more concisely as follows: if α and β are capital Roman letters, then $\alpha = \beta$ and $\alpha \in \beta$ are set-theoretical formulas. Note that some expressions we have already used are not set-theoretical formulas; for example,

$$\forall a \exists B(a \in B) \qquad \text{Corollary 1.4}$$

is not, because the lowercase italic letter a occurs in it. However, 1.4 is equivalent, on the basis of Definition 1.3, to

$$\forall A[\exists C(A \in C) \Rightarrow \exists B(A \in B)],$$

which is a set-theoretical formula. In fact this is a fundamental property of definitions already mentioned: Defined expressions can always be eliminated in favor of primitive ones (see the Appendix). We will frequently make tacit use of this obvious property of definitions, and we will, treat $\forall a\, \exists B(a \in B)$, for example, as a set-theoretical formula. We also use various signs of aggregation freely in order to prevent ambiguity, although Definition 1.6 allows their use in restricted cases only.

Axioms 1.7 (*Class-building axioms*) *If $\varphi(X)$ is a set-theoretical formula not involving the letter A, then the following is an axiom:*

$$\exists A\ \forall X[X \in A \Leftrightarrow X \text{ is a set} \wedge \varphi(X)].$$

Similarly, if $\varphi(X)$ does not involve B, then the following is an axiom:

$$\exists B\ \forall X[X \in B \Leftrightarrow X \text{ is a set} \wedge \varphi(X)],$$

and so on for other letters. Letters other than X may also be used.

Here are some examples of class-building axioms. Letting $\varphi(X)$ be the expression $\neg(X \in X)$, we get, as an axiom,

$$(1) \quad \exists A\ \forall X[X \in A \Leftrightarrow X \text{ is a set} \wedge \neg(X \in X)].$$

Thus the class asserted to exist in Eq. (1) is the class of all sets not members of themselves. Let us try to reproduce Russell's paradox. If $A \in A$, then A is a set and $A \notin A$. Therefore $A \notin A$. Hence A is not a set, or $A \in A$. Knowing that $A \notin A$, we conclude that A is not a set. No contradiction is involved. Next let $\varphi(X)$ be the expression $X = X$. Then an axiom under 1.7 is

$$\exists B\ \forall X(X \in B \Leftrightarrow X \text{ is a set} \wedge X = X),$$

a logical consequence of which is

$$\exists B\ \forall X(X \in B \Leftrightarrow X \text{ is a set}).$$

Thus B is the class of all sets. With $\varphi(X) = \neg(X = X)$ we get

$$\exists A \, \forall X[X \in A \Leftrightarrow X \text{ is a set and } \neg(X = X)],$$

and hence A has no elements. With $\varphi(X) = X \in A \vee X \in B$ we get

$$\exists C \, \forall X[X \in C \Leftrightarrow X \text{ is a set and } (X \in A \vee X \in B)],$$

so that C is the union of the classes A and B.

The class asserted to exist in 1.7 is unique:

Corollary 1.8 *If $\varphi(x)$ is a set-theoretical formula not involving either of the letters A or B and if $\forall X(X \in A \Leftrightarrow X$ is a set $\wedge \varphi(X))$ and $\forall X(X \in B \Leftrightarrow X$ is a set $\wedge \varphi(X))$, then $A = B$. Similarly for formulas not involving either A or C, B or D, etc.*

Proof Under the assumptions of the theorem it follows on purely logical grounds that $\forall X(X \in A \Leftrightarrow X \in B)$. Hence $A = B$ by the extensionality axiom.

Definition 1.9 *For any set-theoretical formula $\varphi(X)$ not involving A, let $\{X : \varphi(X)\}$ be the unique class A such that $\forall X(X \in A \Leftrightarrow X$ is a set $\wedge \varphi(X))$. Similarly if $\varphi(X)$ does not involve B, C, and so on. This definition is justified by Corollary 1.8. Again, letters other than X may be used, and even lowercase letters (see 1.3).*

The symbolism introduced in Definition 1.9 is very convenient in practice. We may read $\{X : \varphi(X)\}$ as "the class of all sets X such that $\varphi(X)$." The classes given in the examples following 1.7 are, in this notation,

$$\{X : X \notin X\}, \quad \{X : X = X\}, \quad \{X : X \neq X\}, \quad \{X : X \in A \vee X \in B\}.$$

The entire force of the class-building axioms is embodied in this symbolism. In what follows, the class-building axioms will always be used simply by defining classes equal to $\{X : \varphi(X)\}$ for some $\varphi(X)$. Furthermore, it will not be necessary to discuss set-theoretical formulas generally any more, and we will use only concrete formulas like those in the examples above.

Definition 1.10 $A \subseteq B \Leftrightarrow \forall C(C \in A \Rightarrow C \in B)$. *$A \subseteq B$ is read "A is **included in** B" or "A is **contained in** B"; $\subseteq$ is called **inclusion**. We say that A is a **subclass** of B and B is a **superclass** of A; if A is a set, A is a **subset** of B; if B is a set, B is a **superset** of A.*

The reader is warned that many people use $A \subset B$ where we use $A \subseteq B$. Our usage seems more in line with traditional symbols indicating order,

since it is a fact that $A \subseteq A$ for any class A. No one would write $x < x$ for a number x to mean x less than or equal to x, and $\subseteq$ is similar to $\leq$.

Corollary 1.11 $A \subseteq B \Leftrightarrow \forall x(x \in A \Rightarrow x \in B)$.

Proof Certainly if $A \subseteq B$, then for all x, $x \in A \Rightarrow x \in B$. Now assume that $\forall x(x \in A \Rightarrow x \in B)$. Suppose that C is a class and that $C \in A$. Then C is a set, so that $C \in B$. Since C is arbitrary, $\forall X(X \in A \Rightarrow X \in B)$. Hence $A \subseteq B$.

Axiom 1.12 (*Power-set axiom*) $\forall a\, \exists b\, \forall C(C \subseteq a \Rightarrow C \in b)$.

This axiom says intuitively that, if a class A is "small," then there is a small class which has every subclass of A as a member. We already know that there is a class which has every subset of A as a member, namely, $\{X : X \subseteq A\}$. The power-set axiom assures us that there is a set with this property, and this is important in what follows. A similar remark applies to almost all the other axioms introduced.

Axiom 1.13 (*Pairing axiom*) $\forall a\, \forall b\, \exists c(a \in c \wedge b \in c)$.

Of course the set c asserted to exist in 1.13 may have many other elements in addition to a and b.

Axiom 1.14 (*Union axiom*) $\forall a\, \exists b\, \forall C(C \in a \Rightarrow C \subseteq b)$.

If we think of a as a family of sets, 1.14 ensures the existence of a set b which includes every member of a.

Definition 1.15 $0 = \{x : x \neq x\}$. *0 is the* ***empty class.***

Corollary 1.16 $\forall X(X \notin 0)$.

Proof If $X \in 0$, then X is a set and $X \neq X$, which is absurd. Hence $\forall X(X \notin 0)$.

Definition 1.17 $A \cap B = \{x : x \in A \wedge x \in B\}$. *$A \cap B$ is called the* ***intersection*** *of A and B.*

Axiom 1.18 (*Regularity axiom*) $\forall A[A \neq 0 \Rightarrow \exists X(X \in A \wedge X \cap A = 0)]$.

Most of mathematics can be developed without the regularity axiom, but it is very convenient. The regularity axiom rules out such counter-intuitive possibilities as the existence of a class A such that $A \in A$ or

the existence of a sequence like $\cdots \in D \in C \in B \in A$. Moreover, the notion of an order type can be defined conveniently by making essential use of this axiom, and even the notion of a cardinal can be so defined, although we will not follow this procedure. Intuitively one can think of the regularity axiom as assuring that $\in$ has a property analogous to a well-ordering: any nonempty class has a "minimal" element. In this connection compare the notion of a well-founded relation defined in Sec. 8.

To indicate the meaning of the regularity axiom, we give the following theorem, which will be useful later.

Theorem 1.19 (*i*) $A \notin A$.
(*ii*) *There do not exist classes* A, B *such that* $A \in B \in A$.
(*iii*) *There do not exist classes* A, B, C *such that* $A \in B \in C \in A$.

Proof Clearly (*iii*) and (*ii*) imply (*i*). Because of the analogous arguments involved we will establish (*iii*) only. Assume that $A \in B \in C \in A$. Let $D = \{x : x = A \vee x = B \vee x = C\}$. Since A, B, C are all sets, clearly for any X, $X \in D$ iff $X = A \vee X = B \vee X = C$. By the regularity axiom (since $D \neq 0$), choose $X \in D$ such that $X \cap D = 0$. Then $X = A \vee X = B \vee X = C$. But

$$X = A \Rightarrow C \in X \cap D,$$
$$X = B \Rightarrow A \in X \cap D,$$
$$X = C \Rightarrow B \in X \cap D,$$

so that we have a contradiction.

Definition 1.20 $\mathcal{S}A = \{x : x \in A \vee x = A\}$. $\mathcal{S}A$ *is called the* ***successor*** *of* A.

Applied to a natural number A, $\mathcal{S}A$ turns out to be $A + 1$. In connection with this definition we note the following consequence, which really expresses the only fact about $\mathcal{S}$ that will be used later.

Corollary 1.21 $\forall x(x \in \mathcal{S}a \Leftrightarrow x \in a \vee x = a)$.

Proof For any class X we have $X \in \mathcal{S}a$ iff X is a set and ($X \in a$ or $X = a$). Thus by logic $x \in \mathcal{S}a$ implies that $x \in a$ or $x = a$. If $x \in a$ or $x = a$, then, since x always denotes a set, x is a set, and $x \in a$ or $x = a$. Hence $x \in \mathcal{S}a$. Hence, x being arbitrary, $\forall x(x \in \mathcal{S}a \Leftrightarrow x \in a \vee x = a)$.

We should note, however, another consequence of Definition 1.20, which is more or less accidental and not very interesting.

Theorem 1.22 *If A is a proper class, then $\S A = A$.*

Proof Assume that X is any class. If $X \in A$, then X is a set, and $X \in A$ or $X = A$; hence $X \in \S A$. If $X \in \S A$, then $X \in A$ or $X = A$, and X is a set; since A is not a set, the possibility $X = A$ is excluded, so that we always have $X \in A$. Thus $\forall X(X \in \S A \Leftrightarrow X \in A)$, so that by the axiom of extensionality $\S A = A$.

Axiom 1.23 *(Infinity axiom)* $\exists a[0 \in a \wedge \forall X(X \in a \Rightarrow \S X \in a)]$.

This axiom gets its name from the fact that the set a asserted to exist is infinite: 0 is in a, $\S 0$ is in a, $\S\S 0$ is in a, etc., and these elements are all distinct, as seen from 9.13(*ix*). Of course using the power-set axiom, we can obtain even larger sets from a.

On the basis of the axioms given so far we still cannot obtain sets as big as we wish. For this and other technical reasons we need the next axiom, the axiom of substitution. To formulate it, we must introduce the notion of a function, which itself depends upon the notion of an ordered pair, and we first work to define the latter notion.

Definition 1.24 $\{A,B\} = \{x : x = A \vee x = B\}$. $\{A,B\}$ *is called the* ***doubleton*** A, B, *or the* ***unordered pair*** A, B.

Theorem 1.25 $\{a,b\}$ *is a set.*

Proof By the pairing axiom (1.13) let c be a set such that $a \in c$ and $b \in c$. Thus $\{a,b\} \subseteq c$. By the power-set axiom (1.12) let d be a set such that $\forall X(X \subseteq c \Rightarrow X \in d)$. In particular we have $\{a,b\} \in d$. Hence $\{a,b\}$ is a set.

Corollary 1.26 $X \in \{a,b\} \Leftrightarrow X = a \vee X = b$.

Definition 1.27 $\{A\} = \{A,A\}$. $\{A\}$ *is called* ***singleton*** A.

Corollary 1.28 $\{a\}$ *is a set.*

Corollary 1.29 $X \in \{a\} \Leftrightarrow X = a$.

Theorem 1.30 *If* $\{a,b\} = \{c,d\}$, *then* $a = c$ *and* $b = d$, *or else* $a = d$ *and* $b = c$.

Proof From 1.26 we see that $a \in \{a,b\}$. Hence by the hypothesis of the theorem, $a \in \{c,d\}$, so that by another application of 1.26, $a = c$ or $a = d$. The two cases are clearly symmetric, so that we assume, say, that $a = c$.

By 1.26, we have $b \in \{a,b\}$, so that arguing as at first, $b = c$ or $b = d$. If $b = d$, the desired conclusion has been reached, so that we assume instead that $b = c$. Thus $a = b = c$. By 1.26, $d \in \{c,d\}$, so that by the familiar argument $d = a$ or $d = b$. Thus $a = b = c = d$, and the desired conclusion has again been reached.

Definition 1.31 $(A,B) = \{\{A\},\{A,B\}\}$. (A,B) *is called the* ***ordered pair*** A,B *with* ***first coordinate*** A *and* ***second coordinate*** B.

Since A and B enter into this definition in nonsymmetric ways, the concept should depend upon the order. A little reflection shows that Theorem 1.33 expresses the essential property one would expect of a reasonable notion of ordered pair of sets.

Corollary 1.32 (a,b) *is a set.*

Proof By 1.25, 1.28, and 1.31.

Theorem 1.33 *If* $(a,b) = (c,d)$, *then* $a = c$ *and* $b = d$.

Proof Since $(a,b) = \{\{a\},\{a,b\}\}$, $(c,d) = \{\{c\},\{c,d\}\}$, and $\{a\}$, $\{a,b\}$, $\{c\}$, $\{c,d\}$ are all sets by 1.25 and 1.28, Theorem 1.30 applies; hence we have two cases.

Case 1 $\{a\} = \{c\}$ and $\{a,b\} = \{c,d\}$. By 1.29, $a \in \{a\}$; hence $a \in \{c\}$, and by 1.29 again, $a = c$. It remains to be shown that $b = d$. We may apply 1.30 to the assumption $\{a,b\} = \{c,d\}$ to infer that either $a = c$ and $b = d$, which is all right, or that $a = d$ and $b = c$, from which it follows that $b = c = a = d$, as desired.

Case 2 $\{a\} = \{c,d\}$ and $\{a,b\} = \{c\}$. Now $c \in \{c,d\}$ by 1.26, so that $c \in \{a\}$, and hence $c = a$ by 1.29. Similarly $d = a$ and $c = b$, so that $a = c$ and $b = d$.

Definition 1.34 (*i*) R *is a* ***relation*** *iff* $\forall A(A \in R \Rightarrow \exists c\, \exists d[A = (c,d)])$.
(*ii*) $Dmn\ R = \{x : \exists y[(x,y) \in R]\}$. $Dmn\ R$ *is called the* ***domain*** *of* R.
(*iii*) $Rng\ R = \{y : \exists x[(x,y) \in R]\}$. $Rng\ R$ *is called the* ***range*** *of* R.
(*iv*) F *is a* ***function*** *iff* F *is a relation and* $\forall x\, \forall y\, \forall z[(x,y) \in F \wedge (x,z) \in F \Rightarrow y = z]$.

A relation is thus simply any class of ordered pairs of sets. If $(a,b) \in R$, R a relation, we say that a is *R-related to b*. Examples of relations are $<$ among real numbers, congruence of two integers modulo a third, the set of solutions of the equation $x^2 - y^2 = 3$, etc. A function is a rule F assigning to each set a in its domain the unique b such that $(a,b) \in F$.

Axiom 1.35 (*Axiom of substitution*) *If F is a function and Dmn F is a set, then Rng F is a set.*

Our next, and last, axiom, the relational axiom of choice, plays a special role in mathematics. Many of the results in this book can be obtained without it. However, the axiom is now generally accepted as a valid set-theoretical principle, so that we make no effort to avoid its use.

Axiom 1.36 (*Relational axiom of choice*) *If R is a relation, then there is a function F such that $F \subseteq R$ and Dmn F = Dmn R.*

Remark 1.37 In summary, we have introduced the following nine-axiom schemata (eight axioms plus one infinite schema of axioms):

Axiom of extensionality
Class-building axioms
Power-set axiom
Pairing axiom
Union axiom
Regularity axiom
Infinity axiom
Axiom of substitution
Relational axiom of choice

For reference, all the axioms are given at the end of the book. The axioms are sufficient to develop almost all of modern mathematics—calculus, geometry, topology, real and complex analysis, etc. Only quite recently have essential modifications and strengthenings of these axioms been seriously considered by any large number of mathematicians; some of these developments will be described briefly in Sec. 23. A comprehensive development of mathematics based on axioms much like ours can be found in Bourbaki 1939 to the present. For a survey of approaches to the foundation of set theory see Fraenkel, Bar-Hillel 1958. Other axiomatic treatments of set theory are Suppes 1960, Bernays, Fraenkel 1958, Halmos 1960, Gödel 1940, Klaua 1964, Rubin 1967, and the appendix of Kelley 1955, where consequences of essentially the present axioms are developed in outline form.

Once the axioms have been given, many natural questions arise concerning, for example, *consistency*, *completeness*, and *independence*. As to consistency, these axioms have not been shown to be consistent, and by a famous result of Gödel there appears to be little hope of doing so. Naturally no inconsistency has been found, and we have faith that the axioms are, in fact, consistent. For a precise exposition of the result of Gödel see Feferman 1960.

As to completeness, the axiom system is incomplete (if consistent)—there is a sentence φ such that neither φ nor $\neg\varphi$ is derivable from the axioms. An example is the continuum hypothesis discussed in Sec. 22.[1] But even if we add this hypothesis, or its negation, to the axioms, they remain incomplete. In fact, even if we add any finite number of new axioms, or an infinite set of new axioms (as long as we can still effectively recognize when an expression is an axiom), the axiom system remains incomplete (if consistent). This result is also, in essence, due to Gödel; for an exposition see Mendelson 1964.

With regard to independence we mention only the recently proved fact that the axiom of choice is independent of the other axioms if they are consistent; see Gödel 1940 and Cohen 1963 and 1964, and the footnote below.

In view of the fact that the axiom system is not complete (if consistent), the choice of axioms may seem rather arbitrary and capricious. We could add more axioms, or take away some, with equal right, it would seem. In defense the author can only appeal to the requirements of present-day mathematics. In order to develop rigorously the overwhelming majority of contemporary mathematics, all of the axioms are needed. Various additional axioms we might take, like the continuum hypothesis, are not needed for most mathematics; and it is not even clear, with regard to the continuum hypothesis, for example, whether it or its negation should be taken as an axiom if one wanted to extend the system.

EXERCISES

1.38 Show that the following statement can replace the pairing axiom in our system:

$$\forall a\, \forall b\, \exists c(a \subseteq c \wedge b \subseteq c).$$

1.39 Prove that the infinity axiom cannot be derived from the other axioms.

1.40 Show that, if $\forall X(x \in X \Rightarrow y \in X)$, then $x = y$.

1.41 Show that the following are proper classes:

(*a*) $\{X : 0 \notin X\}$.

(*b*) $\{X : y \subseteq X\}$ (for any set y).

(*c*) $\{X : y \in X\}$ (for any set y).

1.42 Let $[X,Y] = \{\{0,X\},\{\{0\},Y\}\}$. Show that $[x,y]$ is a set. Show that, if $[a,b] = [c,d]$, then $a = c$ and $b = d$. (Thus $[X,Y]$ could replace

[1] Actually this has not been established for the present axiom system, but only for related ones; but the same proof can very likely be carried out for the present axioms. A similar comment applies to various other metamathematical observations we make later. See Gödel 1940, Cohen 1963–1964, Vopěnka 1964.

(X,Y) in the text. The choice between these two notions was really quite arbitrary; the only criterion used was that the definition of (X,Y) is simpler.)

1.43 Let $\langle X,Y\rangle = \{\{X\},Y\}$. Find sets a, b, c, d such that $\langle a,b\rangle = \langle c,d\rangle$ and $\neg(a = c \wedge b = d)$.

2 BOOLEAN ALGEBRA OF CLASSES

In this section we consider the elementary properties of combinations of classes—those which can be expressed by Venn diagrams. We begin by discussing the empty class, 0, defined in 1.15, and inclusion, defined in 1.10. Figure 1 is a Venn diagram for the inclusion $A \subseteq B$.

Theorem 2.1 *0 is a set.*

Proof By the infinity axiom.

From now on we call 0 the empty set rather than the empty class. From 0 we can build the sets $\{0\}$, $\{0\ \{0\}\}$, $\{\{0\}\}$, etc. In a certain sense every set, and even every class, is built from the empty set (see Sec. 15). This may seem strange, but the implied simplicity is very useful.

Theorem 2.2 *If $A \subseteq b$, then A is a set.*

Proof By the power-set axiom choose c such that $\forall X(X \subseteq b \Rightarrow X \in c)$. Thus $A \in c$, so that A is a set.

This theorem provides one of the main methods for proving that a given class is a set; frequently we construct a class using one of the class-building axioms 1.7 and then apply 2.2 to conclude that the class is a set.

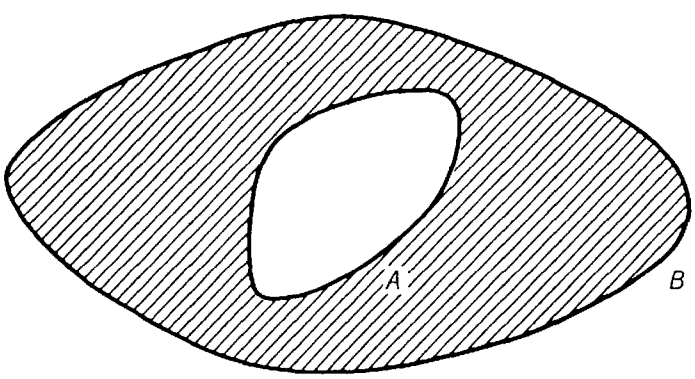


Figure 1

Properties of inclusion are summarized in the following.

Theorem 2.3 (*i*) $0 \subseteq A$.
(*ii*) *If* $A \subseteq 0$, *then* $A = 0$.
(*iii*) $A \subseteq A$.
(*iv*) *If* $A \subseteq B$ *and* $B \subseteq A$, *then* $A = B$.
(*v*) *If* $A \subseteq B$ *and* $B \subseteq C$, *then* $A \subseteq C$.

Proof We prove (*v*) as an example. Assume that $A \subseteq B$ and $B \subseteq C$. Thus by Definition 1.10 we have

(1) $\forall X(X \in A \Rightarrow X \in B)$,
(2) $\forall X(X \in B \Rightarrow X \in C)$.

Now let X be an arbitrary class, and assume that $X \in A$. Then by (1), $X \in B$; by (2), $X \in C$. Thus, X being arbitrary, $\forall X(X \in A \Rightarrow X \in C)$, so that using 1.10, $A \subseteq C$, as desired.

Theorem 2.3(*iv*) is frequently used in proving two classes equal. A good rule is that, if all else fails in trying to prove two classes equal, go back to the method of 2.3(*iv*).

Definition 2.4 $A \subset B$ *iff* $A \subseteq B$ *and* $A \neq B$. $\subset$ *is called* ***proper inclusion;*** *we say that* A *is a* ***proper subclass*** *of* B *and* B *a* ***proper superclass*** *of* A. *Analogously we use the terms* ***proper subset*** *and* ***proper superset*** *when* A, *or* B, *are sets. We also write* $A \supseteq B$ *for* $B \subseteq A$, $A \not\subseteq B$ *for* $\neg(A \subseteq B)$, *etc.*

With regard to the notation for proper inclusion see the comment following Definition 1.10. In Fig. 1 we have $A \subset B$ if there actually are elements in the shaded part, that is, in B but not in A.

Theorem 2.5 (*i*) $0 \subset A$ *iff* $A \neq 0$.
(*ii*) $A \not\subset 0$.
(*iii*) *If* $A \subset B$ *and* $B \subseteq C$, *then* $A \subset C$.
(*iv*) *If* $A \subseteq B$ *and* $B \subset C$, *then* $A \subset C$.
(*v*) *If* $A \subset B$, *then* $B \not\subseteq A$ *and* $B \not\subset A$.
(*vi*) $A \not\subset A$.

Proof We prove (*iii*) as an example. Assume that $A \subset B$ and $B \subseteq C$. Then by 2.4, $A \subseteq B$ and $A \neq B$. Hence by 2.3(*v*), $A \subseteq C$. Now if $A = C$, then $A \subseteq B$ and $B \subseteq A$, so that $A = B$ by 2.3(*iv*), which is impossible. Thus $A \neq C$, so that $A \subset C$, as desired.

An important fact is that there is no largest set; but later in this section we will see that there is a largest class.

Theorem 2.6 (*i*) *For any set a there is a set x such that* $x \notin a$.
(*ii*) *For any set a there is a set b such that* $a \subset b$.

Proof (*i*) By 1.19(*i*).

(*ii*) For a given set a let x be chosen as in (*i*). By the pairing axiom let c be a set such that $a \in c$ and $\{x\} \in c$ (note that $\{x\}$ is a set, by 1.28). By the union axiom let b be a set such that $\forall X(X \in c \Rightarrow X \subseteq b)$. In particular we have $a \subseteq b$ and $\{x\} \subseteq b$. Hence by 1.29, $x \in b$. Thus $a \neq b$, since $x \notin a$ and $x \in b$. Hence $a \subset b$, as desired.

Theorem 2.6(*i*) will be generalized in Theorem 4.16.

We now turn to the discussion of intersection, which was defined in 1.17. Figure 2 is the Venn diagram for an intersection; the shaded part represents $A \cap B$.

Theorem 2.7 $a \cap B$ *is a set.*

Proof If $X \in a \cap B$, then by 1.17, $X \in a$. Thus $a \cap B \subseteq a$, so that by 2.2, $a \cap B$ is a set.

Theorem 2.8 (*i*) $0 \cap A = 0$.
(*ii*) $A \cap A = A$.
(*iii*) $A \cap B = B \cap A$.
(*iv*) $A \cap (B \cap C) = (A \cap B) \cap C$.
(*v*) $A \cap B \subseteq A$.
(*vi*) $A \cap B \subseteq B$.
(*vii*) *If* $X \subseteq A$ *and* $X \subseteq B$, *then* $X \subseteq A \cap B$.
(*viii*) $A \subseteq B$ *iff* $A \cap B = A$.
(*ix*) *If* $A \subseteq C$ *and* $B \subseteq D$, *then* $A \cap B \subseteq C \cap D$.

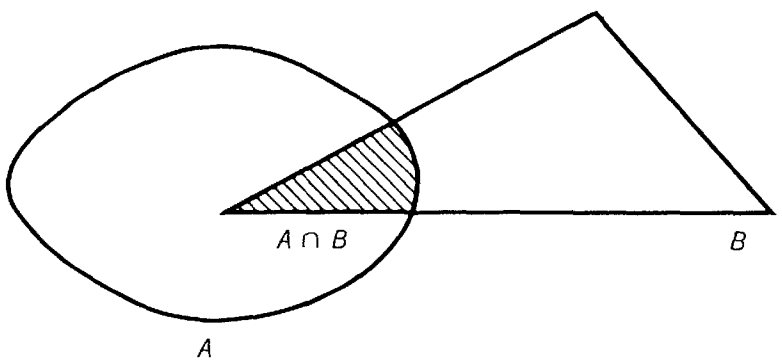


Figure 2

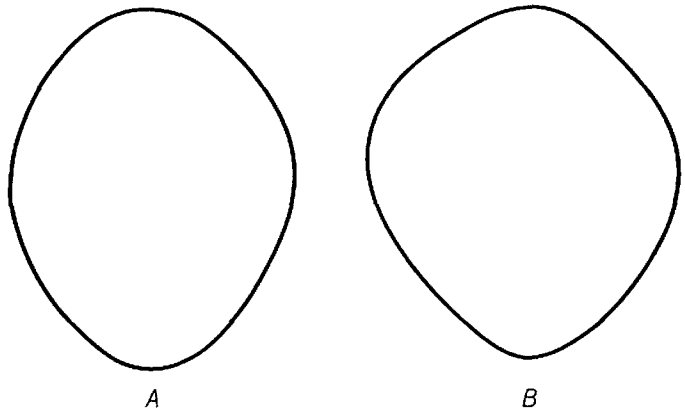


Figure 3

Proof We prove (*iv*) as an example. For any set X we have

$$\begin{aligned} X \in A \cap (B \cap C) \quad &\text{iff} \quad X \in A \wedge X \in B \cap C && \text{by 1.17,} \\ &\text{iff} \quad X \in A \wedge X \in B \wedge X \in C && \text{by 1.17,} \\ &\text{iff} \quad X \in A \cap B \wedge X \in C && \text{by 1.17,} \\ &\text{iff} \quad X \in (A \cap B) \cap C && \text{by 1.17.} \end{aligned}$$

Hence $A \cap (B \cap C) = (A \cap B) \cap C$, by 1.2.

Note that in the terminology of orderings, 2.8(*v*) to (*vii*) expresses that $A \cap B$ is the *greatest lower bound* of the classes A and B with respect to inclusion (see Sec. 8); 2.8(*iii*) and (*iv*) are the *commutative* and *associative* laws for intersection.

Definition 2.9 *Classes A and B are said to be* ***disjoint*** *if $A \cap B = 0$. A class A is called a* ***family of pairwise disjoint sets*** *if any two distinct members of A are disjoint.*

In Fig. 3, the classes A and B are disjoint. Families of pairwise disjoint sets play an important role in mathematics; they will be discussed further in Sec. 7 and will be essential in the definition of addition of cardinals (Sec. 20). For now we content ourselves with a theorem giving some degenerate cases.

Theorem 2.10 (*i*) *0 and A are disjoint for any class A.*
(*ii*) *If $a \notin A$, then A and $\{a\}$ are disjoint.*
(*iii*) *0 and $\{a\}$ are families of pairwise disjoint sets.*
(*iv*) *For $a \neq b$, $\{a,b\}$ is a family of pairwise disjoint sets iff $a \cap b = 0$.*

A Boolean operation that in a sense is dual to that of intersection is union.

Definition 2.11 $A \cup B = \{x : x \in A \vee x \in B\}$. *$A \cup B$ is called the* ***union*** *of A and B.*

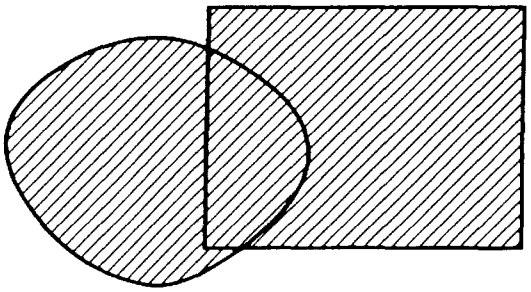


Figure 4

The Venn diagram for a union is given in Fig. 4; the shaded part represents $A \cup B$. It is important to note that elements common to both A and B are also elements of the union.

Theorem 2.12 *$a \cup b$ is a set.*

Proof By 1.25, $\{a,b\}$ is a set. By the union axiom, let c be a set such that $\forall X(X \in \{a,b\} \Rightarrow X \subseteq c)$. By 1.26, we have $a,b \in \{a,b\}$, so that $a \subseteq c$ and $b \subseteq c$. If $X \in a \cup b$, then $X \in a$ or $X \in b$, and hence $X \in c$. Thus $a \cup b \subseteq c$. Thus by 2.2, $a \cup b$ is a set.

Theorem 2.13 (*i*) $A \cup 0 = A$.
(*ii*) $A \cup A = A$.
(*iii*) $A \cup B = B \cup A$.
(*iv*) $A \cup (B \cup C) = (A \cup B) \cup C$.
(*v*) $A \subseteq A \cup B$.
(*vi*) $B \subseteq A \cup B$.
(*vii*) *If* $A \subseteq X$ *and* $B \subseteq X$, *then* $A \cup B \subseteq X$.
(*viii*) $A \subseteq B$ *iff* $A \cup B = B$.
(*ix*) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
(*x*) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(*xi*) *If* $A \subseteq C$ *and* $B \subseteq D$, *then* $A \cup B \subseteq C \cup D$.
(*xii*) $Sx = x \cup \{x\}$, *and* Sx *is a set.*

Proof We prove (*ix*) only. For any set X,

$$\begin{aligned}
X \in A \cup (B \cap C) \quad &\text{iff} \quad X \in A \vee X \in B \cap C && \text{by 2.11,}\\
&\text{iff} \quad X \in A \vee (X \in B \wedge X \in C) && \text{by 1.17,}\\
&\text{iff} \quad (X \in A \vee X \in B) \wedge (X \in A \vee X \in C) && \text{by logic,}\\
&\text{iff} \quad X \in A \cup B \wedge X \in A \cup C && \text{by 2.11,}\\
&\text{iff} \quad X \in (A \cup B) \cap (A \cup C).
\end{aligned}$$

Thus, indeed, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Equation 2.13(*ix*) is illustrated in Fig. 5. On the left $B \cap C$ is shaded /// and A is shaded \\\; $A \cup (B \cap C)$ consists of all regions shaded in any fashion. On the right $A \cup B$ is shaded /// and $A \cup C$ is shaded \\\; $(A \cup B) \cap (A \cup C)$ consists of the crosshatched region ⋇. Hence $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. Equations 2.13(*iii*) and (*iv*) express the commutative and associative laws for union. 2.13(*v*) to (*vii*) express the assertion that $A \cup B$ is the least upper bound of A and B. 2.13(*ix*) and (*x*) are *distributive laws*. With regard to 2.13(*xii*), recall Definition 1.20.

Definition 2.14 $V = \{x : x = x\}$. *V is called the* ***universe.***

Theorem 2.15 (*i*) $\forall x(x \in V)$.
(*ii*) $\forall A(A \subseteq V)$.
(*iii*) $\forall A(A \cap V = A)$.
(*iv*) $\forall A(A \cup V = V)$.

Theorem 2.16 *V is a proper class.*

Proof Suppose that V is a set. Then by 2.6(*i*) and 2.15(*i*) we have a contradiction.

The last Boolean operations which we will introduce in this section are those of complementation and relative complementation.

Definition 2.17 (*i*) $A' = \{x : x \notin A\}$. *A' is the* ***complement*** *of A.*
(*ii*) $A \sim B = \{x : x \in A \wedge x \notin B\}$. *$A \sim B$ is the* ***complement of B relative to*** *A.*

$B \sim A$ is represented by the shaded sector in Fig. 1. It is not required in 2.17 that $B \subseteq A$, however.

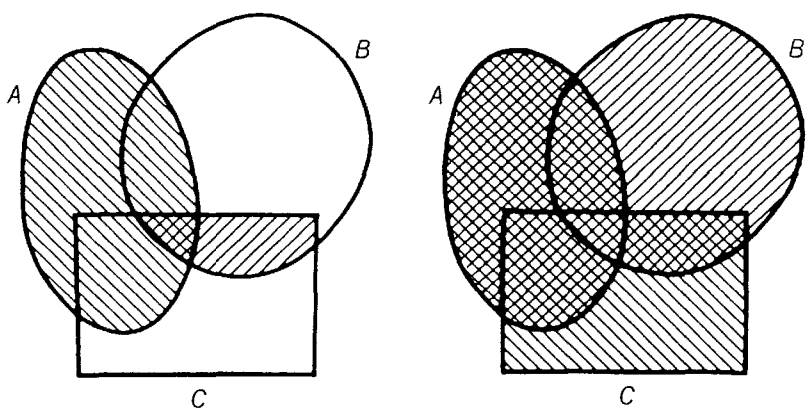


Figure 5

Theorem 2.18 *$a \sim B$ is a set.*

Proof Clearly $a \sim B \subseteq a$, so that the result follows by 2.2.

Theorem 2.19 *(i)* $A' = V \sim A$.
(ii) $A \sim B = A \cap B'$.
(iii) $0' = V$.
(iv) $V' = 0$.
(v) $A'' = A$.
(vi) $A \sim A = 0$.
(vii) $A \subseteq B$ *iff* $B' \subseteq A'$.
(viii) $A \subseteq B$ *iff* $A \sim B = 0$.
(ix) $(A \cap B)' = A' \cup B'$.
(x) $(A \cup B)' = A' \cap B'$.
(xi) $A \cap A' = 0$.
(xii) $A \cup A' = V$.

Proof We prove *(viii)* as an example. First suppose that $A \sim B \neq 0$. Then there is an $x \in A \sim B$, so that by 2.17*(ii)*, $x \in A \wedge x \notin B$. Therefore $A \not\subseteq B$. Second assume that $A \sim B = 0$. Suppose that $x \in A$. Now $x \notin B$ implies that $x \in A \sim B$, by 2.17*(ii)*; hence $x \in B$. Since x is arbitrary, $A \subseteq B$. This completes the proof.

Theorem 2.19*(ix)* and *(x)* are known as the *De Morgan laws.*

Definition 2.20 $\{A,B,C\} = \{A\} \cup \{B,C\}$, $\{A,B,C,D\} = \{A\} \cup \{B,C,D\}$, $(A,B,C) = ((A,B),C)$, $(A,B,C,D) = ((A,B,C),D)$.

Theorem 2.21 *$\{a,b,c\}$, $\{a,b,c,d\}$, (a,b,c), and (a,b,c,d) are sets.*

We could obviously extend Definition 2.20 further, but this is not necessary for our purposes.

Remark 2.22 The operations discussed in this section—union, intersection, and complementation—are called *Boolean operations on classes.* Intersection and union are binary operations; that is, they act upon two classes to produce a third class. These operations will be generalized to act upon arbitrary families of sets in Sec. 5. Many of the simple results of this section have an algebraic flavor. The concepts of this section can be treated purely algebraically; see, e.g., Dwinger 1961, Goodstein 1963, Halmos 1963, and Sikorski 1964. In fact, the theory of Boolean algebras is extensive, with many interesting results as well as many open problems. There are close relationships between this theory and the theory of sentential logic, ring theory, and the theory of switching circuits.

We have dealt in this section mainly with equations involving the Boolean operations. Now there is an automatic method for determining whether or not such an equation can be derived from the axioms (for an explanation of this method see Birkhoff, MacLane 1965). We will not describe the method, but the proof of the following rather complicated equation may give an idea of the tricks involved:

$$(1) \quad (A' \cup B' \cup C)' \cup (A' \cup B)' \cup A' \cup C = V$$

To prove (1), we compute:

$$\begin{aligned}
&(A' \cup B' \cup C)' \cup (A' \cup B)' \cup A' \cup C \\
&\quad = (A'' \cap B'' \cap C') \cup (A'' \cap B') \cup A' \cup C && \text{by 2.19}(x), \\
&\quad = (A \cap B \cap C') \cup (A \cap B') \cup A' \cup C && \text{by 2.19}(v), \\
&\quad = (A \cap B \cap C') \cup (A \cap B') \cup A' \cup (C \cap [V \cup (A \cap B)]) && \text{by 2.15}(iii), (iv), \\
&\quad = (A \cap B \cap C') \cup (A \cap B \cap C) \cup (A \cap B') \cup A' \cup (C \cap V) && \text{by 2.13}(x), \\
&\quad = [A \cap B \cap (C \cup C')] \cup (A \cap B') \cup A' \cup C && \text{by 2.13}(x), \text{ 2.15}(iii), \\
&\quad = (A \cap B \cap V) \cup (A \cap B') \cup A' \cup C && \text{by 2.19}(xii), \\
&\quad = (A \cap B) \cup (A \cap B') \cup A' \cup C && \text{by 2.15}(iii), \\
&\quad = [A \cap (B \cup B')] \cup A' \cup C && \text{by 2.13}(x), \\
&\quad = (A \cap V) \cup A' \cup C && \text{by 2.19}(xii), \\
&\quad = A \cup A' \cup C && \text{by 2.15}(iii), \\
&\quad = V \cup C && \text{by 2.19}(xii), \\
&\quad = V && \text{by 2.15}(iv).
\end{aligned}$$

EXERCISES

2.23 (Concerning proper classes). Prove:

(*a*) If A is a proper class and $A \subseteq B$, then B is a proper class.

(*b*) If A is a proper class, then $A \cup B$ is a proper class.

(*c*) There are proper classes A, B such that $A \cap B = 0$.

(*d*) There are proper classes A, B such that $A \subset B$ and $B \sim A$ is a proper class.

(*e*) $V \sim a$ is a proper class.

2.24 Answer true or false:

(*a*) $0 \subseteq \{0\}$. (*b*) $0 \in \{0\}$.

(*c*) $0 \subset \{0\}$. (*d*) $0 = \{0\}$.

(e) $\{0\} \in \{0\}$.
(f) $\{0\} \notin 0$.
(g) $\{0\} \in \{0,\{0\}\}$.
(h) $\{0\} \subseteq \{0,\{0\}\}$.
(i) $0 \in \{0,\{0\}\}$.
(j) $\{0,0\} = \{0\}$.
(k) $\{0,0\} = \{\{0\}\}$.
(l) $\{0,\{0\}\} = \{\{0\},0\}$.
(m) $\{0\} \in \{\{\{0\}\}\}$.
(n) $\{\{0\}\} \in \{\{\{0\}\}\}$.
(o) $\{0,\{0,\{0,0\}\}\} \subseteq \{0,\{0\}\}$.
(p) $\{0\} \subseteq 0$. ⊬
(q) $\{\{0\},\{0\}\} \subseteq \{0,\{\{0\}\}\}$.

2.25 For any classes A, B let $A \oplus B = (A \sim B) \cup (B \sim A)$ (this is the operation of *symmetric difference*). Show that for any classes A, B, C, $A \oplus (B \oplus C) = (A \oplus B) \oplus C$.

2.26 (Continuing 2.25). Let A be a nonempty set such that for some set a the following conditions hold:

(1) $0,a \in A$.
(2) $\forall x(x \in A \Rightarrow x \subseteq a)$.
(3) $\forall x(x \in A \Rightarrow a \sim x \in A)$.
(4) $\forall x\, \forall y(x \in A \wedge y \in A \Rightarrow x \cup y \in A \wedge x \cap y \in A)$.

Show that with $\oplus$ as addition and $\cap$ as multiplication A forms a ring.

2.27 Prove the following statements:

(a) $(A' \cup B' \cup C)' \cup B' \cup A' \cup C = V$.
(b) $[(A' \cup B)' \cup B]' \cup (B' \cup A)' \cup A = V$.
(c) $([(B' \cup C)' \cup A' \cup C]' \cup D)' \cup (A' \cup B)' \cup D = V$.
(d) $A' \cup (A' \cup B)' \cup B = V$.

3 ALGEBRA OF RELATIONS

In this section we consider the basic properties of binary relations. We will not discuss relations that are not binary, and we reserve special kinds of binary relations, like functions, orderings, and equivalence relations, until later. Besides giving a basis for these more special concepts, the general theory of relations is occasionally applied in its general form in mathematics, for example in the theory of uniform spaces.

The notion of a relation was defined in 1.34(*i*), and we recall that a relation is simply a class of ordered pairs of sets.

A relation can be pictured as in Fig. 6. The lines indicate which pairs are in the relation. Thus (x,y), (z,y), (u,w), (v,w), etc., are in the pictured relation. Of course the two circles of the figure may overlap.

Another way of thinking of relations intuitively is as subsets of a plane, as in Fig. 7. We think of V, the universe, as laid out along both the horizontal and vertical axes. The point with X-coordinate x and Y-coordinate y is identified with the ordered pair (x,y). Relations may thus be thought of as sets of points in the plane.

In virtue of a vacuous implication, we have the following.

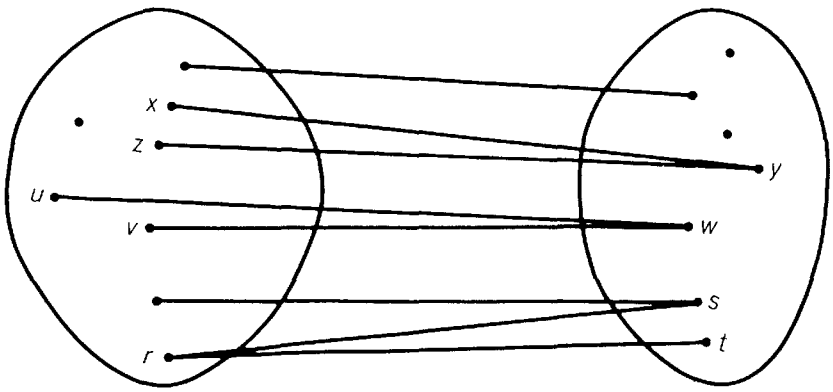


Figure 6

Corollary 3.1 *0 is a relation.*

The following simple theorem is quite useful.

Theorem 3.2 *Let R and S be relations. Then*
(*i*) $R \subseteq S \Leftrightarrow \forall x\, \forall y\, [(x,y) \in R \Rightarrow (x,y) \in S]$.
(*ii*) (*Extensionality principle for relations*). $R = S$ *iff* $\forall x\, \forall y\, [(x,y) \in R \Leftrightarrow (x,y) \in S]$.

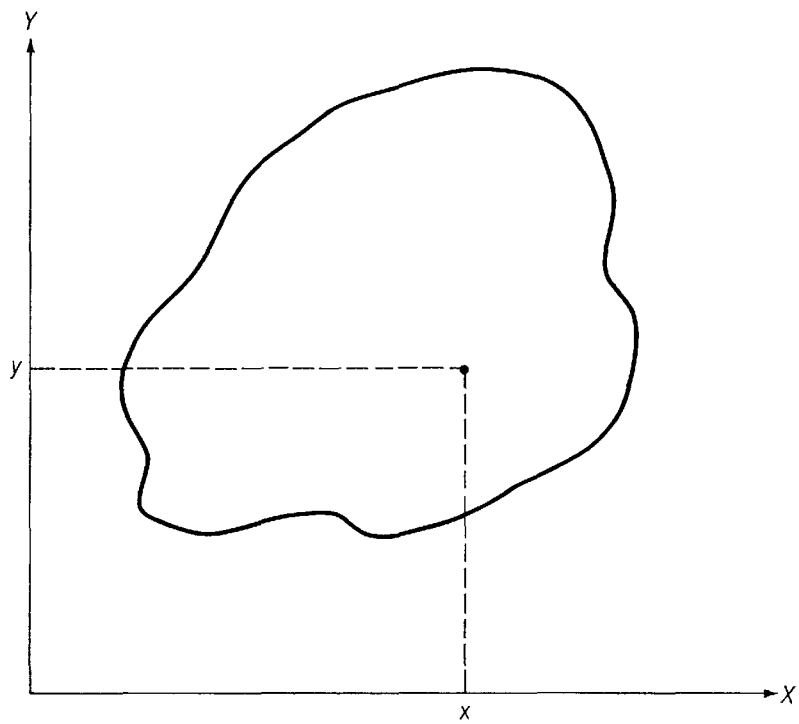


Figure 7

Proof Clearly (*ii*) follows from (*i*), by 2.3(*iv*). The direction $\Rightarrow$ in (*i*) is obvious. Now assume that R and S are relations and $\forall x\, \forall y\, [(x,y) \in R \Rightarrow (x,y) \in S]$. Suppose that $X \in R$. By 1.34(*i*), choose x, y such that $X = (x,y)$. Then $(x,y) \in R$, so that $(x,y) \in S$; that is, $X \in S$. Thus $\forall X\, (X \in R \Rightarrow X \in S)$, so that, by 1.10, $R \subseteq S$, as desired.

Theorem 3.2(*ii*) has an intuitive usefulness because it enables one to prove relations equal by arguing with arbitrary ordered pairs rather than with arbitrary sets. Intuitively speaking, ordered pairs have a different connotation from sets, and 3.2(*ii*) helps in maintaining this intuitive separation. For the same reason, it is helpful to extend the notation of 1.9.

Definition 3.3 (*i*) *For any set-theoretical expression φ not involving A, let $\{(x,y) : \varphi(x,y)\} = \{A : \text{there exist } x, y \text{ such that } A = (x,y) \text{ and } \varphi(x,y)\}$; similarly if φ does not involve B, etc., and similarly for $\{(x,z) : \psi(x,z)\}$, etc.*
(*ii*) *We sometimes write xRy instead of $(x,y) \in R$ and $x \not R y$ instead of $(x,y) \notin R$.*

Thus $\{(x,y) : \varphi(x,y)\}$ consists of all ordered pairs (x,y) such that $\varphi(x,y)$ holds.

Definition 3.4 (*i*) $R|S = \{(x,z) :$ *there is a y such that* $(x,y) \in R$ *and* $(y,z) \in S\}$. *$R|S$ is the* ***relative product*** *of R and S.*
(*ii*) $R^{-1} = \{(x,y) : (y,x) \in R\}$. *$R^{-1}$ is the* ***converse*** *or* ***inverse*** *of R.*

Note that for any u, v, $u(R|S)v$ iff there is a w such that uRw and wSv; $uR^{-1}v$ iff vRu.

Pictorially, representing R as in Fig. 7, R^{-1} is obtained from R by reflecting in the main diagonal. To interpret $R|S$ geometrically, imagine Fig. 7 supplied with a third dimension and a third axis, the Z-axis. Let R and S be two subsets of the XY-plane, with points identified with ordered pairs. Form the cylinders R° and S° with bases R and S, respectively, and main axes parallel to the Z-axis. Thus R° consists of all triples (x,y,z) with $(x,y) \in R$, and S° of all triples (x,y,z) with $(x,y) \in S$. Rotate R° 90° about the X-axis, forming $R^{\circ\circ}$; $R^{\circ\circ}$ is the set of all triples (x,y,z) such that $(x,z) \in R$. Similarly rotate S° 90° about the Y-axis, forming $S^{\circ\circ}$; $S^{\circ\circ}$ is the set of all triples (x,y,z) such that $(z,y) \in S$. Hence the intersection $R^{\circ\circ} \cap S^{\circ\circ}$ of the two cylinders $R^{\circ\circ}$ and $S^{\circ\circ}$ consists of all triples (x,y,z) such that $(x,z) \in R$ and $(z,y) \in S$. The projection of $R^{\circ\circ} \cap S^{\circ\circ}$ on the XY-plane, parallel to the Z-axis, consists of all pairs (x,y) such that for some z, $(x,z) \in R$ and $(z,y) \in S$; i.e., the projection is $R|S$.

Note that Definition 3.4 makes sense even if R and S are not rela-

tions. Usually, in fact, we will define notions in as great a generality as possible.

Relative product and converse are the two basic operations on relations; some fundamental properties of these notions are given in the following.

Theorem 3.5 (*i*) $R|(S|T) = (R|S)|T$.
(*ii*) $R|(S \cup T) = (R|S) \cup (R|T)$ *and* $(R \cup S)|T = (R|T) \cup (S|T)$.
(*iii*) $R|(S \cap T) \subseteq (R|S) \cap (R|T)$ *and* $(R \cap S)|T \subseteq (R|T) \cap (S|T)$.
(*iv*) *If* $R \subseteq S$, *then* $R|T \subseteq S|T$ *and* $T|R \subseteq T|S$.
(*v*) *If* R, S, T *are relations, then* $(R|S) \cap T = 0$ *iff* $(R^{-1}|T) \cap S = 0$.
(*vi*) $(R \cup S)^{-1} = R^{-1} \cup S^{-1}$.
(*vii*) $(R^{-1})^{-1} = R$ *if* R *is a relation.*
(*viii*) $(R|S)^{-1} = S^{-1}|R^{-1}$.

Proof We prove (*i*) and (*v*) as examples. First suppose that $(x,w) \in R|(S|T)$. Choose y such that xRy and $y(S|T)w$. Choose z such that ySz and zTw. Then $x(R|S)z$, so that $x[(R|S)|T]w$. Hence, by 3.2(*i*), $R|(S|T) \subseteq (R|S)|T$. The opposite inclusion is proved similarly, and (*i*) follows.

To prove (*v*), first suppose that $(R|S) \cap T \neq 0$. Say that $(x,y) \in (R|S) \cap T$. By Definition 3.4, there is, then, a z such that $xRzSy$. Hence $zR^{-1}xTy$, so that $z(R^{-1}|T)y$. Also zSy, so that $(z,y) \in (R^{-1}|T) \cap S$, and $(R^{-1}|T) \cap S \neq 0$. Therefore, $(R|S) \cap T \neq 0$ implies $(R^{-1}|T) \cap S \neq 0$, and the converse is similar.

The inclusions in 3.5(*iii*) cannot be replaced by equalities. To show this for the first inclusion, let $R = \{(0,0),(0,\{0\})\}$, $S = \{(0,0)\}$, and $T = \{(\{0\},0)\}$. Then $(0,0) \in (R|S) \cap (R|T)$, but $S \cap T = 0$ and hence $R|(S \cap T) = 0$. Thus $(R|S) \cap (R|T) \neq R|(S \cap T)$. The second inclusion is treated similarly. Usually when stating only inclusions in theorems, it is possible to show that the inclusions cannot be replaced by equalities.

Theorem 3.5 gives the basic and most useful properties of $|$ and $^{-1}$, but there are many more complex facts that are also easily verified. For example, the inclusion

$$(1) \quad (R|S) \cap (T|U) \subseteq R|[(R^{-1}|T) \cap (S|U^{-1})]|U$$

is demonstrated as follows. Assume that $(x,z) \in (R|S) \cap (T|U)$. Choose y such that xRy and ySz, and choose w such that xTw and wUz. Then $yR^{-1}xTw$, so that $y(R^{-1}|T)w$. Also $ySzU^{-1}w$, so that $y(S|U^{-1})w$. Hence $(y,w) \in (R^{-1}|T) \cap (S|U^{-1})$, xRy, and wUz, so that $(x,z) \in R|[(R^{-1}|T) \cap (S|U^{-1})]|U$. The inclusion follows, by 3.2(*i*).

Definition 3.6 $I = \{(x,y) : x = y\}$. *I is called the* ***identity relation,*** *or* ***diagonal.***

Theorem 3.7 (*i*) *For any relation R, $R|I = I|R = R$.*
(*ii*) $I^{-1} = I$.

We now discuss briefly the concepts of domain and range of relations, which were introduced in 1.34. When we think of R pictorially, as in Fig. 6, *Dmn* R is the circle on the left and *Rng* R that on the right. In Fig. 7, *Dmn* R is the projection of R on the X-axis and *Rng* R that on the Y-axis.

Theorem 3.8 *Let R and S be relations.*
(*i*) *Dmn* $(R \cup S) =$ *Dmn* $R \cup$ *Dmn* S *and Rng* $(R \cup S) =$ *Rng* $R \cup$ *Rng* S.
(*ii*) *Dmn* $(R \cap S) \subseteq$ *Dmn* $R \cap$ *Dmn* S *and Rng* $(R \cap S) \subseteq$ *Rng* $R \cap$ *Rng* S.
(*iii*) *Dmn* $R \sim$ *Dmn* $S \subseteq$ *Dmn* $(R \sim S)$ *and Rng* $R \sim$ *Rng* $S \subseteq$ *Rng* $(R \sim S)$.
(*iv*) *If $R \subseteq S$, then Dmn $R \subseteq$ Dmn S and Rng $R \subseteq$ Rng S.*
(*v*) *Dmn* $0 = 0 =$ *Rng* 0.
(*vi*) *Dmn* $I = V =$ *Rng* I.
(*vii*) *Dmn* $R^{-1} =$ *Rng* R *and Rng* $R^{-1} =$ *Dmn* R.
(*viii*) *Dmn* $(R|S) \subseteq$ *Dmn* R *and Rng* $(R|S) \subseteq$ *Rng* S.

Proof We prove the first part of (*iii*) as an example. Suppose that $x \in$ *Dmn* $R \sim$ *Dmn* S. By 1.34(*ii*), choose y such that $(x,y) \in R$. Now $(x,y) \in S$ implies that $x \in$ *Dmn* S, by 1.34(*ii*), but $x \notin$ *Dmn* S since $x \in$ *Dmn* $R \sim$ *Dmn* S, by 2.17(*ii*). Hence $(x,y) \notin S$. Thus $(x,y) \in R \sim S$, by 2.17(*ii*), so that $x \in$ *Dmn* $(R \sim S)$, by 1.34(*ii*). This completes the proof.

Definition 3.9 *Fld* $R =$ *Dmn* $R \cup$ *Rng* R. *Fld R is called the* ***field*** *of R. R is said to be* ***on*** *A if Fld* $R = A$.

The intuitive meaning of the notion of the field of a relation should be clear. *Fld* has properties analogous to 3.8(*i*) to (*vi*). In addition we have the following.

Theorem 3.10 *Let R and S be relations.*
(*i*) *Fld* $(R^{-1}) =$ *Fld* R.
(*ii*) *Fld* $(R|S) \subseteq$ *Dmn* $R \cup$ *Rng* S.

Definition 3.11 $A \times B = \{(a,b) : a \in A \text{ and } b \in B\}$. *$A \times B$ is called the* ***cartesian product*** *of the classes A and B. Further, let $A \times B \times C = (A \times B) \times C$, $A \times B \times C \times D = (A \times B \times C) \times D$, and $A \times B \times C \times D \times E = (A \times B \times C \times D) \times E$.*

$A \times B$ can be thought of as a rectangle with sides A and B; it is thus a subclass of the plane represented in Fig. 7.

We will not state any properties of $A \times B \times C$, or $A \times B \times C \times D$, or $A \times B \times C \times D \times E$; they follow easily from the simple properties of $A \times B$, which we will give.

Theorem 3.12 *$a \times b$ is a set.*

Proof By 2.12, $a \cup b$ is a set. By the power-set axiom, let c be a set such that

(1) $\forall X(X \subseteq a \cup b \Rightarrow X \in c)$.

Using the power-set axiom again, let d be a set such that

(2) $\forall X(X \subseteq c \Rightarrow X \in d)$.

Now we claim that $a \times b \subseteq d$. Let X be any element of $a \times b$. By Definition 3.11, choose $x \in a$ and $y \in b$ such that $X = (x,y)$. Thus, by Definition 2.11, $x \in a \cup b$ and $y \in a \cup b$. Now, from 1.26 and 1.29, we infer that $\{x\} \subseteq a \cup b$ and $\{x,y\} \subseteq a \cup b$. Hence, by (1), $\{x\} \in c$ and $\{x,y\} \in c$. By 1.26 again, $\{\{x\},\{x,y\}\} \subseteq c$; i.e., $X \subseteq c$, by 1.31. Hence, by (2), $X \in d$. Thus, indeed, $a \times b \subseteq d$. Therefore $a \times b$ is a set, by 2.2.

Theorem 3.13 (*i*) *$A \times B$ is a relation.*
(*ii*) $A \times 0 = 0 \times A = 0$.
(*iii*) *If $A \neq 0$ and $B \neq 0$, then $A \times B \neq 0$.*
(*iv*) *If $A \subseteq C$ and $B \subseteq D$, then $A \times B \subseteq C \times D$.*
(*v*) $A \times (B \cup C) = (A \times B) \cup (A \times C)$ *and* $(A \cup B) \times C = (A \times C) \cup (B \times C)$.
(*vi*) $A \times (B \cap C) = (A \times B) \cap (A \times C)$ *and* $(A \cap B) \times C = (A \times C) \cap (B \times C)$.
(*vii*) $A \times (B \sim C) = (A \times B) \sim (A \times C)$ *and* $(A \sim B) \times C = (A \times C) \sim (B \times C)$.
(*viii*) $(A \times B)^{-1} = B \times A$.
(*ix*) *If $B \cap C = 0$, then $(A \times B)|(C \times D) = 0$; if $B \cap C \neq 0$, then $(A \times B)|(C \times D) = A \times D$.*
(*x*) *Dmn* $(A \times B) = A$ *if* $B \neq 0$ *and Rng* $(A \times B) = B$ *if* $A \neq 0$.
(*xi*) *If R is a relation, then $R \subseteq Dmn\ R \times Rng\ R$.*

Theorem 3.14 *If r and s are relations, then $r|s$, r^{-1}, $Dmn\ r$, $Rng\ r$, and $Fld\ r$ are sets.*

Proof By the union axiom, let a be a set such that

(1) $\forall X(X \in r \Rightarrow X \subseteq a)$,

and b a set such that

(2) $\forall X(X \in a \Rightarrow X \subseteq b)$.

We claim that $Dmn\ r \subseteq b$. Suppose that $x \in Dmn\ r$. By 1.34(*ii*), choose y such that $(x,y) \in r$. Then, by (1), $(x,y) \subseteq a$. Now $(x,y) = \{\{x\},\{x,y\}\}$, by 1.31. By 1.26, $\{x\} \in (x,y)$, and hence, by 1.11, $\{x\} \in a$. Hence, according to (2), $\{x\} \subseteq b$. Using 1.11 and 1.29, we see that $x \in b$, as desired. Hence $Dmn\ r \subseteq b$, so that, using 2.2, $Dmn\ r$ is a set. In a similar manner one sees that $Rng\ r$ is a set. $Fld\ r = Dmn\ r \cup Rng\ r$; $Fld\ r$ is a set, by virtue of 2.12. Now $r^{-1} \subseteq Dmn\ (r^{-1}) \times Rng\ (r^{-1}) = Rng\ r \times Dmn\ r$, by 3.13(*xi*) and 3.8(*vii*), so that r^{-1} is a set, by 3.12 and 2.2. Finally, $r|s \subseteq Dmn\ (r|s) \times Rng\ (r|s) \subseteq Dmn\ r \times Rng\ s$, by 3.13(*xi*), 3.8(*viii*), and 3.13(*iv*), so that $r|s$ is a set, by virtue of 3.12 and 2.2.

We consider one more operation on relations, which plays an important role in the discussion of functions.

Definition 3.15 $R^*A = \{y : \exists x \in A \text{ such that } (x,y) \in R\}$. *$R^*A$ is called the R-**image** of A.*

When R is conceived as in Fig. 6, A is usually taken to be a subclass of the left-hand circle; R^*A consists, then, of the subclass of the right-hand circle consisting of sets which are at the right end of a line with the left end in A. In Fig. 7, A will be considered as a class of points on the X-axis; we form the strip with base A parallel to the Y-axis, intersect with R, and then project on the Y-axis parallel to the X-axis, forming R^*A.

Theorem 3.16 *Let R and S be relations.*

(*i*) $0^*A = 0$.
(*ii*) $R^*0 = 0$.
(*iii*) $R^*(A \cup B) = (R^*A) \cup (R^*B)$.
(*iv*) $R^*(A \cap B) \subseteq (R^*A) \cap (R^*B)$.
(*v*) $(R^*A) \sim (R^*B) \subseteq R^*(A \sim B)$.
(*vi*) *If $A \subseteq B$, then $R^*A \subseteq R^*B$.*
(*vii*) $(R|S)^*A = S^*(R^*A)$.
(*viii*) $I^*A = A$.

(*ix*) $Dmn\ (R|S) = R^{-1}{}^*(Dmn\ S)$ *and* $Rng\ (R|S) = S^*(Rng\ R)$.
(*x*) $(A \times B)^*C = B$ *if* $A \cap C \neq 0$; $(A \times B)^*C = 0$ *if* $A \cap C = 0$.
(*xi*) $R^*A \subseteq Rng\ R$.

Proof We prove (*vii*) and the first part of (*ix*). For (*vii*), first suppose that $y \in (R|S)^*A$. By 3.15, choose $a \in A$ such that $a(R|S)y$. By 3.4(*i*), choose x such that aRx and xSy. Then, by 3.15, $x \in R^*A$, and so $y \in S^*(R^*A)$. Conversely, assume that $z \in S^*(R^*A)$. Choose $t \in R^*A$ such that tSz. Choose $b \in A$ such that bRt. Thus $b(R|S)z$. Hence $z \in (R|S)^*A$, and the proof of (*vii*) is complete.

For the first part of (*ix*), first suppose that $x \in Dmn\ (R|S)$. Choose z such that $x(R|S)z$. Choose y such that $xRySz$. Thus $y \in Dmn\ S$ and $yR^{-1}x$, so that $x \in R^{-1}{}^*(Dmn\ S)$. Second, suppose that $a \in R^{-1}{}^*(Dmn\ S)$. Choose $b \in Dmn\ S$ such that $b(R^{-1})a$. Thus aRb. Choose c such that bSc. Thus $a(R|S)c$, so that $a \in Dmn\ (R|S)$, as desired.

Theorem 3.17 *r^*A is a set.*

Proof By 3.16(*xi*) and 3.14.

Remark 3.18 The general theory of relations is highly developed both explicitly, as in this section, and algebraically, where one abstracts from the actual definition of a relation. Many general theorems are given in Schröder 1895 and Russell, Whitehead 1925 to 1927. The algebraic theory is developed, for example, in Tarski 1941, Chin, Tarski 1951, Jónsson, Tarski 1952, Lyndon 1961, and Monk 1964. There is one more very important operation on relations, the operation of forming the transitive closure of a relation, which will be discussed in Sec. 11.

EXERCISES

3.19 Show that the following statements hold for any relations R, S, T:
(*a*) $(R|S) \cap T \subseteq R|[(R^{-1}|T) \cap S]$.
(*b*) $(R|S) \cap T \subseteq R|R^{-1}|T$.
(*c*) $R \cap S \cap T \subseteq R|S^{-1}|T$.
(*d*) If $R|T \subseteq R$ and $R|T^{-1} \subseteq R$, then $R \cap (S|T) = (R \cap S)|T$.

3.20 Show that none of the inclusions in 3.8(*ii*) and (*iii*) and 3.16(*iv*) and (*v*) can be replaced by equalities.

3.21 (Sets and classes). Prove:
(*a*) I is a proper class.
(*b*) If R is a relation and $Dmn\ R$ and $Rng\ R$ are sets, then R is a set.
(*c*) If $Dmn\ R$ is a proper class, then R is a proper class.
(*d*) If $A \neq 0$, then $A \times V$ is a proper class.
(*e*) Give an example of proper classes R, S such that $R|S$ is a set.

(*f*) If R is a relation, then R is a set iff R^{-1} is a set.

(*g*) Give an example of a relation R and a set a such that R^*a is a proper class.

3.22 Prove:

(*a*) If A and B are proper classes, then $\{A,B\} = 0$, $\{A\} = 0$, and $(A,B) = \{0\}$.

(*b*) If A is a proper class, then $\{A,x\} = \{x,A\} = \{x\}$ and $(A,x) = \{0,\{x\}\}$, $(x,A) = \{\{x\}\}$.

(*c*) If A and B are proper classes and R is a relation, then $(A,B) \notin R$ and $(A,x) \notin R$; but a relation R can be constructed for which $(x,V) \in R$ for some x.

3.23 Show that $(a \times b) \times c = a \times (b \times c)$ fails in general.

4 FUNCTIONS

The notion of a function is one of the most important in mathematics. The definition was given in 1.34(*iv*); intuitively we think of a function F as a rule that assigns to each $x \in Dmn\ F$ the unique y such that $(x,y) \in F$. Thus a function is simply a special kind of relation. With a relation R pictured as in Fig. 6, R is a function iff the situation illustrated by r, s, t does not occur, that is, iff, given any element in the left-hand circle, there is at most one line beginning with the element and leading to the right-hand circle. If R is pictured as in Fig. 7, then R is a function iff each vertical line (i.e., each line parallel to the Y-axis) meets R in at most one point (see Fig. 8).

Two important properties of inverses of functions not shared by other relations are the following [cf. 3.16(*iv*) and (*v*)].

Theorem 4.1 *Let F be a function. Then*

(*i*) $F^{-1*}(A \cap B) = (F^{-1*}A) \cap (F^{-1*}B)$.

(*ii*) $F^{-1*}(A \sim B) = (F^{-1*}A) \sim (F^{-1*}B)$.

Proof (*i*) In virtue of 3.16(*iv*), we have only to prove that $(F^{-1*}A) \cap (F^{-1*}B) \subseteq F^{-1*}(A \cap B)$. Suppose, then, that $x \in (F^{-1*}A) \cap (F^{-1*}B)$. Thus $x \in F^{-1*}A$ and $x \in F^{-1*}B$. By 3.15, choose $a \in A$ and $b \in B$ such that $(a,x) \in F^{-1}$ and $(b,x) \in F^{-1}$. By Definition 3.4(*ii*), we then have $(x,a) \in F$ and $(x,b) \in F$, so that the fact that F is a function yields $a = b$. Thus $a \in A \cap B$ and $x \in F^{-1*}(A \cap B)$, as desired.

(*ii*) In view of 3.16(*v*), we have only to show that $F^{-1*}(A \sim B) \subseteq (F^{-1*}A) \sim (F^{-1*}B)$. Assume that $x \in F^{-1*}(A \sim B)$. Since $A \sim B \subseteq A$, we infer from 3.16(*vi*) that $x \in F^{-1*}A$. To complete the proof, it is enough to derive a contradiction from the assumption that $x \in F^{-1*}B$. This assumption implies that there is a $y \in B$ such that $(y,x) \in F^{-1}$; the earlier assumption $x \in F^{-1*}(A \sim B)$ implies the existence of $z \in A \sim B$

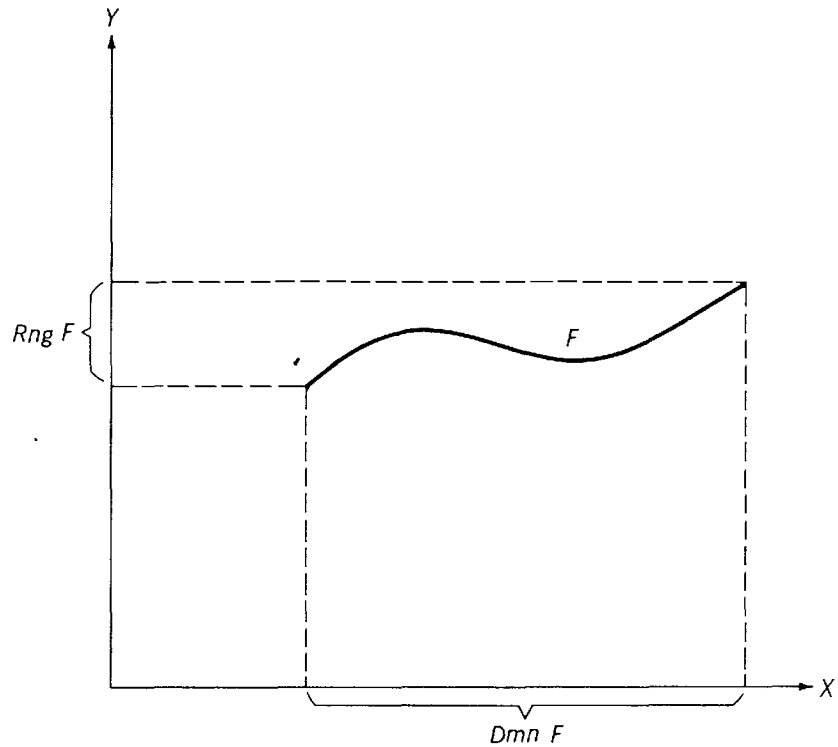


Figure 8

with $(z,x) \in F^{-1}$. Thus, by the definition of $^{-1}$, xFy and xFz; since F is a function, $y = z$. But then $y \in B$ and $y \in A \sim B$, which is impossible.

We now introduce the customary functional notation.

Definition 4.2 $F(A) = \{x : \forall y(A \text{ is a set} \wedge AFy \Rightarrow x \in y)\}$. *Instead of* $F(A)$ *one may write* FA *or* F_A. *In case* $A = (a,b)$ *one may write* $F(a,b)$, Fab, *or* F_{ab} *instead of* $F(A)$.

One reads $F(A)$ as F *of* A or F *applied to* A. We say that F *sends* A *to* $F(A)$, or *assigns* $F(A)$ *to* A, or *makes* $F(A)$ *correspond to* A. When using the notation F_A, one usually says that F *is a function indexed by* I, where $I = Dmn\ F$, or simply that F *is a family of sets indexed by* I.

Note that $F(A)$ is essentially defined to be the intersection of all sets y such that AFy; if F is a function and $A \in Dmn\ F$, then there is only one such y, and $F(A)$ equals that y. Since the general notion of intersection has not been introduced yet, the definition must be spelled out as in 4.2 (see Sec. 5), whose meaning is fully expressed in the following.

Theorem 4.3 *Let* F *be a function.*

(*i*) *If* $x \in Dmn\ F$, *then* $F(x)$ *is the unique* y *such that* $(x,y) \in F$, *and hence* $(x,F(x)) \in F$; *in particular,* $F(x)$ *is a set if* $x \in Dmn\ F$.

(*ii*) *If* $x \in Dmn\ F$, *then* $F(x) = z$ *iff* $(x,z) \in F$.
(*iii*) *If* $A \notin Dmn\ F$, *then* $F(A) = V$.

Proof (*i*) Assume that $x \in Dmn\ F$. Then we may choose y such that $(x,y) \in F$; y is uniquely determined. We need to show that $F(x) = y$. First, suppose that $a \in F(x)$. Now x is a set and xFy, so that, by 4.2, $a \in y$. Hence, a being arbitrary, $F(x) \subseteq y$. Second, suppose that $a \in y$. If z is such that xFz, then $z = y$ and hence $a \in z$; hence $a \in F(x)$. Thus $y \subseteq F(x)$, so that $F(x) = y$.

(*ii*) By (*i*).

(*iii*) Assume that $A \notin Dmn\ F$. If A is a proper class, then the implication in the definition of $F(A)$ is always vacuously true, so that $F(A) = V$. If A is a set, then for any set y it is never true that AFy, so that again the implication is vacuously true, and hence $F(A) = V$.

With regard to part (*iii*) of this theorem, we may mention the peculiar fact that, if we had simplified 4.3 to read $F(A) = \{X : \forall y(AFy \Rightarrow X \in y)\}$, then part (*iii*) would still be true (under the assumption that F is a function), but the proof would be more complex; parts (*i*) and (*ii*) would also remain true, with the same proofs as above.

Note that some symbols we have introduced appear to be functions but are not. Thus $\mathcal{S}$, the successor "function" introduced in 1.20, is not a function. $\mathcal{S}$ is a symbol that has no meaning apart from the context $\mathcal{S}(A)$. The same applies to $\{A\}$, introduced in 1.27. Note that $\{V\} = 0$; 4.3(*iii*) would be contradicted if $\{\ \}$ were treated as a function. Compare the Appendix here. The point is that $\mathcal{S}$, for example, is a part of our symbolism and can be applied even to proper classes, although functions are normally applied only to sets.

Theorem 4.4 *Let* F *and* G *be functions.*
(*i*) $F = G$ *iff* $Dmn\ F = Dmn\ G$ *and* $F(x) = G(x)$ *for all* $x \in Dmn\ F$ (*extensionality principle for functions*).
(*ii*) $F = G$ *iff* $Fx = Gx$ *for every set* x.
(*iii*) $y \in F^*A$ *iff there is an* $x \in A$ *such that* $F(x) = y$.
(*iv*) $x \in F^{-1*}A$ *iff* $F(x) \in A$.
(*v*) *If* $F \subseteq G$ *and* $x \in Dmn\ F$, *then* $Fx = Gx$.

Proof We prove (*i*) only. Note that (*ii*) follows from (*i*), since $x \notin Dmn\ F$ iff $Fx = V$ iff $Gx = V$ iff $x \notin Dmn\ G$. The direction $\Rightarrow$ follows on logical grounds. Now assume that $Dmn\ F = Dmn\ G$ and $\forall x \in Dmn\ F[F(x) = G(x)]$. In order to apply 3.2(*ii*), on grounds of symmetry it is enough to take sets x, y such that xFy and prove that xGy. By 4.3, the assumption xFy implies that $x \in Dmn\ F$ and $F(x) = y$. Hence $G(x) = y$, so that, by 4.3, xGy, as desired.

Theorem 4.5 *I is a function. For any x, $I(x) = x$.*

Theorem 4.6 *If F and G are functions and $Dmn\ F \cap Dmn\ G = 0$, then $F \cup G$ is a function.*

Definition 4.7 *(i)* $F \circ G = G|F$. *$F \circ G$ is called the* ***composition*** *of F and G.*
(ii) F is 1-1 *iff F and F^{-1} are functions. One says, then, that F is* ***one-to-one, one-one,*** *or* ***biunique.***

The definition of composition is made essentially to ensure the property expressed in 4.8(*iii*). The function F is one-one provided that distinct elements of its domain are sent by F onto distinct elements of its range. Again we note that Definition 4.7 makes sense even if F and G are not functions (see the similar remark following 3.4).

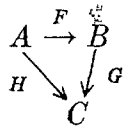
Theorem 4.8 *Let F and G be functions.*
(i) $Dmn\ (F \circ G) = G^{-1*}\ (Dmn\ F) \subseteq Dmn\ G$.
(ii) $Rng\ (F \circ G) = F^{*}(Rng\ G) \subseteq Rng\ F$.
(iii) If $x \in Dmn\ (F \circ G)$, *then* $(F \circ G)\ (x) = F(G(x))$.
(iv) The following three conditions are equivalent:
  (a) F is one-one.
  (b) For all $x, y \in Dmn\ F$, *if* $F(x) = F(y)$, *then* $x = y$.
  (c) For all $x, y \in Dmn\ F$, *if* $x \neq y$, *then* $F(x) \neq F(y)$.

Proof (*i*) and (*ii*) follow directly from 4.4(*iv*), 3.16(*ix*), and 3.8(*viii*). To prove (*iii*), suppose that $x \in Dmn\ (F \circ G)$. Thus, by 4.7(*i*), $x \in Dmn\ (G|F)$. Choose y such that $x(G|F)y$; choose z such that $xGzFy$. Then, using 4.3(*i*), $G(x) = z$ and $F(z) = y$; i.e., $F(G(x)) = y = (F \circ G)(x)$. Finally, to prove (*iv*), first note that (*b*) and (*c*) are equivalent on logical grounds alone. Now assume that (*a*) holds, $x, y \in Dmn\ F$, and $F(x) = F(y)$. Thus 4.3(*i*) implies that $(x,F(x)), (y,F(x)) \in F$ and hence $(F(x),x)$, $(F(x),y) \in F^{-1}$. Since F^{-1} is a function, $x = y$, as desired. Hence $(a) \Rightarrow (b)$. Now assume that (*b*) holds. To prove (*a*), we need to show that F^{-1} is a function; assume, then, that $xF^{-1}y$ and $xF^{-1}z$. Thus yFx and zFx, so that, by 4.3(*i*), $F(y) = x = F(z)$. The assumption (*b*) then yields $y = z$, as desired. This completes the proof.

We now introduce some important ways of speaking about functions.

Definition 4.9 *(i) We say that F is a function* ***from*** *A* ***into*** *B iff F is a function, $Dmn\ F = A$, and $Rng\ F \subseteq B$. We may also say that F* ***maps*** *A* ***into*** *B; we abbreviate this expression by writing* $F : A \to B$, *or* $A \xrightarrow{F} B$.

We say that a triangle



commutes if $F : A \to B$, $G : B \to C$, $H : A \to C$, *and* $G \circ F = H$.

(*ii*) ${}^{A}B = \{f : f \text{ is a function from } A \text{ into } B\}$.

(*iii*) *F is a function from A* ***onto*** *B iff F is a function from A into B and* $Rng\ F = B$.

(*iv*) *F is a* ***one-one correspondence*** *between A and B provided that F is a one-one function mapping A onto B.*

(*v*) *F is a* ***permutation*** *of A if F is a one-one function and maps A onto A.*

Theorem 4.10 (*i*) *If F is a function, then F maps $Dmn\ F$ onto $Rng\ F$.*

(*ii*) $0 : 0 \to A$; *if* $F : 0 \to A$, *then* $F = 0$.

(*iii*) *If* $F : A \to 0$, *then* $A = F = 0$.

(*iv*) *If* $F : A \to B$ *and* $B \subseteq C$, *then* $F : A \to C$.

(*v*) *If* $F : A \to B$ *and* $G : B \to C$, *then* $G \circ F : A \to C$.

(*vi*) *If* $F : A \to B$, *then F is one-one iff for all C and all G, H, if* $G : C \to A$, $H : C \to A$, *and* $F \circ G = F \circ H$, *then* $G = H$.

(*vii*) *If* $F : A \to B$, *then F maps onto B iff for all C and all G, H, if* $G : B \to C$, $H : B \to C$, *and* $G \circ F = H \circ F$, *then* $G = H$.

(*viii*) ${}^{0}B = \{0\}$: *if* $A \neq 0$, *then* ${}^{A}0 = 0$.

(*ix*) *If* $B \neq 0$, *then* ${}^{a}B \neq 0$.

Proof (*i*) to (*v*) are quite easy to prove. To prove (*vi*), assume that $F : A \to B$. Direction $\Rightarrow$. Suppose that F is one-one. Assume that $G : C \to A$, $H : C \to A$, and $F \circ G = F \circ H$. Thus $Dmn\ G = C = Dmn\ H$, by 4.9(*i*). If $x \in C$, then $F(G(x)) = (F \circ G)\ (x) = (F \circ H)\ (x) = F(H(x))$, by 4.8(*iii*); since F is one-one, we obtain, by 4.8(*iv*), $G(x) = H(x)$. Since x is arbitrary, it follows, by 4.4(*i*), that $G = H$. Direction $\Leftarrow$. Suppose that F is not one-one. Then there exist distinct $a,b \in A$ such that $F(a) = F(b)$, by 4.8(*iv*). Let $G = \{(x,y) : x \in A \text{ and } y = a\}$, and let $H = \{(x,y) : x \in A \text{ and } y = b\}$. Then $G : A \to A$, $H : A \to A$, and for any $x \in A$, $(F \circ G)\ (x) = F(G(x)) = F(a) = F(b) = F(H(x)) = (F \circ H)\ (x)$; furthermore, by (*v*), $Dmn\ (F \circ G) = A = Dmn\ (F \circ H)$. Hence, by 4.4(*i*), $F \circ G = F \circ H$. But $G \neq H$.

The proof of (*vii*) is analogous. Direction $\Rightarrow$. Suppose that F maps onto B. Assume that $G : B \to C$, $H : B \to C$, and $G \circ F = H \circ F$. Then $Dmn\ G = Dmn\ H$. For any $b \in B$ choose $a \in A$ such that $F(a) = b$, by 4.9(*iii*); then $G(b) = G(F(a)) = (G \circ F)\ (a) = (H \circ F)\ (a) = H(F(a)) = H(b)$. Hence, b being arbitrary, $G = H$. Direction $\Leftarrow$. Suppose that F does not map onto B. Choose $b \in B \sim Rng\ F$. Let $G = \{(x,y) : x \in B$

$\sim \{b\}, y = 0\} \cup \{(b,0)\}$ and $H = \{(x,y) : x \in B \sim \{b\}, y = 0\} \cup \{(b,\{0\})\}$. Then $G : B \to \{0,\{0\}\}$, $H : B \to \{0,\{0\}\}$, and hence $Dmn\ (G \circ F) = Dmn\ (H \circ F)$, by (v). If $a \in A$, then $F(a) \in B \sim \{b\}$ and hence $(G \circ F)(a) = G(F(a)) = 0 = H(F(a)) = (H \circ F)(a)$. Thus $G \circ F = H \circ F$. But obviously $G \neq H$.

Condition $(viii)$ is easily proved. To prove (ix), suppose that $B \neq 0$, and choose $b \in B$. Let $F = \{(x,y) : x \in a \text{ and } y = b\}$. Then clearly F is a function from a into B. Furthermore, $F \subseteq a \times \{b\}$, so that, by 3.12, F is a set. Hence $F \in {}^aB$, so that ${}^aB \neq 0$.

Note that, if A is a proper class, then ${}^AB = 0$, since there is no function f that is a set and maps A into B.

The following theorem is a very useful variant of the substitution axiom.

Theorem 4.11 *If F is a one-one function from A into B and B is a set, then A is a set.*

Proof $Rng\ F \subseteq B$, so that $Rng\ F$ is a set, by Theorem 2.2. F^{-1} is a function mapping $Rng\ F$ onto A, so that A is a set, by the substitution axiom.

Definition 4.12 $F \restriction A = F \cap (A \times Rng\ F)$. *$F \restriction A$ is called the* ***restriction*** *of F to A.*

Theorem 4.13 *If F is a function, then $F \restriction A$ is a function, $Dmn\ (F \restriction A) = A \cap Dmn\ F$, $Rng\ (F \restriction A) = F^*(A)$, and for any $x \in A \cap Dmn\ F$, $(F \restriction A)(x) = F(x)$.*

The following gives a useful method for proving functions one-one or onto.

Theorem 4.14 *Suppose that F maps A into B.*

(i) F maps onto B iff there is a function G from B into A such that $F \circ G = I \restriction B$.

(ii) F is one-one iff $A = 0$ or else $A \neq 0$ and there is a function G from B into A such that $G \circ F = I \restriction A$.

(iii) F is one-one onto B iff there is a function G from B into A such that $F \circ G = I \restriction B$ and $G \circ F = I \restriction A$; in this case we have $G = F^{-1}$.

Proof (i) $\Rightarrow$. By the relational axiom of choice, let G be a function such that $G \subseteq F^{-1}$ and $Dmn\ G = Dmn\ F^{-1}$. Thus $Dmn\ G = Rng\ F = B$, by 3.8(vii). If $y \in Rng\ G$, choose x such that $(x,y) \in G$; then $(x,y) \in F^{-1}$ and consequently $(y,x) \in F$, so that $y \in A$. Since y is arbitrary, $Rng\ G \subseteq A$. Thus G maps B into A. By 4.10(v), $F \circ G$ has domain B. Assume that

$x \in B$. Then $(x,G(x)) \in G$, by 4.3(i), so that $(x,G(x)) \in F^{-1}$ and hence $(G(x),x) \in F$. Thus, by 4.3(ii), $(F \circ G)(x) = F(G(x)) = x$. x being arbitrary, it follows that $\forall x(x \in B \Rightarrow (F \circ G)(x) = (\mathrm{I} \restriction B)(x))$ (using 4.5 and 4.13), and hence $F \circ G = \mathrm{I} \restriction B$, by 4.4($i$). $\Leftarrow$. For this direction we do not need the relational axiom of choice. Suppose that $b \in B$. Then $b = (\mathrm{I} \restriction B)\ (b) = (F \circ G)\ (b) = F(G(b))$. Thus $(G(b),b) \in F$, so that $b \in Rng\ F$. Since b is arbitrary and we are given that $Rng\ F \subseteq B$, it follows that $Rng\ F = B$; that is, F maps onto B.

(ii) $\Rightarrow$. Assume that F is one-one and $A \neq 0$. Choose $a \in A$. Let $G = F^{-1} \cup \{(x,y) : x \in B \sim Rng\ F \text{ and } y = a\}$. Then G is a function from B into A. Clearly $Dmn\ G = B$, and so, by 4.10(v), $Dmn\ (G \circ F) = A$. For any $a \in A$ we have $a \in Dmn\ F$, $(a,F(a)) \in F$, $(F(a),a) \in F^{-1}$, $(F(a),a) \in G$, and hence $a = G(F(a)) = (G \circ F)(a) = (\mathrm{I} \restriction A)(a)$. Thus, by 4.4($i$), $G \circ F = \mathrm{I} \restriction A$. $\Leftarrow$. To prove that F is one-one, we apply 4.8(iv) (b); suppose, then, that $a,b \in A$ and $F(a) = F(b)$. Then $a = (\mathrm{I} \restriction A)\ (a) = (G \circ F)\ (a) = G(F(a)) = G(F(b)) = (G \circ F)\ (b) = (\mathrm{I} \restriction A)\ (b) = b$, as desired.

(iii) $\Rightarrow$. Let $G = F^{-1}$; then G is a function from B into A. If $b \in B$, then, since F maps onto B, there is an $a \in A$ such that $(a,b) \in F$. Thus $(b,a) \in F^{-1}$. Hence $G(b) = a$, and $(F \circ G)\ (b) = F(G(b)) = F(a) = b$. Hence $F \circ G = \mathrm{I} \restriction B$. Similarly $G \circ F = \mathrm{I} \restriction A$. $\Leftarrow$. See the corresponding proofs in (i) and (ii).

In proving Theorem 4.14(i), it is essential to use the relational axiom of choice; see Rubin, Rubin 1963, pp. 5–7.

Theorem 4.15 (i) *If F is a function, then F^*a is a set.*
(ii) *If f is a function, then f^*A is a set.*
(iii) *If F is a function, then $F \restriction a$ is a set.*
(iv) *${}^b a$ is a set.*
(v) *If F is a function and $Dmn\ F$ is a set, then F and $Rng\ F$ are sets.*

Proof (i) Let $G = F \restriction a$. Then $Dmn\ G = a \cap Dmn\ F$ is a set, by 2.2. Hence, by the axiom of substitution, $F^*(a) = Rng\ G$ is a set.

(ii) Similar to the proof of (i).

(iii) Clearly $(F \restriction a) \subseteq a \times F^*(a)$, so that $F \restriction a$ is a set, by 3.12 and 2.2.

(iv) By 3.12, $b \times a$ is a set. From the power-set axiom, we obtain a set c such that $\forall X(X \subseteq b \times a \Rightarrow X \in c)$. Thus ${}^b a \subseteq c$, so that ${}^b a$ is a set, by 2.2.

(v) Since $F = F \restriction Dmn\ F$, F is a set, by (iii). Hence $Rng\ F$ is a set, by 3.14 (or the substitution axiom).

Theorem 4.16 *If A and B are sets, then there is a set C and a one-one function F mapping B onto C such that $A \cap C = 0$.*

Proof Let $F = \{(x,y) : x \in B \text{ and } y = (x,A)\}$. Clearly F is a function and $Dmn\ F = B$. Let $C = Rng\ F$. By the substitution axiom, C is a set. If $x,y \in B$ and $x \neq y$, then $Fx = (x,A) \neq (y,A) = Fy$. Thus F is one-one. Suppose that $x \in A \cap C$. Thus $x \in A$, and $x = Fy$ for some $y \in B$, and hence $x = (y,A)$. Thus $A \in \{y,A\} \in \{\{y\},\{y,A\}\} = (y,A) = x \in A$, contradicting 1.19(*iii*). Hence there is no such x; that is, $A \cap C = 0$. This completes the proof.

Theorem 4.16 is useful in some applications; see, for example, the replacement theorem in algebra (van der Waerden 1940, pp. 39–42). One may say that in the theorem we simply "rename" the elements of B in order to "make" A and B disjoint.

To conclude this section, we introduce some special notation for functions, expanding Definitions 1.6, 1.9, and 3.3. It is again necessary to discuss the very foundations of our development; namely, we need to single out a different collection of expressions from that delineated in Definition 1.6.

Definition 4.17 (*i*) *We define the notion of a* ***set-theoretical term****:* (*a*) *any capital italic letter is a term, and* 0, *V, and I are terms;* (*b*) *if* $\sigma, \tau, \rho, \xi, \pi$ *are terms, then so are* $(\sigma \cap \tau)$, $(\S\sigma)$, $\{\sigma,\tau\}$, $\{\sigma\}$, (σ,τ), $Dmn\ \sigma$, $Rng\ \sigma$, $\sigma \cup \tau$, σ', $\sigma \sim \tau$, $\{\sigma,\tau,\rho\}$, $\{\sigma,\tau,\rho,\xi\}$, (σ,τ,ρ), (σ,τ,ρ,ξ), $\sigma|\tau$, σ^{-1}, $Fld\ \sigma$, $\sigma \times \tau$, $\sigma \times \tau \times \rho$, $\sigma \times \tau \times \rho \times \xi$, $\sigma \times \tau \times \rho \times \xi \times \pi$, $\sigma^*\tau$, $\sigma(\tau)$, $\sigma\tau$, σ_τ, $\sigma \circ \tau$, ${}^\tau\sigma$, and $\sigma\restriction\tau$.

(*ii*) *If* $\sigma(X)$ *is a term and* $\varphi(X)$ *an expression, then* $\{\sigma(X) : \varphi(X)\} = \{Y : \exists X(\varphi(X) \wedge Y = \sigma(X))\}$, *where* Y *does not occur in* $\sigma(X)$ *or* $\varphi(X)$. *Similarly we use the notations* $\{\sigma(X,Y) : \varphi(X,Y)\}$, $\{\sigma(Z) : \varphi(Z)\}$, *etc.*

(*iii*) $\langle\sigma(X) : X \in I\rangle$ *denotes the function* $\{(X,\sigma(X)) : X \in I\}$ *for any term* $\sigma(X)$ *such that* $\sigma(X)$ *is a set for each* $X \in I$.

A term is thus an expression that denotes a class, and a formula an expression representing an assertion about classes. The notion of a term has to be expanded as we give more definitions, but we will not do so explicitly. The notion of a term will be used only in the context of 4.17(*ii*) and (*iii*) anyway, and we will be concerned only with concrete examples. Some examples of 4.17(*ii*) are:

$$\{(x,y) : \varphi(x,y)\} = \{z : \exists x\, \exists y[\varphi(x,y) \wedge z = (x,y)]\}$$

[thus Definition 3.3 is a special case of 4.17(*ii*)];

$$\{\S x : x \in A\} = \{y : \exists x(x \in A \wedge y = \S x)\};$$
$$\{a \times b : a \in A, b \in B\} = \{x : \exists a\, \exists b(a \in A \wedge b \in B \wedge x = a \times b)\}.$$

We meet many more examples in the course of the development of set

theory. Note that we do not hesitate to use lowercase letters, although Definition 4.17 allows only capital letters. Some examples of 4.17(*iii*) are:

$\langle a : b \in A \rangle = \{(b,a) : b \in A\}$ = the function with domain A that assigns a to each $b \in A$;

$\langle \S x : x \in A \rangle = \{(x,\S x) : x \in A\}$ = the function with domain A that assigns $\S x$ to each $x \in A$;

$\langle x \cup a : x \in A, x \cap y = 0 \rangle = \langle x \cup a : x \in A \cap \{z : z \cap y = 0\} \rangle = \{(x,x \cup a) : x \in A, x \cap y = 0\}$ = the function with domain $A \cap \{x : x \cap y = 0\}$ assigning $x \cup a$ to each element x of its domain.

Note that, if F is a function, then $F = \langle F_i : i \in Dmn\ F \rangle$. Again there will be many examples of the use of 4.17(*iii*) later.

In case $\sigma(X)$ is a proper class for some $X \in I$, $\langle \sigma(X) : X \in I \rangle$ does not have its intended meaning. Letting $F = \langle \sigma(X) : X \in I \rangle$, it is easily seen that $FX = X$ if $\sigma(X)$ is a proper class. Naturally we are really interested in $\langle \sigma(X) : X \in I \rangle$ only when $\sigma(X)$ is a set for every $X \in I$.

Remark 4.18 There are many general and abstract accounts of functions. With regard in particular to Definition 4.17, see Church 1941 [the exact notation of 4.17(*iii*) is due to Alfred Tarski]. Functions, and in particular functions that preserve structure on sets, are abstractly considered in the rapidly developing theory of categories; for an introductory account see Kuroš, Livšic, Šulgeĭfer 1960. For the theory of categories, a special concept of a mapping is introduced. A *mapping* is a triple (A,f,B) such that f is a function mapping A into B. A composition of mappings (A,f,B) and (C,g,D) is defined if and only if $A = D$, and then, by definition, $(A,f,B) \circ (C,g,D) = (C, f \circ g, B)$. (A,f,B) is *injective* if f is one-one; *surjective* if $B = Rng\ f$; *bijective* if both injective and surjective. Frequently the distinction between the mapping (A,f,B) and the function f is ignored, although it is important in many considerations of category theory.

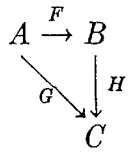
EXERCISES

Prove the following statements.

4.19 If F and G are functions from A into A, then $(A \times A) \sim ([(A \times A) \sim F] \circ [(A \times A) \sim G])$ is a function.

4.20 Suppose that F maps A onto B, G maps A into C, and for all $x,y \in A$, $F(x) = F(y)$ implies that $G(x) = G(y)$. Then there is an H such

that $H : B \to C$ and the triangle



commutes.

4.21 If F is a one-one function from A onto C and G is a one-one function from B onto D, then

$$\langle (F(a),G(b)) : (a,b) \in A \times B \rangle$$

is a one-one function from $A \times B$ onto $C \times D$.

4.22 If f is a one-one function from a onto c and g is a one-one function from b onto d, then

$$\langle g \circ h \circ f^{-1} : h \in {}^{a}b \rangle$$

is a one-one function from ${}^{a}b$ onto ${}^{c}d$.

4.23 For any sets A, B, C,

$$\langle \langle \langle f(a,b) : b \in B \rangle : a \in A \rangle : f \in {}^{(A \times B)}C \rangle$$

is a one-one function from ${}^{(A \times B)}C$ onto ${}^{A}({}^{B}C)$.

4.24 If $F \subseteq A \times B$, then (i) F is a function iff (ii) for every $X \subseteq B$, $F^*(F^{-1*}(X)) \subseteq X$, iff (iii) $F^{-1} | F \subseteq I$.

4.25 (a) If F and G are functions and $F \subseteq G$, then $F = G \restriction (Dmn\ F)$.

(b) If G is a function and $F \subseteq G$, then F is a function.

(c) If F and G are functions, $A = Dmn\ F \cap Dmn\ G$, and $F \restriction A = G \restriction A$, then $F \cup G$ is a function.

4.26 If a function f maps a set A onto B, then $\langle f^{-1*}\{b\} : b \in B \rangle$ maps B one-one into SA.

5 INFINITE BOOLEAN OPERATIONS

The binary notions of union and intersection discussed in Sec. 2 are not sufficient when we want to take the union or intersection of an infinite family of sets. The general notions are introduced in this section. The elementary properties of the general notions are analogous to those of the binary notions.

Definition 5.1 $\bigcup A = \{x : \textit{there is a } y \in A \textit{ such that } x \in y\}$. *We call* $\bigcup A$ *the* ***union*** *of the family* A.

If we represent A as in Fig. 9, the circles being the various members of A, then $\bigcup A$ consists of all the shaded part.

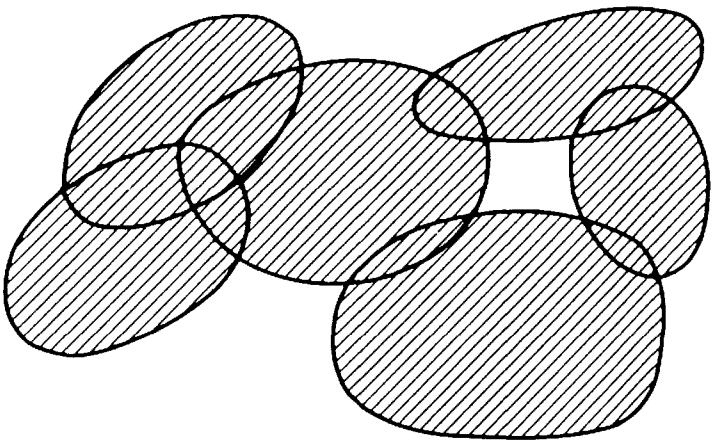


Figure 9

Theorem 5.2 (*i*) $\bigcup 0 = 0$.
(*ii*) $\bigcup\{a\} = a$.
(*iii*) $\bigcup\{a,b\} = a \cup b$.
(*iv*) *If* $A \subseteq B$, *then* $\bigcup A \subseteq \bigcup B$.
(*v*) $\bigcup(A \cup B) = (\bigcup A) \cup (\bigcup B)$.
(*vi*) $\bigcup(A \cap B) \subseteq (\bigcup A) \cap (\bigcup B)$.
(*vii*) *If* $x \in A$, *then* $x \subseteq \bigcup A$.
(*viii*) *If* $\forall x(x \in A \Rightarrow x \subseteq B)$, *then* $\bigcup A \subseteq B$.

Proof We prove (*ii*) as an example. First, suppose that $x \in \bigcup\{a\}$. Choose $y \in \{a\}$ such that $x \in y$, by 5.1. By 1.29, $y = a$. Hence $x \in a$. Second, suppose that $x \in a$. By 1.29, $a \in \{a\}$, so that, by 5.1, $x \in \bigcup\{a\}$, as desired.

Theorem 5.3 $\bigcup a$ *is a set.*

Proof By the union axiom, let b be a set such that $\forall X(X \in a \Rightarrow X \subseteq b)$. Now for any x, $x \in \bigcup A \Rightarrow \exists y(x \in y \in a) \Rightarrow \exists y(x \in y \wedge y \subseteq b) \Rightarrow x \in b$. Thus $\bigcup A \subseteq b$, so that, by 2.2, $\bigcup a$ is a set.

We now introduce a more standard notation for the infinite union.

Definition 5.4 $\bigcup_{i \in I} A_i = \bigcup Rng\ A$.

This notation is usually used when A is a function with domain I; frequently, then, we use the alternate terminology, speaking of A as an *indexed family of sets* (see 4.2). Recall that A_i is simply the function value $A(i)$. We will occasionally stray from the exact form of Definition 5.4. Thus we may use $\bigcup_{a \in A} a$ for $\bigcup A$, $\bigcup_{i \in I, i \neq 0} A_i$ for $\bigcup_{j \in J} B_j$, where $J = I \sim \{0\}$ and $B = A \restriction J$, etc.

Theorem 5.5 *If A is a function with domain I, then*

$$x \in \bigcup_{i \in I} A_i \text{ iff } \exists i \in I (x \in A_i).$$

We can now formulate commutative and associative laws for infinite unions.

Theorem 5.6 (*Commutative law for infinite unions*) *Let A be an indexed family of sets, with $Dmn\ A = I$. Let F be a function from a class J onto I. Then*

$$\bigcup_{i \in I} A_i = \bigcup_{j \in J} A_{Fj}.$$

Proof Observe, as in the remarks following 5.4, that $\bigcup_{j \in J} A_{Fj} = \bigcup_{j \in J} (A \circ F)_j = \bigcup Rng\ (A \circ F)$, and $\bigcup_{i \in I} A_i = \bigcup Rng\ A$. Since F maps onto $Dmn\ A$, $Rng\ A = Rng\ (A \circ F)$, and the desired result follows.

Theorem 5.7 (*Associative law for infinite unions*) *Let A be an indexed family of sets, with $Dmn\ A = I \times J$. Then*

$$\bigcup_{i \in I, j \in J} A_{ij} = \bigcup_{i \in I} \left(\bigcup_{j \in J} A_{ij}\right).$$

Proof The unions involved may be interpreted as follows:

$$\bigcup_{i \in I, j \in J} A_{ij} = \bigcup_{k \in I \times J} A_k,$$
$$\bigcup_{i \in I} \left(\bigcup_{j \in J} A_{ij}\right) = \bigcup_{i \in I} B_i,$$

where for each $i \in I$, $B_i = \bigcup_{j \in J} (C_i)_j$, $C_i = \langle A_{ij} : j \in J \rangle$. From now on we assume that the reader can put infinite unions in the standard of form 5.4, if he thinks it necessary.

Suppose that $x \in \bigcup_{i \in I, j \in J} A_{ij}$. Choose $i_0 \in I$, $j_0 \in J$ such that $x \in A_{i_0 j_0}$. Hence $x \in \bigcup_{j \in J} A_{i_0 j}$, so that $x \in \bigcup_{i \in I} \left(\bigcup_{j \in J} A_{ij}\right)$. The converse is proved similarly.

As in 3.16(*iii*), we have the following.

Theorem 5.8 *If R is a relation and A is an indexed family of sets with domain I, then*

$$R^* \left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} R^*(A_i).$$

Theorem 5.9 *If M is a family of functions having the property that (i) for all $f,g \in M$, $f\restriction(Dmn\ f \cap Dmn\ g) = g\restriction(Dmn\ f \cap Dmn\ g)$, then $\bigcup M$ is a function.*

Proof Suppose that $(x,y) \in \bigcup M$ and $(x,z) \in \bigcup M$. Then there exist $f,g \in M$ such that $(x,y) \in f$ and $(x,z) \in g$. Thus $x \in Dmn\ f \cap Dmn\ g$, so that, by (*i*), $fx = gx$; that is, $y = z$.

Theorem 5.9 is useful in constructing functions; see, for example, the proof of the important Theorem 13.1. Note that (*i*) holds in particular if $Dmn\ f \cap Dmn\ g = 0$ whenever $f \neq g$, and that 5.9 can be applied to show that $\bigcup M$ is one-one when *Dmn* is replaced by *Rng* and f, g by f^{-1}, g^{-1}, where both f and g are one-one.

Definition 5.10 $\bigcap A = \{x : x \in y \text{ for all } y \in A\}$. $\bigcap A$ *is called the* ***intersection*** *of* A.

This notion is illustrated in Fig. 10; the circles represent members of A, and the shaded part is the intersection of A.

Theorem 5.11 (*i*) $\bigcap 0 = V$.
(*ii*) $\bigcap\{a\} = a$.
(*iii*) $\bigcap\{a,b\} = a \cap b$.
(*iv*) $A \subseteq B \Rightarrow \bigcap B \subseteq \bigcap A$.
(*v*) $\bigcap(A \cup B) = (\bigcap A) \cap (\bigcap B)$.
(*vi*) $(\bigcap A) \cup (\bigcap B) \subseteq \bigcap(A \cap B)$.
(*vii*) $x \in A \Rightarrow \bigcap A \subseteq x$.
(*viii*) $\forall x(x \in A \Rightarrow b \subseteq x) \Rightarrow b \subseteq \bigcap A$.

Proof We prove (*iv*) as an example. Assume that $A \subseteq B$, and $x \in \bigcap B$. To prove that $x \in \bigcap A$, let y be an arbitrary member of A. Then $y \in B$, so that, since $x \in \bigcap B$, $x \in y$. Thus $\forall y(y \in A \Rightarrow x \in y)$, so that $x \in \bigcap A$, as desired.

Theorem 5.12 (*i*) $\bigcap\bigcap(a,b) = a$.
(*ii*) $\bigcap\bigcap\bigcap\{(a,b)\}^{-1} = b$.

Proof $\bigcap\bigcap(a,b) = \bigcap\bigcap\{\{a\},\{a,b\}\} = \bigcap\{a\} = a$, by 5.11(*iii*) and (*ii*). *Thus* (*i*) holds, and (*ii*) follows easily.

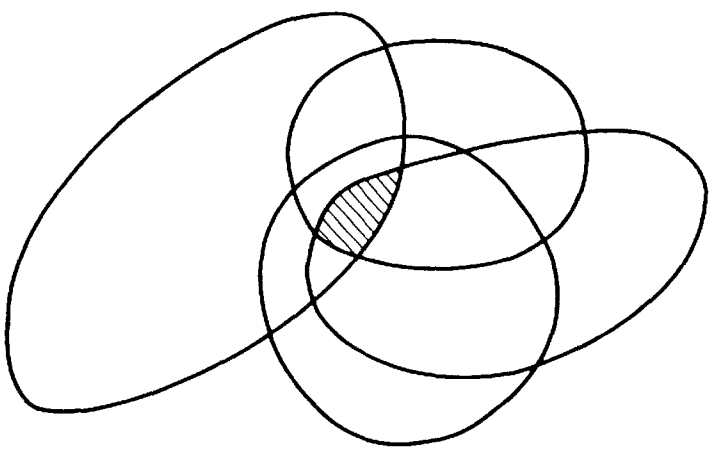


Figure 10

Theorem 5.12 provides an easy method to pick out the first and second coordinates of an ordered pair (a,b).

Definition 5.13 $1^{st}x = \bigcap\bigcap x$. $2^{nd}x = \bigcap\bigcap\bigcap\{x\}^{-1}$.

Of course Definition 5.13 is really interesting only when x is an ordered pair. Directly from 5.12 we obtain the following.

Theorem 5.14 (*i*) $1^{st}(a,b) = a$.
(*ii*) $2^{nd}(a,b) = b$.
(*iii*) $1^{st}1^{st}(a,b,c) = a$.
(*iv*) $2^{nd}1^{st}(a,b,c) = b$.
(*v*) $2^{nd}(a,b,c) = c$.

Theorem 5.15 *If* $A \neq 0$, *then* $\bigcap A$ *is a set.*

Proof Choose $a \in A$. Then, by 5.11(*vii*), $\bigcap A \subseteq a$, so that, by 2.2, $\bigcap A$ is a set.

As in the case of unions, a more common notation for intersections will now be introduced.

Definition 5.16 $\bigcap_{i\in I} A_i = \bigcap Rng\ A$.

With regard to this definition remarks apply as to Definition 5.4. As in the case of generalized union, we have commutative and associative laws for generalized intersections; we state them without proof.

Theorem 5.17 (*Commutative law for infinite intersections*) *Let* A *be an indexed family of sets, with* $Dmn\ A = I$. *Let* F *be a function from a class* J *onto* I. *Then*

$$\bigcap_{i\in I} A_i = \bigcap_{j\in J} A_{F(j)}.$$

Theorem 5.18 (*Associative law for infinite intersections*) *Let* A *be an indexed family of sets, with* $Dmn\ A = I \times J$. *Then*

$$\bigcap_{i\in I, j\in J} A_{ij} = \bigcap_{i\in I} \left(\bigcap_{j\in J} A_{ij}\right).$$

The following theorem gives analogs to 3.16(*iv*) and 4.1(*i*).

Theorem 5.19 *Let* R *be a relation, and* A *an indexed family of sets with domain* I. *Then*

$$R^*\left(\bigcap_{i\in I} A_i\right) \subseteq \bigcap_{i\in I} R^*(A_i).$$

If F is a function, then

$$F^{-1*}\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} F^{-1*}(A_i).$$

We now turn to relationships between general unions and general intersections. First, we have De Morgan's laws, generalizing 2.19(ix) and (x).

Theorem 5.20 *Let A be an indexed family of sets with domain I. Then*
(i) $\left(\bigcup_{i\in I} A_i\right)' = \bigcap_{i\in I} A_i'$.
(ii) $\left(\bigcap_{i\in I} A_i\right)' = \bigcup_{i\in I} A_i'$.

Second, we give the general distributive laws.

Theorem 5.21 *Let A be an indexed family of sets with domain $I\times J$, where I and J are sets. Then*
(i) $\bigcap_{i\in I}\bigcup_{j\in J} A_{ij} = \bigcup_{F\in {}^I J}\bigcap_{i\in I} A_{i,F(i)}$.
(ii) $\bigcup_{i\in I}\bigcap_{i\in J} A_{ij} = \bigcap_{F\in {}^I J}\bigcup_{i\in I} A_{i,F(i)}$.

Proof We will prove (*i*) only. First, suppose that $x \in \bigcap_{i\in I}\bigcup_{j\in J} A_{ij}$. Let $R = \{(i,j) : x \in A_{ij}\}$. By the relational axiom of choice, let $F_0 \subseteq R$ be a function with $Dmn\ F_0 = Dmn\ R$. Clearly $Dmn\ R = I$, so that $F_0 \in {}^I J$ (note that F_0 is a set, since R is by virtue of $R \subseteq I\times J$). For any $i \in I$ we have $x \in A_{i,F_0(i)}$. Hence $x \in \bigcup_{F\in {}^I J}\bigcap_{i\in I} A_{i,F(i)}$.

Conversely, suppose that $x \in \bigcup_{F\in {}^I J}\bigcap_{i\in I} A_{i,F(i)}$. Choose $F_0 \in {}^I J$ such that $x \in \bigcap_{i\in I} A_{i,F_0(i)}$. For any $i \in I$ we then have $x \in A_{i,F_0(i)}$, and so $x \in \bigcup_{j\in J} A_{ij}$. Thus $x \in \bigcap_{i\in I}\bigcup_{j\in J} A_{ij}$.

Remark 5.22 The discussion of infinite unions and intersections can also be carried out within the framework of Boolean algebra; see Sikorski 1964, for example. For many mathematical purposes, only countable unions and intersections are important (see Sec. 19 for the definition). For collateral reading we suggest Hausdorff 1914, the introduction of Kuratowski 1966, and Chap. I of Halmos 1950.

EXERCISES

Prove the following statements.

5.23 Let A be a function with domain I and B a function with domain J. Assume that $A_i \subseteq B_j$ for all $i \in I$, $j \in J$. Then $\bigcup_{i\in I} A_i \subseteq \bigcap_{j\in J} B_j$.

5.24 If $\forall x \in A \exists y \in B(x \subseteq y)$, then $\bigcup A \subseteq \bigcup B$.

5.25 If $\forall x \in A \exists y \in B(y \subseteq x)$, then $\bigcap B \subseteq \bigcap A$.

5.26 Let $\langle R_i : i \in I\rangle$ be a system of relations. Then
(*a*) $\left(\bigcup_{i\in I} R_i\right)^* A = \bigcup_{i\in I} (R_i^* A)$.
(*b*) $\left(\bigcap_{i\in I} R_i\right)^* \{x\} = \bigcap_{i\in I} (R_i^* \{x\})$.

6 DIRECT PRODUCTS, POWER CLASSES

Definition 6.1 *Let A be a function with domain I. Let $\mathsf{P}A = \mathsf{P}_{i\in I}A_i = \{f : f \text{ is a function with domain } I, \text{ and } f_i \in A_i \text{ for all } i \in I\}$. $\mathsf{P}A$ is called the* ***direct product*** *of the family A.*

Note that, if I is a proper class, then $\mathsf{P}A = 0$. The notion of a direct product will play a large role in our later development of cardinal arithmetic, and we will prove many of the basic properties of direct products at that time, since then they will appear more natural.

Theorem 6.2 (*i*) $\mathsf{P}0 = \{0\}$.

(*ii*) *If A is a function with domain I such that A is a set (and hence I is a set) and $A_i \neq 0$ for all $i \in I$, then $\mathsf{P}_{i\in I} A_i \neq 0$, and $\mathsf{P}_{i\in I} A_i$ is a set.*

Proof (*i*) is obvious. To prove (*ii*), let $R = \{(i,x) : i \in I \text{ and } x \in A_i\}$. If $i \in I$, then there is an $x \in A_i$, so that $(i,x) \in R$; thus $Dmn\ R = I$. By the relational axiom of choice, let F be a function with domain I such that $F \subseteq R$. Thus $F_i \in A_i$ for all $i \in I$; by 4.15(*v*), F is a set, so that $F \in \mathsf{P}_{i\in I} A_i$. Now $\mathsf{P}_{i\in I} A_i \subseteq {}^I \bigcup Rng\ A$, so that $\mathsf{P}_{i\in I} A_i$ is a set, by 3.14 and 5.3.

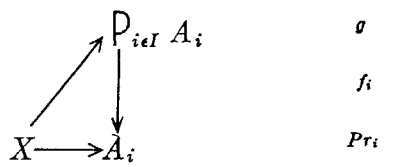
In proving 6.2(*ii*), it is essential to use the relational axiom of choice (see Sec. 16).

We may think of $\mathsf{P}_{i\in I} A_i$ as an I-dimensional space; elements f of $\mathsf{P}_{i\in I} A_i$ are I-tuples, with ith coordinates f_i for $i \in I$. It is natural to consider the operation of projection into the ith-coordinate space A_i.

Definition 6.3 $Pr_i = \langle f_i : f \text{ a function}, i \in Dmn\ f\rangle$. *$Pr_i$ is called the* ***ith projection*** *function.*

Thus for each $i \in V$, Pr_i is a function whose domain is $\{f : f \text{ is a function}, i \in Dmn\ f\}$, and for each f in its domain, $Pr_i f = f_i$.

Theorem 6.4 *Let $A = \langle A_i : i \in I\rangle$ be a family of nonempty sets, I a set. Then $Pr_i{\restriction}\mathsf{P}_{i\in I} A_i$ maps $\mathsf{P}_{i\in I} A_i$ onto A_i, for each $i \in I$. Furthermore, if X is a set and f_i maps X into A_i for each $i \in I$, then there is a unique g, $g : X \to \mathsf{P}_{i\in I} A_i$, such that the diagram*



commutes for each i, i.e., such that $Pr_i \circ g = f_i$ for each i.

Proof We may define g as $g = \langle \langle f_i(x) : i \in I \rangle : x \in X \rangle$; the various assertions of the theorem are then easily checked.

Definition 6.5 $SA = \{B : B \subseteq A\}$. *$SA$ is called the* ***power class*** *of A.*

If A is a proper class, then so is SA (as is easily checked). We are really interested in this notion only when A is a set.

Theorem 6.6 (*i*) $0 \in SA$.
(*ii*) $S0 = \{0\}$.
(*iii*) $A \subseteq B \Rightarrow SA \subseteq SB$.
(*iv*) $S(A \cap B) = SA \cap SB$.
(*v*) $SA \cup SB \subseteq S(A \cup B)$.

Theorem 6.7 *Sa is a set.*

Proof Power-set axiom.

About the most fundamental fact in elementary set theory is the following theorem of Cantor. The reader will recognize another variant of the argument involved in Russell's paradox. The theorem implies that, given any set, there is another set with more elements than it, interpreting "more" rather strictly—the large set cannot even be put in one-one correspondence with the smaller (see Chap. 4).

Theorem 6.8 (*Cantor*) *There does not exist a function mapping a onto Sa.*

Proof Suppose, on the contrary, that F maps a onto Sa. Let $b = \{x : x \in a, x \notin F(x)\}$. Then $b \subseteq a$, so choose $x \in a$ such that $F(x) = b$. Then $x \in b \Rightarrow x \notin F(x) \Rightarrow x \notin b$; $x \notin b \Rightarrow x \notin F(x) \Rightarrow x \in b$. This is a contradiction.

EXERCISES

Prove the following statements.

6.9 If A is a function with domain $\{a,b\}$, $a \neq b$, then there is a function f mapping $\mathsf{P}A$ one-one onto $A_a \times A_b$.
6.10 Suppose that A and B are indexed families of sets with the same nonempty domain I, A and B sets, such that $B_i \subseteq A_i$ for each $i \in I$. Then $\mathsf{P}_{i \in I} B_i = \bigcap_{i \in I} [(Pr_i \restriction \mathsf{P}_{i \in I} A_i)^{-1*} B_i]$.
6.11 If A and B are functions with domain I, I a set, and $f_i : A_i \to B_i$ for each $i \in I$, then there is a unique function g, $g : \mathsf{P}_{i \in I} A_i \to \mathsf{P}_{i \in I} B_i$,

such that for each $i \in I$ the diagram

$$\begin{array}{ccc} \mathsf{P}_{i\in I} A_i & \xrightarrow{g} & \mathsf{P}_{i\in I} B_i \\ \Big\downarrow{\scriptstyle Pr_i} & & \Big\downarrow{\scriptstyle Pr_i} \\ A_i & \xrightarrow{f_i} & B_i \end{array}$$

commutes; i.e., $f_i \circ (Pr_i \restriction \mathsf{P}_{i\in I} A_i) = (Pr_i \restriction \mathsf{P}_{i\in I} B_i) \circ g$.

6.12 $SS0 = \{0,\{0\}\}$.

6.13 $SSS0 = \{0,\{0\},\{\{0\}\},\{0,\{0\}\}\}$.

6.14 Let A be a set, and F a function mapping SA into SA such that for all $B,C \in SA$, $B \cap FC = 0$ iff $FB \cap C = 0$. Let $\langle B_i : i \in I\rangle$ be an indexed family of members of SA, with I a set. Then $F(\bigcap_{i\in I} B_i) = \bigcap_{i\in I} FB_i$.

7 EQUIVALENCE RELATIONS

Among the most important concepts of elementary set theory is that of an equivalence relation. In this section we give the simplest properties of equivalence relations; in particular we prove the important fact that equivalence relations and partitions amount to the same thing.

Equivalence relations arise in the following way. Frequently we work within a given set A in mathematics, and for certain purposes we wish to "identify" various elements of A; we are interested in only certain properties of the members of A and would like to forget that there are distinct elements having the same properties of interest. We define R to be the set of all pairs (x,y) such that we wish to identify x and y. R is then an equivalence relation. If x/R is the set of all y such that xRy, then $\{x/R : x \in A\}$ is a partition of A; it is a set in which the identification has taken place. For a proper class A the class x/R may also be a proper class and hence in general cannot be treated like the set x. Thus, in this case, a more involved identification device is needed (see Sec. 15). We now want to make these considerations precise.

Definition 7.1 (*i*) *R is **transitive** iff R is a relation and* $\forall x,y,z[(x,y) \in R \wedge (y,z) \in R \Rightarrow (x,z) \in R]$.
(*ii*) *R is **symmetric** iff R is a relation and* $\forall x,y[(x,y) \in R \Rightarrow (y,x) \in R]$.
(*iii*) *R is an **equivalence relation** iff R is transitive and symmetric.*
(*iv*) *R is **reflexive on** A iff* $\forall x \in A$, $(x,x) \in R$.

Equivalent formulations of this definition are given in the following.

Theorem 7.2 *Let R be a relation. Then*
(*i*) *R is transitive iff $R|R \subseteq R$.*
(*ii*) *R is symmetric iff $R^{-1} \subseteq R$.*

(*iii*) *R is an equivalence relation iff* $R|R^{-1} = R$.
(*iv*) *R is reflexive on A iff* $I\restriction A \subseteq R$.

Proof (*i*), (*ii*), and (*iv*) are trivial. To prove (*iii*), first suppose that R in an equivalence relation. Then $R|R^{-1} \subseteq R|R \subseteq R$, by (*i*) and (*ii*). On the other hand, if xRy, then $xRyR^{-1}x$, and hence xRx, by what was just shown; hence $xRxR^{-1}y$, by symmetry of R, so that $xR|R^{-1}y$.

Conversely, suppose that $R|R^{-1} = R$. Then $R^{-1} = (R|R^{-1})^{-1} = (R^{-1})^{-1}|R^{-1} = R|R^{-1} = R$, so that R is symmetric. And $R|R = R|R^{-1} = R$, so that R is transitive.

In most treatments of elementary set theory the notion of an equivalence relation is made relative to some class A, but we have defined the notion without doing so. Mostly, however, we discuss equivalence relations on a class. Recalling from 3.9 what it means for a relation to be on a class, we see that R is an equivalence relation on A iff R is transitive, symmetric, and $Fld\ R = A$. By 7.4(*ii*), R is then also reflexive on A, so that the notion of equivalence relation on A has its usual meaning.

We now introduce some standard notation concerning equivalence relations.

Definition 7.3 (*i*) $x/R = \{y : (x,y) \in R\}$.
(*ii*) $A/R = \{a : \exists x \in A, a = x/R\}$.
(*iii*) $\pi_R = \langle x/R : x \in Fld\ R\rangle$.

Definition 7.3 contains a couple of ambiguities we hope do not confuse the reader. $/$ is used in different senses in (*i*) and (*ii*); two different symbols, say $/$ and $/\!\!/$, should have been used. We usually use a lowercase letter, like x, for the first sense and a capital letter for the second. Note also that, if $Fld\ R$ is a proper class, then π_R may also be a proper class. Thus π_R is sometimes a proper class, although π is a lowercase letter; this small conflict with the general convention of 1.3 should not lead to any confusion. We reserve Π for cardinal multiplication (see Sec. 21).

For an equivalence relation R and elements $x \in Fld\ R$, the sets x/R are called *equivalence classes*, and x is a *representative* of x/R. The most useful facts about equivalence relations are summarized in the following.

Theorem 7.4 *Let R be an equivalence relation. Then*
(*i*) $Dmn\ R = Rng\ R = Fld\ R$.
(*ii*) *R is reflexive on $Fld\ R$.*
(*iii*) *For any $x,y \in Fld\ R$, xRy iff $x/R = y/R$.*
(*iv*) *For any $x,y \in Fld\ R$, if $x/R \cap y/R \neq 0$, then $x/R = y/R$.*
(*v*) *For any $x \in Fld\ R$, $x \in x/R$.*

(*vi*) *For any* $x,y \in Fld\ R$, *if* $x \in y/R$, *then* $x/R = y/R$.
(*vii*) *For any* $x,y,z \in Fld\ R$, *if* $x,y \in z/R$, *then* xRy *and* $x/R = y/R$.
(*viii*) *If* R *is a set, then* π_R *maps* $Fld\ R$ *onto* $Fld\ R/R$ *and* $\pi_R(x) = x/R$ *for all* $x \in Fld\ R$.

Proof (*i*) follows by symmetry of R. To prove (*ii*), assume that $x \in Fld\ R$; say, by (*i*), xRy. Then $xRyRx$, by symmetry of R, so that xRx, by transitivity of R.

For (*iii*), let xRy. First suppose that $z \in x/R$. Thus xRz. Hence $yRxRz$, by symmetry of R, and yRz, by transitivity of R, so that $z \in y/R$. Suppose, conversely, that $z \in y/R$. Then $xRyRz$, so that xRz and $z \in x/R$. Conversely, if $x/R = y/R$, then yRy, by (*ii*), implies that $y \in y/R$; hence $y \in x/R$; hence xRy.

As to (*iv*), say that $z \in x/R \cap y/R$. Then xRz and yRz. By symmetry of R, $xRzRy$, so that xRy. Thus, by (*iii*), $x/R = y/R$.

Condition (*v*) follows from (*ii*). Condition (*vi*) is an easy consequence of (*iii*), and (*vii*) follows from (*vi*) and (*iii*). Finally, (*viii*) is obvious.

Thus two equivalence classes coincide if they have a common representative; and x and y are representatives of the same equivalence class iff xRy.

Definition 7.5 *P is a **partition of** A if P is a family of pairwise disjoint nonempty sets and* $\bigcup P = A$.

Recall the definition of *family of pairwise disjoint sets* from 2.9. Thus P is a partition of A iff the following three conditions hold:

(1) $\forall x \in P\ \forall y \in P(x \neq y \Rightarrow x \cap y = 0)$.
(2) $\forall x \in P(x \neq 0)$.
(3) $\bigcup_{x \in P} x = A$.

In the next theorem we show that for any set A there is a natural one-one correspondence between equivalence relations with field A and partitions of A.

Theorem 7.6 *Let A be any set. Let $\boldsymbol{E}(A) = \{R : R$ is an equivalence relation with field $A\}$ and $\boldsymbol{P}(A) = \{P : P$ is a partition of $A\}$. Let*

$$\mathcal{P} = \langle A/R : R \in \boldsymbol{E}(A)\rangle;$$
$$\mathcal{E} = \langle \{(x,y) : \exists M \in P,\ x,y \in M\} : P \in \boldsymbol{P}(A)\rangle.$$

Then
(*i*) *For every* $R \in \boldsymbol{E}(A)$, $\mathcal{P}(R) \in \boldsymbol{P}(A)$.
(*ii*) *For every* $P \in \boldsymbol{P}(A)$, $\mathcal{E}(P) \in \boldsymbol{E}(A)$.

(*iii*) $\mathcal{E} \circ \mathcal{P} = I \restriction E(A)$.
(*iv*) $\mathcal{P} \circ \mathcal{E} = I \restriction P(A)$.
(*v*) $\mathcal{E}$ *maps* $P(A)$ *one-one onto* $E(A)$, *and* $\mathcal{P}$ *maps* $E(A)$ *one-one onto* $P(A)$.

Proof (*i*) $\mathcal{P}(R)$ is a family of pairwise disjoint sets, by 7.4(*iv*); it then follows easily that $\mathcal{P}(R) \in P(A)$.

(*ii*) $\mathcal{E}(P)$ is symmetric: if $x\mathcal{E}(P)y$, then there is an $M \in P$ with $x,y \in M$; hence $y\mathcal{E}(P)x$. $\mathcal{E}(P)$ is transitive: if $x\mathcal{E}(P)y\mathcal{E}(P)z$, then there exist $M,N \in P$ with $x,y \in M$ and $y,z \in N$. Thus $M \cap N \neq 0$, so that $M = N$ and $x,z \in M$, so that $x\mathcal{E}(P)z$. Therefore $\mathcal{E}(P)$ is an equivalence relation. Clearly $Fld\ \mathcal{E}(P) \subseteq \bigcup P = A$. Now suppose that $x \in A$; say $x \in M \in P$. Then $x\mathcal{E}(P)x$. Thus $\mathcal{E}(P) \in E(A)$.

(*iii*) Let $R \in E(A)$. If xRy, then $y \in x/R$, and $x \in x/R$, by 7.4(*ii*), so that $x\mathcal{E}\mathcal{P}(R)y$. Suppose conversely that $x\mathcal{E}\mathcal{P}(R)y$; say $x,y \in z/R$. By 7.4(*vii*), xRy. Thus $R = \mathcal{E}\mathcal{P}(R)$.

(*iv*) Let $P \in P(A)$. If $M \in P$, then it is easily checked that $M = x/E(P)$ for each $x \in M$; thus $M \in \mathcal{P}\mathcal{E}(P)$ since M is nonempty. This shows that $P \subseteq \mathcal{P}\mathcal{E}(P)$. For any $x \in A$ there is an $M \in P$ with $x \in M$, and hence, by what was "easily checked" above, $x/\mathcal{E}(P) = M \in P$. Thus $\mathcal{P}\mathcal{E}(P) \subseteq P$, so that, by what was already proved, $P = \mathcal{P}\mathcal{E}(P)$.

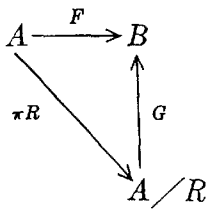
(*v*) Follows from (*iii*) and (*iv*), by 4.14(*iii*).

Remark 7.7 The notion of an equivalence relation is found useful in almost every branch of mathematics. For relatively "deep" facts about equivalence relations we suggest the articles Jónsson 1953 and Janiczak 1952.

EXERCISES

Prove the following statements.

7.8 Let F be a function mapping a set A onto B. Let $R = \{(x,y) : x,y \in A \text{ and } F(x) = F(y)\}$. Then R is an equivalence relation on A. Moreover, there is a one-one function G mapping A/R onto B such that the following diagram commutes:



7.9 Let A be a set, and R an equivalence relation on A. For S an equivalence relation on A with $R \subseteq S$ let $F(S) = \{(x,y) : \exists a,b \in A$ such that $x = a/R$, $y = b/R$, and $aSb\}$. Then F is a function mapping

$\{S : S$ is an equivalence relation on A, $R \subseteq S\}$ one-one onto $\{T : T$ is an equivalence relation on $A/R\}$.

7.10 Let R be an equivalence relation on A.

(*a*) If $B \subseteq A$, then $R \cap (B \times B)$ is an equivalence relation on B.

(*b*) If F maps B into A, then $\{(x,y) : x,y \in B$ and $(F(x),F(y)) \in R\}$ is an equivalence relation on B.

(*c*) If S is an equivalence relation on B, then $\{(x,y) : \exists a,b \in A, \exists c,d \in B, x = (a,c), y = (b,d), aRb$ and $cSd\}$ is an equivalence relation on $A \times B$.

7.11 If $A \neq 0$ and each member of A is an equivalence relation on a set B, then $\bigcap A$ is also an equivalence relation on B.

7.12 A union of equivalence relations is not necessarily an equivalence relation. A union of a set of equivalence relations with pairwise disjoint fields is an equivalence relation.

7.13 Let R and S be equivalence relations on a set A, and $\mathcal{P}(R)$ and $\mathcal{P}(S)$, respectively, the associated partitions of A. Then $R \subseteq S$ iff every $M \in \mathcal{P}(R)$ is included in some $N \in \mathcal{P}(S)$.

8 ORDERING

In this section we want to discuss at a very elementary level the notion of ordering, generalizing the mathematically familiar idea of the ordering of the integers or of the real numbers. We begin with the weakest, most general notion of ordering—partial ordering—and proceed to the most restrictive notion—that of a well-ordering.

Definition 8.1 (*i*) *R is **antisymmetric** iff R is a relation and for all x, y, if $xRyRx$, then $x = y$.*

(*ii*) *R is a **partial ordering** iff R is a relation, R is reflexive on Fld R, R is transitive—by Definition 7.1(i)—and R is antisymmetric.*

(*iii*) *A **is partially ordered by** R iff $(A \times A) \cap R$ is a partial ordering with field A.*

Frequently we will use a symbol suggesting ordering, such as $\leq$ or $\preceq$, instead of the letter R, in discussing partial orderings. We will then always use the symbol without the line, such as $<$ or $\prec$, to denote the relation

$$\{(x,y) : xRy \wedge x \neq y\}.$$

Clearly the relation $<$ is transitive and irreflexive: $x \not< x$ for all x. [Recall from 3.3(*ii*) that $x \not< x$ simply means $\neg(x < x)$.] Actually transitive and irreflexive relations stand in a very close relationship with

partial orderings. Given any set A, one can assign to each partial ordering $\leq$ with field A the transitive and irreflexive relation $F(\leq) = <$ with field $\subseteq A$; and given any transitive and irreflexive relation R with field $\subseteq A$, one can assign the relation $G(R) = R \cup (I \restriction A)$. $G(R)$ is then a partial ordering with field A. Furthermore, $F \circ G$ is the identity on $\{R : R \text{ is transitive and irreflexive, and } Fld\ R \subseteq A\}$, and $G \circ F$ is the identity on $\{R : R \text{ is a partial ordering with field } A\}$. Thus, according to 4.14, both F and G are one-one and onto; $F = \langle < : \leq \text{ a partial ordering on } A\rangle$ establishes a one-one correspondence between partial orderings on A and transitive and irreflexive relations with field $\subseteq A$.

Partial orderings can intuitively be represented by diagrams consisting of nodes and lines arranged in levels, as in Fig. 11. The points represent elements of the field of the partial ordering, and the lines indicate the ordering itself. If we denote the ordering by $\preceq$, we thus have $l \preceq a$, $c \preceq b \preceq a$, $f \preceq d$, $h \preceq e$, $f \preceq b$, and $k \preceq k$, for example, but $b \not\preceq c$, $f \not\preceq g$, $h \not\preceq k$, etc.

The conditions defining a partial ordering can be more concisely stated thus: $I \restriction Fld\ R \subseteq R$, $R \restriction R \subseteq R$, and $R \cap R^{-1} \subseteq I$. These conditions always hold upon replacing R by R^{-1}, so we see that R^{-1} is a partial ordering whenever R is a partial ordering. With regard to Definition 8.1(*iii*), note that there exist A, R such that R is not a partial ordering, although $(A \times A) \cap R$ is; by our definition, A is still partially ordered *by* R. This is the case for $A = 0$, no matter what R is. For a less trivial example, let $R = \{(x,y) : x \in y \text{ or } x = y\}$, and let $A =$

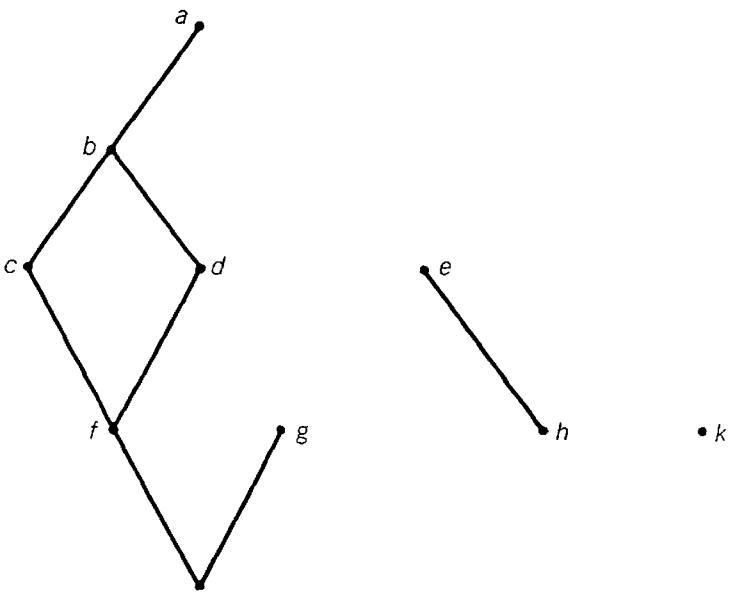


Figure 11

$\{0,\{0\},\{0,\{0\}\}\}$. (R is not transitive since $0 \in \{0\} \in \{\{0\}\}$, although $0 \notin \{\{0\}\}$.)

Theorem 8.2 *If R is a partial ordering, then also $(A \times A) \cap R$ is a partial ordering.*

The most typical example of a partial ordering is *inclusion*. To indicate precisely the sense in which inclusion is typical, we need the notion of isomorphism of relations.

Definition 8.3 *F is an **isomorphism from** R **onto** S iff R and S are relations, F is a one-one function mapping Fld R onto Fld S, and $\forall a,b \in$ Fld $R[aRb$ iff $(Fa)S(Fb)]$.*

Theorem 8.4 *(i) $\{(x,y) : x \subseteq y\}$ is a partial ordering.*
(ii) If $R \in V$ is any partial ordering, then there is a set A and an isomorphism F from R onto $(A \times A) \cap \{(x,y) : x \subseteq y\}$.

Proof *(i) Trivial.*

(ii) We first define F. For any $x \in$ Fld R, let $Fx = \{y : yRx\}$. Let $A =$ Rng F. Thus F maps Fld R onto Fld S, where $S = (A \times A) \cap \{(x,y) : x \subseteq y\}$. If $x,y \in$ Fld R and $x \neq y$, then, by antisymmetry, $x \not\subseteq y$ or $y \not\subseteq x$, and hence $x \in Fx \sim Fy$ or $y \in Fy \sim Fx$, so that $Fx \neq Fy$. Thus F is one-one. If $x,y \in$ Fld R and xRy, then $(Fx)S(Fy)$. Indeed, this means that $Fx \subseteq Fy$, and for any $z \in Fx$ we have zRx, which, combined with the hypothesis xRy, gives zRy, and so $z \in Fy$. If $x,y \in$ Fld R and $x \not R y$, then $x \in Fx \sim Fy$, and so $(Fx)\not S(Fy)$. This completes the proof.

If we let $R = \{(x,y) : x \subseteq y\}$, then any class A is partially ordered by R; we may say that A *is partially ordered by inclusion*. By abuse of notation, we may even say that $\subseteq$ itself is a partial ordering, although this is not strictly correct.

A method of obtaining partial orderings that is frequently used is given in the following.

Theorem 8.5 *Let R be a transitive relation with field a set A, and suppose that R is reflexive on A. Let $S = \{(x,y) : xRy$ and $yRx\}$. Then S is an equivalence relation with field A. Further, let $\preceq = \{(a,b) : \exists x,y \in A\ (a = x/S \wedge b = y/S \wedge xRy)\}$. Then $\preceq$ is a partial ordering with field A/S.*

Proof If $xSySz$, then xRy, yRx, yRz, and zRy; hence, by the transitivity of R, xRz and zRx, so that xSz. Thus S is transitive. Clearly S is symmetric, so that S is an equivalence relation; also, its field is clearly A (since R is reflexive on A).

Next, suppose that $a \preceq b \preceq c$. Then there exist $w,x,y,z \in A$ such that $a = w/S$, $b = x/S$, wRx, $b = y/S$, $c = z/S$, and yRz. Since $x/S = y/S$, we have xSy, and hence xRy. Thus $wRxRyRz$, so that, by the transitivity of R, wRz. Therefore, $a \preceq c$. Hence $\preceq$ is transitive. Using the notation already established, if $c = a$, then $w/S = z/S$, from which it follows that wSz and zRw. By the above $wRxRy$ and $yRzRw$; transitivity of R gives wRy and yRw, i.e., wSy. Thus $a = w/S = y/S = b$. This shows that $\preceq$ is antisymmetric. Clearly $Fld\,(\preceq) \subseteq A/S$. Now suppose that $x \in A$. Then xRx, so that $x/S \preceq x/S$. Hence $A/S \subseteq Fld\,(\preceq)$, and $\preceq$ is reflexive on A/S. Therefore $\preceq$ is a partial ordering with field A/S, as desired.

We now introduce some special concepts associated with partial orderings.

Definition 8.6 *Let $\preceq$ be a partial ordering with field A, and suppose that $X \subseteq A$ and $a \in A$.*

(*i*) *a is a $\preceq$-**upper bound of** X if $x \preceq a$ for all $x \in X$.*

(*ii*) *a is a $\preceq$-**lower bound of** X if $a \preceq x$ for all $x \in X$.*

(*iii*) *a is a $\preceq$-**greatest element of** X if a is a $\preceq$-upper bound of X and $a \in X$.*

(*iv*) *a is a $\preceq$-**least element of** X if a is a $\preceq$-lower bound of X and $a \in X$.*

(*v*) *a is a $\preceq$-**least upper bound of** X, or for brevity a $\preceq$-l.u.b. of X, if a is a $\preceq$-upper bound of X and a is a $\preceq$-lower bound of the class of all $\preceq$ upper bounds of X.*

(*vi*) *a is a $\preceq$-**greatest lower bound of** X, or for brevity a $\preceq$-g.l.b. of X, if a is a $\preceq$-lower bound of X and a is a $\preceq$-upper bound of the class of all $\preceq$-lower bounds of X.*

(*vii*) *a is a $\preceq$-**minimal element of** X if $a \in X$ and $x \not\prec a$ for all $x \in X$.*

(*viii*) *a is a $\preceq$-**maximal element of** X if $a \in X$ and $a \not\prec x$ for all $x \in X$.*

There are some immediate properties of these concepts we state informally. Every element of A is a $\preceq$-upper bound of 0 and a $\preceq$-lower bound of 0. There are sets without $\preceq$-upper bounds or $\preceq$-lower bounds, for some partial orderings $\preceq$. Such examples are not easily available at the present state of our formal development of set theory, but the ordering of the rationals may be cited informally. There is at most one $\preceq$-greatest element of X, but X may have many $\preceq$-maximal elements; similarly with least elements and minimal elements. There is at most one $\preceq$-l.u.b., and at most one $\preceq$-g.l.b. of a set X. a is a $\preceq$-l.u.b of 0 iff a is a $\preceq$-least element of A, and a is a $\preceq$-g.l.b. of 0 iff a is a $\preceq$-greatest element of A; a is a $\preceq$-l.u.b. of A iff a is a $\preceq$-greatest element of A, and a is a $\preceq$-g.l.b. of A iff a is a $\preceq$-least element of A.

The following *fixed-point theorem* has many applications.

Theorem 8.7 *Let $\preceq$ be a partial ordering with field A, and suppose that every subclass $B \subseteq A$ has a $\preceq$-l.u.b. Suppose that F maps A into A in such a way that for all $x,y \in A$, $x \preceq y$ implies that $Fx \preceq Fy$. Then $Fx = x$ for some $x \in A$.*

Proof Let a be the $\preceq$-l.u.b. of $B = \{x : x \in A, x \preceq Fx\}$. For any $x \in B$ we have both $x \preceq Fx$ and $x \preceq a$, and hence both $x \preceq Fx$ and $Fx \preceq Fa$, so that $x \preceq Fa$. Thus Fa is a $\preceq$-upper bound for B, so that $a \preceq Fa$. Hence $Fa \preceq FFa$, so that $Fa \in B$. From this it follows that $Fa \preceq a$, and hence $Fa = a$.

Definition 8.8 *Let $\preceq$ be a partial ordering. A is **directed** by $\preceq$ if $A \subseteq Fld\ (\preceq)$, and for all $a,b \in A$ there is a $c \in A$ such that $a \preceq c$ and $b \preceq c$.*

Theorem 8.9 *(i) If A is a class of functions directed by inclusion, then $\bigcup A$ is a function.*
(ii) If A is a class of partial orderings directed by inclusion, then $\bigcup A$ is a partial ordering.

Proof (*i*) Clearly $\bigcup A$ is a relation. Suppose that $x(\bigcup A)y$ and $x(\bigcup A)z$. Say $(x,y) \in f \in A$ and $(x,z) \in g \in A$. Since A is directed by inclusion, choose $h \in A$ such that $f \subseteq h$ and $g \subseteq h$. Then xhy and xhz, so that $y = z$, as desired.

(*ii*) Again, it is clear that $\bigcup A$ is a relation. Suppose now that $x \in Fld\ (\bigcup A)$. Then xRy or yRx for some y and some $R \in A$, i.e., $x \in Fld\ R$ for some $R \in A$, hence xRx for some $R \in A$, so that $x(\bigcup A)x$. Therefore, $\bigcup A$ is reflexive on $Fld\ (\bigcup A)$. Next, suppose that $x(\bigcup A)y(\bigcup A)z$; say $xRySz$ with $R,S \in A$. Since A is directed by inclusion, there is a $T \in A$ such that $R \subseteq T$ and $S \subseteq T$. Thus $xTyTz$, so that xTz and hence $x(\bigcup A)z$. Therefore, $\bigcup A$ is transitive. Finally, suppose that $x(\bigcup A)y(\bigcup A)x$; say $xRySx$ with $R,S \in A$. As before, we find a $T \in A$ such that $xTyTx$, and hence $x = y$. Therefore, $\bigcup A$ is antisymmetric and so is a partial ordering.

One can similarly show that $\bigcup A$ is an equivalence relation if A is a class of equivalence relations directed by inclusion.

Definition 8.10 *(i) $\preceq$ is a **simple ordering** (or a **linear ordering**) iff $\preceq$ is a partial ordering and for all $x,y \in Fld\ (\preceq)$, $x \preceq y$ or $y \preceq x$.*
*(ii) A is **simply ordered by** R iff $(A \times A) \cap R$ is a simple ordering.*

With regard to 8.10(*ii*) remarks apply as to 8.1(*iii*). The notion of isomorphism is frequently applied with respect to simply orderings. If

we call two simple orderings $\leq$ and $\preceq$ *isomorphic* if there is an isomorphism between them, and let $\equiv$ consist of all isomorphic pairs $(\leq, \preceq)$ with $\leq$ and $\preceq$ *sets*, then it is easy to check that $\equiv$ is an equivalence relation. For each simple ordering $R \in V$, $R/\equiv$ is, however, a proper class, as is easily seen. Thus the situation discussed at the beginning of Sec. 7 has arisen; as we mentioned there, we will give a method for "identifying" isomorphic simple orderings in Sec. 15.

Theorem 8.11 *Let $\leq$ and $\preceq$ be simple orderings, and suppose that F is a mapping from $Fld\ (\leq)$ onto $Fld\ (\preceq)$ such that for all $a,b \in Fld\ (\leq)$, if $a < b$, then $Fa \prec Fb$. Then F is an isomorphism from $\leq$ onto $\preceq$.*

Proof If $a,b \in Fld\ (\leq)$ and $a \neq b$, then either $a < b$ or $b < a$, and hence either $Fa \prec Fb$ or $Fb \prec Fa$; in particular, $Fa \neq Fb$. Thus F is one-one. Next, given $a,b \in Fld\ (\leq)$, we know, by hypothesis, that $a < b$ implies $Fa \prec Fb$. If $a \not< b$, then $b \leq a$, and hence, by hypothesis, $Fb \preceq Fa$, and so $Fa \not\prec Fb$, by antisymmetry. Thus $a \leq b$ iff $Fa \preceq Fb$, as desired.

Definition 8.12 *(i) R is **well-founded** iff R is a relation and for every nonempty class $A \subseteq Fld\ R$ there is an $x \in A$ such that $A \cap \{y : yRx\} = 0$.*
*(ii) $\preceq$ is a **well-ordering** iff $\preceq$ is a simple ordering and $\prec$ is well-founded (cf. the remark following 8.1).*

Note the similarity of the definition of a well-founded relation with our formulation of the regularity axiom. In fact, the regularity axiom really says simply that the relation $\{(x,y) : x \in y\}$ is well-founded.

The ordinary ordering of the positive integers furnishes a simple example of a well-ordering. Well-orderings will be considered in great detail in Chap. 2, since they are intimately related to the notion of an ordinal number. Here we content ourselves with exhibiting two equivalent definitions of a well-ordering.

Theorem 8.13 *For any partial ordering $\preceq$, the following three conditions are equivalent:*
(i) $\preceq$ is a well-ordering.
(ii) $\preceq$ is a simple ordering, and every nonempty class $A \subseteq Fld\ (\preceq)$ has a $\preceq$-least elements.
(iii) every nonempty class $A \subseteq Fld\ (\preceq)$ has a $\preceq$-least element.

Proof (i) $\Rightarrow$ (ii) We need to show only that an arbitrary nonempty class $A \subseteq Fld\ (\preceq)$ has a $\preceq$-least element. By (i), choose $x \in A$ such that $A \cap \{y : y \prec x\} = 0$. For any $y \in A$ we then have $y \not\prec x$, and so, $\preceq$ being a simple ordering, $x \preceq y$. Thus x is the desired $\preceq$-least element of A.

(*ii*) $\Rightarrow$ (*iii*) This is obvious.

(*iii*) $\Rightarrow$ (*i*) We first have to show that $\leq$ is a simple ordering. Given two elements x, y of $Fld\ (\leq)$, let z be a $\leq$-least element of $\{x,y\}$. This means that $z = x$ or $z = y$, and hence $x \leq y$ or $y \leq x$. Thus, indeed, $\leq$ is a simple ordering. Next, let A be a nonempty subclass of $Fld\ (\leq)$. By (*iii*), let x be a $\leq$-least element of A. Thus $x \leq y$, and hence $y \not< x$, by antisymmetry, for each $y \in A$. This means that $A \cap \{y : y < x\} = 0$. Thus $<$ is well-founded, and (*i*) holds.

Remark 8.14 Much of the more advanced theory of ordering is due to Hausdorff; his basic book, Hausdorff 1914, is still very readable.

EXERCISES

8.15 How many partial orderings with field $\{a,b,c,d\}$ are there, where a,b,c,d are all distinct?

8.16 Show that, if A is a set of well-orderings directed by inclusion, then $\bigcup A$ is not necessarily a well-ordering.

8.17 Exhibit a well-founded relation R that is a proper class and is not a partial ordering.

8.18 Let $R = \{(m,n) : m,n$ are positive integers and m divides $n\}$. Show that R is a partial ordering such that any two positive integers have an R-*l.u.b.* and an R-*g.l.b.*

8.19 Let R be a simple ordering with field A, and S a simple ordering with field B. Let $T = \{(x,y) :$ there exist $a,a' \in A$ and $b,b' \in B$ such that $x = (a,b)$, $y = (a',b')$, and either $a \neq a'$ and aRa', or else $a = a'$ and $bSb'\}$. Show that T is a simple ordering with field $A \times B$ and that, if R and S are well-orderings, then so is T.

2
Ordinals

The more advanced aspect of set theory begins with this chapter. Ordinals are certain classes which are associated with well-orderings; they yield typical examples of well-orderings. Here we first give the basic properties of ordinals and then proceed to the discussion of transfinite induction. Special ordinals, the natural numbers, are then discussed. The later development is facilitated by the description of normal functions, which is the next topic. We then rigorously justify recursive definitions and follow this with an exposition of the elements of ordinal arithmetic. The chapter closes with a survey of a few more advanced topics.

9 ORDINALS: BASIC PROPERTIES

Definition 9.1 (*i*) *A is* ϵ-***transitive*** *iff for all* x *and* y, $x \in y \in A \Rightarrow x \in A$.
(*ii*) *A is an* ***ordinal*** *iff A is* ϵ*-transitive and each member of A is* ϵ*-transitive.*
(*iii*) $Ord = \{x : x \text{ is an ordinal}\}$.

This simple definition of an ordinal will gain meaning after we have developed the properties of ordinals extensively; then we shall be able to appreciate the elegance of the definition.

We may anticipate the rigorous development by an intuitive comment. Ordinals constitute an extension into the infinite of the order properties of the natural number sequence

$$0, 1, 2, \ldots .$$

In fact, we simply affix a new number ω that comes after all the natural numbers and begin again, counting from ω:

$$0, 1, 2, \ldots , \omega, \omega \dotplus 1, \omega \dotplus 2, \ldots .$$

Furthermore, there is no reason to stop at any point, and we continue indefinitely:

$$0, 1, 2, \ldots , \omega, \omega \dotplus 1, \omega \dotplus 2, \ldots , \omega \cdot 2, \omega \cdot 2 \dotplus 1, \omega \cdot 2 + 2, \ldots .$$

The main use of ordinals in mathematics is in labeling sets so as to systematize constructions or proofs. This is one of the uses of natural numbers, as in proof by induction. The other main use of natural numbers—indicating magnitude—is extended to the transfinite in the theory of cardinals, which will be treated in Chap. 4.

More formally, for any $A \in Ord$, $\{(x,y) : x,y \in A \text{ and } (x = y \text{ or } x \in y)\}$ turns out to be a well-ordering, distinct ordinals give rise to nonisomorphic well-orderings, and every well-ordering $R \in V$ is isomorphic to some well-ordering derived as indicated from an ordinal (cf. the discussion following 8.10). These statements will all be proved in the course of this chapter (see Theorems 13.10 and 13.11).

It may not be obvious that $\in$-transitive classes, much less ordinals, even exist. We first give some theorems showing that there is a profusion of ordinals.

Theorem 9.2 $0 \in Ord$.

Proof 0 is a set, by the axiom of infinity; 0 is $\in$-transitive, by a vacuous implication; and every member of 0 is $\in$-transitive, by a vacuous implication. Thus $0 \in Ord$.

Theorem 9.3 *If* $x \in Ord$, *then* $\mathcal{S}x \in Ord$.

Proof Assume that $x \in Ord$. By 2.13(*xii*), $\mathcal{S}x$ is a set, and $\mathcal{S}x = x \cup \{x\}$. If $y \in z \in \mathcal{S}x$, then either $z \in x$, $y \in z \in x$, $y \in x$ (since x is $\in$-transitive, in virtue of being an ordinal), and $y \in \mathcal{S}x$, or $z = x$, $y \in x$, $y \in \mathcal{S}x$; thus $y \in \mathcal{S}x$

in either case. Hence $\mathcal{S}x$ is ϵ-transitive. If $y \in \mathcal{S}x$, then either $y \in x$ or $y = x$; x being an ordinal, y is ϵ-transitive.

By 9.2 and 9.3, 0, $\mathcal{S}0$, $\mathcal{S}\mathcal{S}0$, . . . are all members of *Ord*. The following theorem then enables us to obtain even more members of *Ord*.

Theorem 9.4 *If $A \subseteq Ord$, then $\bigcup A$ is an ordinal.*

Proof To show that $\bigcup A$ is ϵ-transitive, assume that $x \in y \in \bigcup A$. Thus there is a $z \in A$ such $y \in z$. Since $A \subseteq Ord$, we see that $z \in Ord$, and so z is ϵ-transitive; but $x \in y \in z$, so it follows that $x \in z$. Now $z \in A$, so that, finally, $x \in \bigcup A$. Thus, indeed, $\bigcup A$ is ϵ-transitive.

Now take any element w of $\bigcup A$. Then there is a $v \in A$ such that $w \in v$. Now $A \subseteq Ord$, so that $v \in Ord$. By 9.1 w, as a member of v, is ϵ-transitive. Hence any member of $\bigcup A$ is ϵ-transitive.

This completes the proof.

Now it may appear that almost all sets are ordinals. This is not true. For example, $\{\{0\}\}$ is not even ϵ-transitive, since $0 \in \{0\} \in \{\{0\}\}$ but $0 \notin \{\{0\}\}$. The set $a = \{0,\{0\},\{\{0\}\}\}$ is ϵ-transitive, but its member $\{\{0\}\}$ is not, so that a is not an ordinal. A further example, which is sometimes useful in applications of set theory, is the following.

Theorem 9.5 $(a,b) \notin Ord$.

Proof Recall that $(a,b) = \{\{a\},\{a,b\}\}$. Thus $a \in \{a\} \in (a,b)$. However, $a \notin (a,b)$, so that (a,b) is not even ϵ-transitive. Suppose, on the contrary, that $a \in (a,b)$. Then either $a = \{a\}$ or $a = \{a,b\}$, but

$a = \{a\}$ implies that $a \in a$,
$a = \{a,b\}$ implies that $a \in a$,

contradicting 1.19(i), in each case.

We now give some useful properties of ordinals, leading up to the well-ordering property.

Theorem 9.6 *If A is an ordinal, then $A \subseteq Ord$.*

Proof Assume that A is an ordinal and $x \in A$. Then, by Definition 9.1, x is ϵ-transitive. If y is any member of x, then, since A is ϵ-transitive, $y \in A$; then, again by 9.1, y is ϵ-transitive. Thus $x \in Ord$.

Combining 9.5 and 9.6, we see that no nonempty relation is an ordinal, and no nonempty function is an ordinal.

Theorem 9.7 *Ord is an ordinal.*

Proof Obviously every member of *Ord*, by virtue of being an ordinal, is ϵ-transitive. All we need to show is that *Ord* itself is ϵ-transitive. Assume that $x \in y \in Ord$. Thus y is an ordinal, so that, by 9.6, $x \in Ord$, as desired.

Theorem 9.8 *Ord is not a set.*

Proof If *Ord* were a set, by Theorem 9.7, we should have $Ord \in Ord$, which contradicts Theorem 1.19(i).

Theorems 9.7 and 9.8 are related to the *Burali-Forti paradox*, a contradiction that arises in nonaxiomatic set theory. In axiomatic set theory these theorems do not lead to any obvious contradictions; they simply constitute minor facts about the class *Ord*.

The next theorem about ordinals has the only difficult proof in our development of the basics; the theorem is essential in discussing ordering below.

Theorem 9.9 *If $x,y \in Ord$, then $x = y$ or $x \in y$ or $y \in x$.*

Proof Let $A = \{x : x \in Ord \wedge (\exists y)\ (y \in Ord \wedge x \neq y \wedge x \notin y \wedge y \notin x)\}$. Thus we want to show that $A = 0$. Assume, on the contrary, that $A \neq 0$. By the regularity axiom, choose $a \in A$ such that $a \cap A = 0$. Since $a \in A$, the class $B = \{y : y \in Ord \wedge a \neq y \wedge a \notin y \wedge y \notin a\}$ is nonempty. By the axiom of regularity again, choose $b \in B$ such that $b \cap B = 0$. Note that $a \notin B$, so that $a \neq b$.

We reach a contradiction by proving that $a = b$. We first show that $a \subseteq b$. Let $z \in a$. Now $a \in A$ implies that a is an ordinal, so that, by Theorem 9.6, z is also an ordinal. Since $a \cap A = 0$, z is not a member of A, so that $\forall y\ (y \in Ord \Rightarrow z = y \vee z \in y \vee y \in z)$; in particular, $z = b \vee z \in b \vee b \in z$. Now

$$z = b \Rightarrow b \in a \Rightarrow b \notin B,$$

a contradiction. Also, using the fact that a is ϵ-transitive,

$$b \in z \Rightarrow b \in a \Rightarrow b \notin B,$$

again a contradiction. The only remaining possibility is that $z \in b$. Since z is arbitrary, $a \subseteq b$.

To prove that $b \subseteq a$, suppose, conversely, that $z \in b$. Thus $z \notin B$, since $b \cap B = 0$, while $z \in Ord$, by 9.6, since $b \in B \subseteq Ord$. It follows that $a = z \vee a \in z \vee z \in a$. We have

$$a = z \Rightarrow a \in b \Rightarrow b \notin B,$$

a contradiction, and, using the fact that b is ϵ-transitive,

$$a \in z \Rightarrow a \in b \Rightarrow b \notin B,$$

the same contradiction. The only remaining possibility is that $z \in a$. Thus $b \subseteq a$, so that, by the preceding paragraph, $a = b$. This is the contradiction that establishes the theorem.

The next theorem essentially gives the least-element principle in the well-ordering of ordinals.

Theorem 9.10 *If A is a nonempty class of ordinals, then $\bigcap A$ is an ordinal, and in fact $\bigcap A \in A$.*

Proof If $x \in \bigcap A$, then $x \in y \in A$ for some ordinal y (since $A \neq 0$), so that x is ϵ-transitive. Assume that $x \in y \in \bigcap A$. For any $z \in A$ we have $x \in y \in z$, so that $x \in z$, since z is ϵ-transitive. Hence $x \in \bigcap A$. Thus $\bigcap A$ is ϵ-transitive, so that $\bigcap A$ is an ordinal.

For the last part of the theorem, to get a contradiction, assume that $\bigcap A \notin A$. For any $y \in A$ we have $\bigcap A \subseteq y$, and hence $y \notin \bigcap A$, by 1.19(i). Since $y, \bigcap A \in Ord$, it follows from 9.9 that $\bigcap A \in y$. y being arbitrary, we get $\bigcap A \in \bigcap A$, which contradicts 1.19(i).

We now state some minor set-theoretical properties of ordinals that will be found useful later.

Theorem 9.11 *Let A and B be ordinals. Then*
(i) $A \in Ord$ *or* $A = Ord$.
(ii) $A \in B$ *iff* $A \subset B$.
(iii) *If* $A \in B$, *then* $\mathcal{S}A = B$ *or* $\mathcal{S}A \in B$.
(iv) *If* $C \subseteq A$, *then* $\bigcup C = A$ *or* $\bigcup C \in A$.
(v) *If* $A \in Ord$, *then* $\bigcup \mathcal{S}A = A$.
(vi) $A = \mathcal{S}\bigcup A$ *or* $A = \bigcup A$.

Proof (i) Suppose that $A \notin Ord$; thus A is an ordinal, but it is a proper class. By 9.6, $A \subseteq Ord$. Assume now that $Ord \sim A \neq 0$; say $x \in Ord \sim A$. Then for any $y \in A$ we have $y \in x$, by 9.9 (since $x \in y$ implies that $x \in A$), so that $A \subseteq x$ and A is a set—a contradiction. Hence $Ord \sim A = 0$ and $A = Ord$.

(ii) $\Rightarrow$. By ϵ-transitivity of B, $A \subseteq B$. Since $A \in B \sim A$, $A \neq B$. $\Leftarrow$. $B \in A$ and $B = A$ are ruled out, by 1.19(i) and the definition of proper inclusion. Thus if B is a set, $A \in B$ follows, by 9.9 (since A is then also a set by virtue of $A \subset B$). If B is not a set, $B = Ord$, by (i), and $A \in Ord$, by (i), since $A \subset B$ precludes the possibility that $A = Ord$.

(iii) Assume that $\mathcal{S}A \neq B$. By 9.3, $\mathcal{S}A \in Ord$. We may assume

that B is a set; otherwise $\mathcal{S}A \in B$, by (i). Thus by 9.9, $\mathcal{S}A \in B$, since $B \in \mathcal{S}A$ is ruled out, by 1.19, and we are assuming that $\mathcal{S}A \neq B$.

(iv) We know that $\bigcup C$ is an ordinal, by 9.4. If $A = Ord$, the desired conclusion follows from (i), applied to $\bigcup C$. If $A \neq Ord$, then $A \in Ord$, by (i), and the desired conclusion follows from 9.9, by virtue of the fact that $A \in \bigcup C$ easily implies that $A \in \bigcup A$, which contradicts 1.19 again.

(v) Assume first that $x \in \bigcup \mathcal{S}A$; say $x \in y \in \mathcal{S}A$ for a certain y. If $y \in A$, then $x \in A$, by the $\in$-transitivity of A, while $y = A$ logically implies that $x \in A$. Thus $x \in A$, so that $\bigcup \mathcal{S}A \subseteq A$. If $x \in A$, then the fact that $x \in A \in \mathcal{S}A$ implies that $x \in \bigcup \mathcal{S}A$. Hence $\bigcup \mathcal{S}A = A$, as desired.

Finally, as to (vi), condition (iv) gives the two cases $\bigcup A = A$ and $\bigcup A \in A$. The first case gives a desired conclusion. The second case gives the following possibilities, by virtue of (iii): $\mathcal{S} \bigcup A = A$ or $\mathcal{S} \bigcup A \in A$. The last possibility, however, is in fact impossible; indeed, $\mathcal{S} \bigcup A \in A$ implies that $\bigcup A \in \mathcal{S} \bigcup A \in A$ and hence $\bigcup A \in \bigcup A$. This completes the proof.

Definition 9.12 (i) *Lowercase Greek letters* $\alpha, \beta, \gamma, \ldots$ *are used to denote ordinals* $\in Ord$ *unless otherwise indicated.*
(ii) $\leq = \{(x,y) : x,y \in Ord \text{ and } (x \in y \text{ or } x = y)\}$.

By the convention following 8.1, we have

$$< = \{(x,y) : x,y \in Ord,\ x \in y\}.$$

Thus $\alpha \in \beta$ is equivalent to $\alpha < \beta$; this will be tacitly assumed in much of what follows. The symbols $\leq$ and $<$ will be reserved from now on for their use with ordinals unless stated to the contrary.

We will feel free to define classes of ordinals by expressions like $\{\alpha : \varphi(\alpha)\}$—*the class of all ordinals* α *such that* $\varphi(\alpha)$—and relations by expressions like $\{(\alpha,\beta) : \varphi(\alpha,\beta)\}$. This extends the convention of Definition 1.9.

In the last theorem of this section we gather together all of the order properties of ordinals that will be used in the future.

Theorem 9.13 $\leq$ *is a well-ordering with field Ord. Furthermore:*
(i) *For any nonempty set* $A \subseteq Ord$, $\bigcap A$ *is the* $\leq$*-least element of* A.
(ii) 0 *is the* $\leq$*-least element of Ord.*
(iii) *For any* α, $\alpha = \{\beta : \beta < \alpha\}$.
(iv) $\beta < \mathcal{S}\alpha$ *iff* $\beta \leq \alpha$.
(v) *For any* α, *there is no* β *such that* $\alpha < \beta < \mathcal{S}\alpha$.
(vi) $\alpha \leq \beta$ *iff* $\alpha \subseteq \beta$.

(*vii*) *For any set A of ordinals, $\bigcup A$ is the $\leq$-least upper bound of A.*
(*viii*) $\alpha < \beta$ *iff* $\S\alpha \leq \beta$.
(*ix*) $\S\alpha = \S\beta$ *iff* $\alpha = \beta$.
(*x*) $\alpha < \beta$ *iff* $\S\alpha < \S\beta$.

Proof From 1.19 we easily infer that $\leq$ is antisymmetric; it is transitive, by the ϵ-transitivity of ordinals, and it is obviously reflexive on *Ord*, which is its field. Thus $\leq$ is a partial ordering with field *Ord*. Hence if we establish (*i*), by 8.13 it will also follow that $\leq$ is a well-ordering. Suppose, then, that $0 \neq A \subseteq Ord$. By 9.10, $\bigcap A \in A$. If $y \in A$ and $y \in \bigcap A$, then $y \in y$ since $\bigcap A \subseteq y$, contradicting 1.19. Thus, by 9.9, $\bigcap A \leq y$ for all $y \in A$, as desired.

(*ii*) is an immediate consequence of (*i*), and (*iii*) follows from the remarks after 9.12. (*iv*) is obvious from the definitions involved. If $\alpha < \beta < \S\alpha$, then $\alpha < \beta \leq \alpha$, by (*iv*), and so $\alpha < \alpha$, by transitivity, which is impossible; thus (*v*) holds. (*vi*) is immediate from 9.11(*ii*). To prove (*vii*), first note that if $x \in A$, then $x \subseteq \bigcup A$, and hence further, by (*vi*) and 9.4, $x \leq \bigcup A$. Thus $\bigcup A$ is a $\leq$-upper bound for A. Suppose that α is any $\leq$-upper bound for A. Thus $\forall\beta\ (\beta \in A \Rightarrow \beta \leq \alpha)$, so that since α is ϵ-transitive, $\forall\beta\ \forall\gamma\ (\beta \in \gamma \in A \Rightarrow \beta < \alpha)$. Hence $\bigcup A \subseteq \alpha$, so that, by (*vi*), $\bigcup A \leq \alpha$. This proves (*vii*).

If $\alpha < \beta$, then $\S\alpha \subseteq \beta$, by the ϵ-transitivity of β, so that $\S\alpha \leq \beta$, by (*vi*); conversely, $\S\alpha \leq \beta \Rightarrow \alpha < \S\alpha \leq \beta \Rightarrow \alpha < \beta$. Thus (*viii*) holds. To prove (*ix*), note that $\alpha < \S\alpha = \S\beta$, and hence $\alpha \leq \beta$, by (*iv*); by symmetry, $\beta \leq \alpha$, so that $\alpha = \beta$. Finally, $\alpha < \beta$ iff $\S\alpha \leq \beta$, by (*viii*); iff $\S\alpha < \S\beta$, by (*iv*). And so (*x*) holds.

Remark 9.14 Two standard references for detailed information on ordinal and cardinal numbers are Bachmann 1967 and Sierpinski 1965.

EXERCISES

9.15 Show that for any set x the following statements are equivalent:
(*a*) x is an ordinal.
(*b*) x is ϵ-transitive, and for all $y,z \in x$, either $y \in z$, $y = z$, or $z \in y$.
(*c*) x is ϵ-transitive, and for all y, if $y \subset x$ and y is ϵ-transitive, then $y \in x$.
(*d*) $x = 0$ or $0 \in x$; for all $y \in x$, either $\S y = x$ or $\S y \in x$; for all $y \subseteq x$, either $\bigcap y = x$ or $\bigcap y \in x$.
(*e*) $R = \{(y,z) : y \in x, z \in x, \text{ and } y \in z \text{ or } y = z\}$ is a well-ordering with field x such that for all $y \in x$, $y = \{z : zRy, z \neq y\}$.

Hint: It is easily seen that (*a*) $\Rightarrow$ (*b*), (*a*) $\Rightarrow$ (*c*), (*a*) $\Rightarrow$ (*d*), (*a*) $\Rightarrow$ (*e*),

$(b) \Rightarrow (a)$, and $(e) \Rightarrow (a)$. To show that $(c) \Rightarrow (a)$, let x satisfy (c) and define $y = \{z : z \in x,\ z \in Ord\}$, and show that $y = x$. To show that $(d) \Rightarrow (a)$, let α be the least element of $Ord \sim x$, where x satisfies (d), and examine $\bigcup\alpha$.

9.16 If f maps α onto $\bigcup A$, then the function $g = \langle f\,(\bigcap f^{-1*}x) : x \in A\rangle$ has the property that $gx \in x$ for all $x \in A$.

10 TRANSFINITE INDUCTION

Theorem 10.1 *(First principle of transfinite induction) If A is an ordinal, and B is a class such that*
(*) for every $\alpha \in A$, if $\beta \in B$ for every $\beta < \alpha$, then $\alpha \in B$, and $A \subseteq B$.

Proof Suppose, on the contrary, that $A \not\subseteq B$; then $A \sim B \neq 0$, and hence it has a least element α (by 9.13). Thus $\alpha \in A$, and for every $\beta < \alpha$, $\beta \in A$ (since A is $\in$-transitive), and hence, since $\beta \notin A \sim B$, also $\beta \in B$. Thus by (*), $\alpha \in B$, which is a contradiction.

Theorem 10.1 is a generalization of the complete-induction principle for natural numbers (see Sec. 11). To illustrate the use of the theorem, we will prove the following statement.

1 *If F is a function mapping an ordinal A into Ord and if for all $\alpha,\beta \in A$ the condition $\alpha < \beta$ implies that $F\alpha < F\beta$, then $\alpha \leq F\alpha$ for each $\alpha \in A$.*

To prove this, let $B = \{\alpha : \alpha \in A, \alpha' \leq F\alpha\}$. Suppose that $\alpha \in A$ and $\beta \in B$ for every $\beta < \alpha$. Thus for any $\beta < \alpha$ we have $\beta \leq F\beta < F\alpha$, so that, by the $\in$-transitivity of $F\alpha$, $\beta < F\alpha$. Hence $\alpha \subseteq F\alpha$, so that, by 9.13*(vi)*, $\alpha \leq F\alpha$; that is, $\alpha \in B$. Hence, by 10.1, $A = B$, as desired.

Usually, rather than explicitly defining B when applying 10.1, we simply assume a statement $\varphi(\beta)$ true for all $\beta < \alpha$ and prove it true for α. In our example, for $\varphi(\beta)$ we would take the statement $\beta \in A \wedge \beta \leq F\beta$. A similar remark applies to the other induction principles that we will introduce.

The form of transfinite induction most similar to ordinary induction on integers is a step-by-step process formulated in terms of successor ordinals and limit ordinals. We now discuss these latter notions.

Definition 10.2 *(i) α is a **successor ordinal** if $\alpha = S\beta$ for some β. One then writes $\beta = \alpha - 1$; $\alpha - 1 = \alpha$ if α is not a successor ordinal.*
*(ii) α is a **limit ordinal** if $\alpha \neq 0$ and α is not a successor ordinal.*

Theorem 10.3 (*i*) α *is a successor ordinal iff* $\bigcup\alpha < \alpha$.
(*ii*) α *is a limit ordinal iff* $\bigcup\alpha = \alpha \neq 0$.
(*iii*) α *is a limit ordinal iff* $\forall\beta[\beta < \alpha \Rightarrow \exists\gamma(\beta < \gamma < \alpha)] \wedge \alpha \neq 0$.

Proof (*i*) $\Rightarrow$. Say $\alpha = \S\beta$. Then $\bigcup\alpha = \bigcup\S\beta = \beta < \alpha$, by 9.11(*v*). $\Leftarrow$. Assuming $\bigcup\alpha < \alpha$, we then have $\S\bigcup\alpha \leq \alpha$, by 9.13(*viii*). If $\S\bigcup\alpha < \alpha$, then $\bigcup\alpha < \S\bigcup\alpha < \alpha$ and hence $\bigcup\alpha \in \bigcup\alpha$, contradicting 1.19(*i*). Thus $\S\bigcup\alpha = \alpha$, as desired.

To prove (*ii*), note that $\bigcup\alpha \leq \alpha$, by 9.11(*iv*), and apply (*i*). The right-hand condition of (*iii*) is equivalent to $\alpha \subseteq \bigcup\alpha \wedge \alpha \neq 0$; by 9.11(*iv*), it is equivalent to $\alpha = \bigcup\alpha \neq 0$, so that (*iii*) follows from (*ii*).

In the intuitive description of ordinals given at the beginning of Sec. 9, successor ordinals are ordinals immediately following other ordinals; examples are 1, $\omega + 2$, $\omega\cdot 2 + 1$. Limit ordinals are ordinals coming after an ellipsis, like ω and $\omega\cdot 2$.

Theorem 10.4 (*Second principle of transfinite induction*) *Assume that* A *is an ordinal, and* B *is a class such that*
(*i*) $0 \in B$.
(*ii*) *For all* $\alpha \in A$, *if* $\alpha \in B$ *and* $\S\alpha \in A$, *then* $\S\alpha \in B$.
(*iii*) *For all* $\alpha \in A$, *if* α *is a limit ordinal and* $\beta \in B$ *for each* $\beta < \alpha$, *then* $\alpha \in B$.
Then $A \subseteq B$.

Proof We will apply the first principle of transfinite induction. To verify 10.1(*), suppose that $\alpha \in A$ and $\beta \in B$ for every $\beta < \alpha$. Definition 10.2 then gives three cases.
Case 1 $\alpha = 0$. Then by (*i*), $\alpha \in B$.
Case 2 α is a successor ordinal, say $\alpha = \S\beta$. Then $\beta < \alpha$, so that $\beta \in B$, by assumption. By (*ii*), $\alpha = \S\beta \in B$.
Case 3 α is a limit ordinal. Then $\alpha \in B$, by (*iii*).
Thus $\alpha \in \beta$ in any case. Hence 10.1(*) holds, so that, by 10.1, $A \subseteq B$.

Note that, if A is a limit ordinal or if $A = Ord$, then the hypothesis $\S\alpha \in A$ may be omitted from 10.4(*ii*). If $A = Ord$, then the hypothesis $\alpha \in A$ may be omitted from both (*ii*) and (*iii*). Now we will illustrate the use of 10.4 by proving the following statement.

2 *Let* F *be a function with domain Ord satisfying the following conditions:*
(*a*) $F0 = 0$.
(*b*) $F\S\alpha = SF\alpha$ (*recall Definition 6.5*).
(*c*) *if* α *is a limit ordinal, then* $F\alpha = \bigcup_{\beta<\alpha} F\beta$.
Then for every α *and every* $x \in F\alpha$, $x \subseteq F\alpha$.

We prove this "by induction on α," using 10.4. Let $B = \{\alpha : \text{for every } x \in F\alpha \text{ we have } x \subseteq F\alpha\}$. Since $F0 = 0$, vacuously $0 \in B$. Assume that $\alpha \in B$; to show that $S\alpha \in B$, assume that $x \in FS\alpha$ and $y \in x$. By 2(b), $FS\alpha = SF\alpha$, so that $x \in SF\alpha$; i.e., $x \subseteq F\alpha$. Therefore, $y \in F\alpha$. Since $\alpha \in B$, it follows that $y \subseteq F\alpha$, so that $y \in SF\alpha = FS\alpha$. Since y is arbitrary, $x \subseteq FS\alpha$. This shows that $S\alpha \in B$. Finally, suppose that α is a limit ordinal, $\beta \in B$ for every $\beta < \alpha$, and $x \in F\alpha$. By 2(c), choose $\beta < \alpha$ such that $x \in F\beta$. Since $\beta \in B$, it follows that $x \subseteq F\beta$, so that, by 2(c) again, $x \subseteq F\alpha$, as desired. The conclusion of 2 now follows from Theorem 10.4.

EXERCISES

10.5 Let F be as in Item 2. Show:

(a) $F\alpha \subset FS\alpha$ for every α.

(b) If $\alpha < \beta$, then $F\alpha \subset F\beta$.

(c) If $\alpha < \beta$, then $F\alpha \in F\beta$.

10.6 Suppose that $\mu, \nu \in {}^{\alpha}Ord$ and the following conditions hold:

(a) $\alpha > 0$, and $\mu 0 = \nu 0 = 0$.

(b) For any β with $S(\beta) < \alpha$, $\mu S(\beta) = S(S(\mu\beta))$ and $\nu S(\beta) = S(S(S(\nu\beta)))$.

(c) For any limit ordinal $\beta < \alpha$, $\mu\beta = \bigcup_{\beta<\gamma} \mu\gamma$ and $\nu\beta = \bigcup_{\gamma<\beta} \nu\gamma$.

Show that for any $\beta < \alpha$, $\nu\beta > \mu\beta$ if β is a successor ordinal, although $\nu\beta = \mu\beta = \beta$ otherwise.

10.7 Suppose that μ maps $Ord \times Ord$ into Ord and satisfies the following conditions (for any α, β):

(a) $\mu(\alpha,0) = \alpha$.

(b) $\mu(\alpha,S(\beta)) = S(\mu(\alpha,\beta))$.

(c) $\mu(\alpha,\beta) = \bigcup_{\gamma<\beta} \mu(\alpha,\gamma)$ if $\beta = \bigcup\beta \neq 0$ (i.e., if β is a limit ordinal).

Show that $\mu(\mu(\alpha,\beta),\gamma) = \mu(\alpha,\mu(\beta,\gamma))$ for all α, β, γ. *Hint:* Use induction on γ.

10.8 Assume that $\alpha < \beta$ and that B is a class such that for all γ, if $\alpha \leq \gamma < \beta$ and if $\delta \in B$ for all δ such that $\alpha \leq \delta < \gamma$, then $\gamma \in B$. Show that $\{\gamma : \alpha \leq \gamma < \beta\} \subseteq B$.

10.9 Exercise 10.8 is an induction principle "from α to β" analogous to 10.1. Formulate and prove an induction principle from α to β analogous to 10.4.

10.10 Suppose that A is a set, α is an ordinal, $x \in {}^{\alpha}A$, and $y \in {}^{\alpha}A$. Assume that

$$y_\beta = x_\beta \sim \bigcup_{\gamma<\beta} x_\gamma$$

for every $\beta < \alpha$. Prove that $\bigcup_{\gamma<\beta} y_\gamma = \bigcup_{\gamma<\beta} x_\gamma$ for every $\beta < \alpha$.

11 THE NATURAL NUMBERS

Definition 11.1 (*i*) *Let*

$$\omega = \bigcap\{A : 0 \in A, \text{ and } \forall x(x \in A \Rightarrow \mathcal{S}x \in A)\}.$$

Members of ω *are called* ***natural numbers.***

(*ii*) *The letters* i, j, k, l, m, n, q, p *are used for natural numbers unless otherwise stated.*

(*iii*) $1 = \mathcal{S}0$, $2 = \mathcal{S}1$, . . . , $9 = \mathcal{S}8$, $10 = \mathcal{S}9$.

As for ordinals, we can define classes of natural numbers by expressions like $\{m : \varphi(m)\}$.

Theorem 11.2 (*i*) ω *is a set.*

(*ii*) $0 \in \omega$, $1 \in \omega$, . . . , $9 \in \omega$, $10 \in \omega$.

(*iii*) *If* $x \in \omega$, *then* $\mathcal{S}x \in \omega$.

Theorem 11.2 is immediate from 11.1; for (*i*) we use the infinity axiom. ω is an ordinal; in fact,

Theorem 11.3 ω *is the smallest limit ordinal.*

Proof Let $A = \{x : x \text{ is an ordinal} \wedge x \in \omega \wedge x \subseteq \omega\}$. By 11.2(*ii*), $0 \in A$. If $x \in A$, then $\mathcal{S}x$ is an ordinal (9.3), $\mathcal{S}x \in \omega$ [11.2(*iii*)], and $\mathcal{S}x = x \cup \{x\} \subseteq \omega$; thus, $\mathcal{S}x \in A$. Hence, by 11.1, $\omega \subseteq A$. This implies that every member of ω is $\in$-transitive, in virtue of being an ordinal, and also that ω is $\in$-transitive; i.e., ω is an ordinal. Comparing 11.2(*iii*) with 10.3(*iii*), we see that ω is a limit ordinal. If α is any limit ordinal, then $0 \in \alpha$ and $\forall x(x \in \alpha \Rightarrow \mathcal{S}x \in \alpha)$. Hence, by 11.1, $\omega \subseteq \alpha$, so that, by 9.13(*vi*), $\omega \leq \alpha$.

From 11.3 we see a characteristic property of natural numbers among ordinals: α is a natural number iff α is not a limit ordinal, and no $\beta < \alpha$ is a limit ordinal. In particular, for any $m \neq 0$ there is an n such that $m = \mathcal{S}n$.

Specializing 10.1 and 10.4 to $A = \omega$, we obtain the usual induction principles for natural numbers.

Theorem 11.4 (*Complete induction principle*) *If* B *is a class such that* $m \in B$ *whenever* $n \in B$ *for all* $n < m$, *then* $\omega \subseteq B$.

Theorem 11.5 (*Ordinary induction principle*) *If* B *is a class such that* $0 \in B$, *and for all* m, $m \in B \Rightarrow \mathcal{S}m \in B$, *then* $\omega \subseteq B$.

Now that the set ω is available, we can formulate an important consequence of the regularity axiom that generalizes Theorem 1.19.

Theorem 11.6 *There does not exist a function f with domain ω such that $f\S i \in fi$ for every $i \in \omega$.*

Proof Assume that there is such a function f. By the axiom of regularity, choose $x \in Rng\ f$ such that $x \cap Rng\ f = 0$. Then $x = fi$ for a certain $i \in \omega$. By assumption, $f\S(i) \in x \cap Rng\ f$, which is a contradiction.

Note that the situation described in 11.6 may informally be indicated thus: $\cdots \in f3 \in f2 \in f1 \in f0$.

In the remainder of this section we discuss some topics that may be omitted without loss of continuity. The first notion is that of the *transitive closure* of a relation.

Definition 11.7 *For any relation R let*

$$TR = \{(x,y) : x,y \in Fld\ R \wedge (\exists m \in \omega \sim 1)(\exists f \in {}^{\S m}Fld\ R \text{ such that } f0 = x \wedge fm = y \wedge \forall i < m[(fi, f\S i) \in R])\}$$

TR is called the ***transitive closure*** *of R.*

Thus $x(TR)y$ iff $x = (f0)R(f1)R \cdots R(f(m-1))Rfm = y$ for some finite sequence $f0, \ldots, fm$.

Theorem 11.8 *Let R be any relation.*
(*i*) $R \subseteq TR$.
(*ii*) $Fld\ R = Fld\ (TR)$.
(*iii*) *TR is transitive.*
(*iv*) *If S is a transitive relation and $R \subseteq S$, then $TR \subseteq S$.*
(*v*) *If R* is transitive, then $R = TR$.

Proof (*i*) Assume that xRy. Let $f \in {}^2Fld\ R$ be such that $f0 = x$ and $f1 = y$. The conditions of 11.7 are met, so that $x(TR)y$.

(*ii*) Obviously $Fld\ (TR) \subseteq Fld\ R$, and the reverse inclusion follows from (*i*).

(*iii*) Assume that $x(TR)y$. We now show, by induction on n, that

(1) For every $f \in {}^{\S n}Fld\ R$, if $n > 0$, $f0 = y$, $fn = z$, and $\forall i < n[(fi, f\S i) \in R]$, then $x(TR)z$.

First we apply 11.7 to obtain $m \in \omega \sim 1$ and $g \in {}^{\S m}Fld\ R$ such that $g0 = x$, $gm = y$, and $\forall i < m[(gi, g\S i) \in R]$. For $n = 0$, (1) vacuously holds. Assume (1) true for n. Suppose that $f \in {}^{\S\S n}Fld\ R$, $f0 = y$, $f\S n = z$, and $\forall i < \S n[(fi, f\S i) \in R]$. If $n = 0$, then we merely have yRz, and setting

$h = g \cup \{(\S m,z)\}$, we see by 11.7 that $x(TR)z$. If $n > 0$, then, by (1), for n, $x(TR)(fn)$. Thus by 11.7 there is a $p \in \omega \sim 1$ and an $h \in {}^{\S p}Fld\ R$ such that $h0 = x$, $hp = fn$, and $\forall i < p[(hi,h\S i) \in R]$. Letting $k = h \cup \{(\S p,z)\}$, we again see by 11.7 that $x(TR)z$. This completes the inductive proof of (1). From (1), condition (*iii*) of the theorem follows in an obvious manner.

To prove (*iv*), suppose that S is a transitive relation and $R \subseteq S$. To show that $TR \subseteq S$, it is clearly enough to show by induction on m that

(2) For every $f \in {}^{\S m}Fld\ R$, if $m > 0$, $f0 = x$, $fm = y$, and $\forall i < m[(fi,f\S i) \in R]$, then xSy.

We omit the straightforward proof of (2). Finally, (*v*) follows directly from (*iv*) and (*i*).

Theorem 11.9 *If A is a family of equivalence relations with field a set B, then $T(\bigcup A)$ is an equivalence relation with field B, and*

$$T(\bigcup A) = \bigcap\{S : \bigcup A \subseteq S \text{ and } S \text{ is an equivalence relation with field } B\}.$$

Proof The following statement is easily shown by induction:

(1) $\forall m \in \omega \sim 1\ \forall f \in {}^{\S m}B(\forall i < m[(fi,f\S i) \in \bigcup A] \Rightarrow (fm)[T(\bigcup A)](f0))$.

Thus $T(\bigcup A)$ is symmetric and hence is an equivalence relation, by 11.8(*iii*). Clearly $Fld\ T(\bigcup A) = B$. By 11.8(*iv*), we have

$T(\bigcup A) \subseteq \bigcap\{S : \bigcup A \subseteq S$ and S is an equivalence relation with field $B\}$.

The converse inclusion follows from the fact that $T(\bigcup A)$ itself satisfies the condition in braces.

The final theorem of this section finds many important applications. We do not state the most general result of this type, since the proof of the result is so simple that the reader can easily verify other generalizations (see, e.g., Exercise 11.14).

Theorem 11.10 *Let A be a set and $R \in V$ a relation. Then the following two conditions are equivalent, for any x:*

(*i*) $x \in \bigcap\{C : A \subseteq C \text{ and } R^*C \subseteq C\}$.

(*ii*) *There exist $m \in \omega \sim 1$ and $f \in {}^mV$ such that $f(m - 1) = x$, $f0 \in A$, and for each $i \in m \sim 1$ there is a $j < i$ such that $(fj)R(fi)$.*

Proof (*i*) $\Rightarrow$ (*ii*) Let C be the set of all x such that there exist m and f

as indicated in (*ii*). The implication (*i*) $\Rightarrow$ (*ii*) follows as soon as we show that $A \subseteq C$ and $R^*C \subseteq C$. First, suppose that $x \in A$. Then with $m = 1$ and $f = \{(0,x)\}$ we see that $x \in C$. Thus $A \subseteq C$. Next, suppose that $x \in C$, say via m and f, and that xRy. Then $y \in C$ via $\mathcal{S}m$ and $f \cup \{(m,y)\}$. Thus $R^*C \subseteq C$, as desired.

(*ii*) $\Rightarrow$ (*i*) It is easily seen by induction that, if m and f satisfy the conditions of (*ii*), then $Rng\, f \subseteq \bigcap\{C : A \subseteq C \text{ and } R^*C \subseteq C\}$.

EXERCISES

11.11 Suppose that f is a function mapping $\omega \times \omega$ into ω such that the following conditions hold for all m, n:

(*a*) $f(0,n) = \mathcal{S}(n)$.
(*b*) $f(\mathcal{S}(m),0) = f(m,1)$.
(*c*) $f(\mathcal{S}(m),\mathcal{S}(n)) = f(m,f(\mathcal{S}(m),n))$.

Show that these conditions then follow for all m, n:

(*a*) $n < f(m,n)$.
(*b*) $f(m,n) < f(m,\mathcal{S}(n))$.
(*c*) $f(m,\mathcal{S}(n)) \leq f(\mathcal{S}(m),n)$.
(*d*) $f(m,n) < f(\mathcal{S}(m),n)$.

11.12 Show that the following conditions are equivalent:

(*a*) x is a natural number.
(*b*) For all $y \in \mathcal{S}x$, $y = 0$ or $y = \mathcal{S} \cup y$.

11.13 Show that, if R is a relation, S is a function with domain ω, $S_0 = R$, and $S_{\mathcal{S}m} = R|S_m$ for every $m \in \omega$, then $\boldsymbol{T}R = \bigcup_{m \in \omega} S_m$.

11.14 Suppose that A, B, and M are sets, $B \subseteq A$, and suppose that for every $R \in M$ there is an $m \in \omega$ such that $R \subseteq {}^{\mathcal{S}m}A$. Let

$$C = \bigcap\{D : B \subseteq D \subseteq A, \text{ and for every } R \in M, f \in R, \text{ and } m \in \omega, \text{ if } R \subseteq {}^{\mathcal{S}m}A \text{ and } Rng\,(f|m) \subseteq D, \text{ then } fm \in D\}.$$

Show that for all $a \in A$, $a \in C$ iff there is a $p \in \omega$ and a $g \in {}^{\mathcal{S}p}A$ such that $gp = a$ and for every $i < \mathcal{S}p$ one of the following two conditions holds:

(*a*) $gi \in B$.
(*b*) There exist R, m, j, f such that $R \in M$, $R \subseteq {}^{\mathcal{S}m}A$, $f \in R$, $j \in {}^{m}i$, $f|m = g \circ j$, and $gi = fm$.

11.15 If R is a well-founded relation, and $S = \{(x,y) : x,y \in Fld\ R \text{ and } (x = y \text{ or } xRy)\}$, then $\boldsymbol{T}S$ is a well-founded partial ordering.

12 SEQUENCES AND NORMAL FUNCTIONS

In this section we discuss notions quite analogous to the continuous functions of a real variable, which play an important role in analysis. These notions will be important in the arithmetic of ordinal numbers.

Definition 12.1 *Let A be an ordinal (thus $A = Ord$ or $A \in Ord$).*

(*i*) *A function f with domain A is called an A-**termed sequence** and will sometimes be denoted by $\langle f_\xi \rangle_{\xi \in A}$, $\langle f_\xi \rangle_{\xi < A}$ (if $A \in Ord$), $\langle f_0, \ldots, f_\xi, \ldots \rangle_{\xi \in A}$, or $\langle f_0, \ldots, f_\xi, \ldots \rangle_{\xi < A}$ (if $A \in Ord$). If $A \in \omega$, f will be denoted by $\langle f_0, \ldots, f_{A-1} \rangle$. f_ξ will be called the **ξth term** of f, for $\xi \in A$. For an A-termed sequence of ordinals (i.e., with range $\subseteq Ord$) the letter μ, ν, or ρ is used.*

For the remainder of this definition let μ be an A-termed sequence of ordinals.

(*ii*) *μ is **nondecreasing** if $\forall \alpha\, \forall \beta (\alpha < \beta \in A \Rightarrow \mu\alpha \leq \mu\beta)$.*

(*iii*) *μ is **strictly increasing** if $\forall \alpha\, \forall \beta (\alpha < \beta \in A \Rightarrow \mu\alpha < \mu\beta)$.*

(*iv*) *μ is **limiting** if $\forall \alpha\, (0 \neq \alpha = \bigcup \alpha \in A \Rightarrow \mu\alpha = \bigcup_{\beta < \alpha} \mu\beta)$.*

(*v*) *μ is **half-normal** if μ is limiting and nondecreasing.*

(*vi*) *μ is **normal** if μ is limiting and strictly increasing.*

Note that the two-termed sequence $\langle a,b \rangle$ is different from the ordered pair (a,b). ${}^2A$ is the set of all two-termed sequences of elements of A, while $A \times A$ is the set of all ordered pairs of elements of A. Usually, however, it does not lead to confusion if (a,b) and $\langle a,b \rangle$ are identified. Note that $\langle a,b \rangle = \langle c,d \rangle$ implies that $a = c$ and $b = d$, since $a = (\langle a,b \rangle)(0) = (\langle c,d \rangle)(0) = c$ and $b = (\langle a,b \rangle)(1) = (\langle c,d \rangle)(1) = d$.

If $\mu\alpha = 0$ for all $\alpha \in A$, then μ is nondecreasing, limiting, and half-normal, but not strictly increasing or normal. If $\mu\alpha = \alpha$ for all $\alpha \in A$, then μ satisfies all the conditions 12.1(*ii*) to (*vi*). If μ is nondecreasing, then $\mu 0$ is the smallest element of the range of μ. Note, with regard to 12.1(*iv*), that $0 \neq \alpha = \bigcup \alpha$ simply means that α is a limit ordinal.

We want to establish various simple properties of these notions. Recalling item 1 from Sec. 10, we have the following.

Theorem 12.2 *Let A be an ordinal. If μ is a strictly increasing A-termed sequence of ordinals, then $\alpha \leq \mu\alpha$ for every $\alpha \in A$.*

In 13.9 we shall see that every *Ord*-termed normal function μ has a fixed point, i.e., that there is an ordinal α such that $\alpha = \mu\alpha$.

Theorem 12.3 *If μ is a strictly increasing α-termed sequence of ordinals and $Rng\ \mu \subseteq \beta$, then $\alpha \leq \beta$.*

Proof For all $\gamma < \alpha$ we have $\gamma \leq \mu\gamma \in \beta$, and hence $\gamma < \beta$. Thus $\alpha \subseteq \beta$, so that $\alpha \leq \beta$, by 9.13(*vi*).

Theorem 12.4 *Let A be an ordinal. If μ is a strictly increasing A-termed sequence of ordinals, and $Rng\ \mu = A$, then $\mu = I\restriction A$.*

Proof Suppose $\mu\alpha = \alpha$ for every $\alpha < \beta$, where $\beta \in A$. Choose $\gamma \in A$ such that $\mu\gamma = \beta$. By 12.2, $\gamma \leq \mu\gamma = \beta$. $\gamma < \beta$ implies that $\mu\gamma = \gamma < \beta$, a contradiction, so that $\gamma = \beta$ and $\mu\beta = \beta$. This completes the inductive proof.

We now prove some useful sufficient conditions for half-normality and normality.

Theorem 12.5 *Let A be an ordinal, and let μ be a limiting A-termed sequence of ordinals such that $\mu\beta \leq \mu\mathcal{S}\beta$ whenever $\mathcal{S}\beta \in A$. Then μ is half-normal.*

Proof By induction on γ, using 10.4, we prove

(1) $\forall\gamma\,\forall\beta(\beta < \gamma \in A \Rightarrow \mu\beta \leq \mu\gamma)$.

The case $\gamma = 0$ is trivial, by a vacuous implication. Assume that $\forall\beta(\beta < \gamma \in A \Rightarrow \mu\beta \leq \mu\gamma)$, and also assume that $\beta < \mathcal{S}\gamma \in A$. Then either $\beta < \gamma$ and so $\mu\beta \leq \mu\gamma \leq \mu\mathcal{S}\gamma$, by the assumption of the theorem, or $\beta = \gamma$ and so $\mu\beta \leq \mu\mathcal{S}\gamma$, again by the assumption of the theorem. Thus under the assumption that $\forall\beta(\beta < \gamma \in A \Rightarrow \mu\beta \leq \mu\gamma)$ and $\beta < \mathcal{S}\gamma \in A$ we have shown that $\mu\beta \leq \mu\mathcal{S}\gamma$.

Finally, suppose that γ is a limit ordinal, $\gamma \in A$, and that for every $\delta < \gamma$, and every β, $\beta < \delta \in A$ implies that $\mu\beta \leq \mu\delta$. Suppose that $\beta < \gamma$. Then $\beta < \delta$ for some $\delta < \gamma$, and hence $\mu\beta \leq \mu\delta$. Now $\mu\gamma = \bigcup_{\varepsilon<\gamma} \mu\varepsilon \supseteq \mu\delta$, so that $\mu\delta \leq \mu\gamma$. Hence $\mu\beta \leq \mu\gamma$. This completes the inductive proof.

By an analogous argument, we have the following.

Theorem 12.6 *Let A be an ordinal, and let μ be a limiting A-termed sequence of ordinals such that $\mu\beta < \mu\mathcal{S}\beta$ whenever $\mathcal{S}\beta \in A$. Then μ is normal.*

A particular consequence of 12.6 is that the function $\langle \mathcal{S}m : m \in \omega \rangle$ is normal and hence one-one.

Theorem 12.7 *Let A be an ordinal, μ an A-termed normal function, and let $\alpha \in A$ be a limit ordinal. Then $\mu\alpha$ is also a limit ordinal.*

Proof We apply 10.3(*iii*). To this end, suppose that $\beta < \mu\alpha$. Now, by 12.1, $\mu\alpha = \bigcup_{\gamma<\alpha} \mu\gamma$, so that there is a $\gamma < \alpha$ such that $\beta < \mu\gamma$. But also $\mu\gamma < \mu\alpha$ since μ is strictly increasing. Thus $\exists\delta\,(\beta < \delta < \mu\alpha)$. Also, $\mu\alpha \neq 0$, since $\mu 0 < \mu\alpha$. Hence, by 10.3(*iii*), $\mu\alpha$ is a limit ordinal.

Theorem 12.8 *Let A be an ordinal, μ a nondecreasing A-termed sequence of ordinals, and suppose that $\beta < \alpha \in A$. Then $\bigcup_{\gamma<\alpha} \mu\gamma = \bigcup_{\beta\leq\gamma<\alpha} \mu\gamma$.*

Proof If $\delta < \beta$, then $\mu\delta \leq \mu\beta$ and hence $\mu\delta \subseteq \mu\beta$. Thus $\bigcup_{\gamma<\alpha} \mu\gamma \subseteq \bigcup_{\beta\leq\gamma<\alpha} \mu\gamma \subseteq \bigcup_{\gamma<\alpha} \mu\gamma$, so that the desired equality follows.

For any class of functions, it is interesting to determine whether the class is closed under composition or not. We now consider this question with regard to the kinds of functions introduced in 12.1.

Theorem 12.9 *Let A, B, C be ordinals, and let μ, ν be nondecreasing functions mapping A into B and B into C respectively. Then $\nu\circ\mu$ is a nondecreasing function.*

Proof Clearly $\nu\circ\mu$ is an A-termed sequence of ordinals. Suppose that $\alpha < \beta \in A$. Then $\mu\alpha \leq \mu\beta$ and hence $\nu\mu\alpha \leq \nu\mu\beta$.

Similarly, one shows the following.

Theorem 12.10 *Let A, B, C be ordinals, and let μ, ν be strictly increasing functions mapping A into B and B into C respectively. Then $\nu\circ\mu$ is a strictly increasing function.*

Theorem 12.11 *Let A, B, C be ordinals, and let μ, ν be half-normal functions mapping A into B and B into C respectively. Then $\nu\circ\mu$ is half-normal.*

Proof By 12.9, we know that $\nu\circ\mu$ is nondecreasing, so that we have simply to check that $\nu\circ\mu$ is limiting. Suppose that α is a limit ordinal and $\alpha \in A$. We distinguish two cases.
Case 1 There is a $\beta < \alpha$ such that $\mu\beta = \mu\gamma$ for all γ such that $\beta < \gamma < \alpha$. Then, by 12.8, $\mu\alpha = \bigcup_{\gamma<\alpha} \mu\gamma = \bigcup_{\beta\leq\gamma<\alpha} \mu\gamma = \mu\beta$, and so, by the assumption of this case, $\mu\gamma = \mu\alpha$ for every γ for which $\beta \leq \gamma < \alpha$. Consequently,

$$\begin{aligned} \bigcup_{\gamma<\alpha} \nu\mu\gamma &= \bigcup_{\beta\leq\gamma<\alpha} \nu\mu\gamma && \text{since } \nu\circ\mu \text{ is nondecreasing, using 12.8;} \\ &= \nu\mu\alpha && \text{by what was just noted.} \end{aligned}$$

Thus, in this case, $\nu\circ\mu$ is limiting.
Case 2 For every $\beta < \alpha$ there is a γ such that $\beta < \gamma < \alpha$ and $\mu\beta < \mu\gamma$. Then we have

(1) $\forall\beta < \mu\alpha\, \exists\gamma < \alpha(\beta < \mu\gamma < \mu\alpha)$; in particular, $\mu\alpha$ is a limit ordinal.

Indeed, suppose that $\beta < \mu\alpha$. Now, μ being limiting, $\mu\alpha = \bigcup_{\gamma<\alpha} \mu\gamma$. Hence choose $\gamma < \alpha$ such that $\beta < \mu\gamma$. By the assumption of this case, there is a δ such that $\gamma < \delta < \alpha$ and $\mu\gamma < \mu\delta$. Thus $\mu\gamma < \mu\alpha$. This establishes the first part of (1). Furthermore, $\mu\alpha \neq 0$ under the assumption of this case, and $\forall\beta < \mu\alpha\, \exists\varepsilon(\beta < \varepsilon < \mu\alpha)$, so that, by 10.3(*iii*), $\mu\alpha$ is a limit ordinal.

From (1) we obtain, since ν is limiting,

(2) $\quad \nu\mu\alpha = \bigcup_{\beta<\mu\alpha} \nu\beta$

If $\beta < \mu\alpha$, by (1), choose $\gamma < \alpha$ such that $\beta < \mu\gamma < \mu\alpha$; then $\nu\beta \leq \nu\mu\gamma \leq \bigcup_{\delta<\alpha} \nu\mu\delta$. Hence

$$\begin{aligned} \nu\mu\alpha &\leq \bigcup_{\delta<\alpha} \nu\mu\delta && \text{by (2),} \\ &\leq \nu\mu\alpha && \text{since } \nu\circ\mu \text{ is nondecreasing.} \end{aligned}$$

This completes the proof in this case also.

Theorem 12.12 *Let A, B, C be ordinals, and let μ, ν be normal functions mapping A into B and B into C respectively. Then $\nu\circ\mu$ is normal.*

Proof By 12.11, $\nu\circ\mu$ is limiting, and by 12.10 it is strictly increasing.

We conclude this section with an important "bracketing" condition.

Theorem 12.13 *Let A be an ordinal and μ a half-normal A-termed sequence of ordinals. Suppose that there exists a $\beta \in A$ such that $\mu\beta \leq \alpha$, and that there exists a $\gamma \in A$ such that $\alpha < \mu\gamma$. Then there is a unique $\delta \in A$ such that $\mu\delta \leq \alpha < \mu\mathcal{S}\delta$.*

Proof Let ε be the least member of A such that $\alpha < \mu\varepsilon$; ε exists by the hypothesis of the theorem. Now $\varepsilon \neq 0$, for $\mu 0 \leq \mu\beta \leq \alpha$. Also ε is not a limit ordinal, for if it were, we would have $\alpha < \bigcup_{\xi<\varepsilon} \mu\xi$, and hence $\alpha < \mu\xi$ for some $\xi < \varepsilon$, contradicting the choice of ε. Hence ε is a successor ordinal, say $\varepsilon = \mathcal{S}\delta$. We then have $\mu\delta \leq \alpha < \mu\mathcal{S}\delta$, which proves existence. If ζ is any ordinal in A but different from δ, then either $\zeta < \delta$, and hence $\mathcal{S}\zeta \leq \delta$ and $\mu\mathcal{S}\zeta \leq \mu\delta \leq \alpha$, or $\delta < \zeta$, and hence $\mathcal{S}\delta \leq \zeta$ and $\alpha < \mu\mathcal{S}\delta \leq \mu\zeta$. Thus δ is unique, since $\mu\zeta \leq \alpha < \mu\mathcal{S}\zeta$ is impossible.

Remark 12.14 For more on normal functions and related functions see Bachmann 1967.

Having the notion of α-termed sequence available, it is natural to speak also of α-ary relations and operations. R is an *α-ary relation* iff $R \subseteq {}^{\alpha}V$. For $\alpha = 0$ this means that $R = 0$ or $R = 1$. For $\alpha = 1, 2, 3, \ldots$, we speak of *unary, binary, ternary, . . . relations.* Binary relations may be identified with relations as introduced in Sec. 1 (see the comments following 12.1). f is an *α-ary operation* on A iff $f \in {}^{({}^{\alpha}A)}A$. A 0-ary operation on A has the form $\{(0,a)\}$ for some $a \in A$; this operation is usually identified with a itself. Since ${}^{1}A$ consists of all pairs $(0,a)$, it is frequently tempting to identify ${}^{1}A$ with A itself, and hence 1-ary operations on A with mappings from A into A. 2-ary operations on A are frequently called *binary operations;* 3-ary operations, *ternary operations*, etc. We

shall not have occasion to discuss α-ary relations and operations further. The theory of such relations and operations is essentially the same as the general theory of algebra.

EXERCISES

12.15 Give examples of *Ord*-termed sequences of ordinals showing that there are no implications among the concepts of 12.1(*ii*) to (*vi*) except those implied by the obvious implications (*iii*) $\Rightarrow$ (*ii*), (*v*) $\Rightarrow$ (*iv*), (*v*) $\Rightarrow$ (*ii*), (*vi*) $\Rightarrow$ (*v*), and (*vi*) $\Rightarrow$ (*iii*).

12.16 Prove 12.6.

12.17 Construct *Ord*-termed sequences μ, ν such that neither μ nor ν is nondecreasing, while $\nu \circ \mu$ is nondecreasing.

12.18 Prove 12.12 without appealing to 12.11, by simplifying the proof of the latter.

13 RECURSION

In this section we will discuss *definition by recursion*, often called *definition by induction*. Actually it is not a matter of definition in the usual sense, as precisely described in the Appendix, in which, typically, the defined notion is expressed explicitly in terms of known notions. For example, recall the definition of inclusion:

$$A \subseteq B \quad \text{iff} \quad \forall x(x \in A \Rightarrow x \in B).$$

Note that in this *explicit* definition, the symbol defined, $\subseteq$, does not appear at all on the right-hand side of the equivalence. Recursive definitions are usually restricted to ordinals, and most familiarly to ω, the set of nonnegative integers. A notion is defined for all nonnegative integers by defining it first for 0 and then for $\mathcal{S}n$, assuming that it has been defined for n. For example, addition is usually defined as follows. For any $m,n \in \omega$,

$$m + 0 = m; \qquad m + \mathcal{S}n = \mathcal{S}(m + n).$$

Here the defined term, $+$, appears on both sides of the equation.

These two methods of definition are quite different. Using the first method, it is clear that the defined term does not add any strength of its own to the discussion, since the defined notion can always be eliminated in favor of known notions. It is not so clear that recursive definitions do not give rise to an essentially stronger theory, although we will show in this section that they do not (the situation here is special for set theory, since in an autonomous development of number theory, within the first-order logic developed in the Appendix, the above "definition" of addition

would not be eliminable). In fact, recursive definitions can be reduced to ordinary definitions. For example, the operation of addition, $+$, of natural numbers can be introduced by an explicit definition, provided that we can show that there is a unique function f mapping $\omega \times \omega$ into ω such that for all $m,n \in \omega$ $f(m,0) = m$ and $f(m,\mathcal{S}n) = \mathcal{S}(f(m,n))$. For then, we simply let

$$+ = \bigcap\{f : f \in {}^{\omega\times\omega}\omega \wedge \forall m\, \forall n\, (m,n \in \omega \Rightarrow f(m,0) = m \wedge f(m,\mathcal{S}n) = \mathcal{S}f(m,n))\}.$$

Note that the set of f's on the right-hand side of this equation is then a singleton $\{f\}$, f the unique function mentioned above, so that $+ = \bigcap\{f\} = f$, this unique function.

It is not so easy to show that such an f exists and is unique as one might hope. A "proof" frequently used in elementary textbooks runs as follows, applied, for example, to justify the recursive definition of addition. "We show by induction that $m + n$ is defined for all $m,n \in \omega$. $m + 0 = m$, and so $m + 0$ is defined. Assuming that $m + n$ is defined, $m + \mathcal{S}n = \mathcal{S}(m + n)$, and so $m + \mathcal{S}n$ is defined. Thus $m + n$ is defined for all $m,n \in \omega$." This argument, however, is erroneous. It mixes language and metalanguage, since the argument talks about an expression's being defined on the same level as the integers themselves. Such a mixing of language and metalanguage leads to obvious contradictions, for example, to Richard's paradox: "Let m be the least nonnegative integer not definable by fewer than 20 words. But we have just defined m by fewer than 20 words." The moral of this little discussion is that we must try to justify recursive definitions by using the ordinary set-theoretical apparatus that has been developed. This amounts to showing that certain functions f exist and are unique; the proof consists in building up such a function from "approximations."

There is another consideration with regard to recursive definitions. We want to be able to do recursion not only over ω but over any ordinal. There are even more general situations that frequently occur in mathematics where some kind of recursive definition is called for. Although these more general situations can always be reduced to recursion over ordinals, it is best simply to give a general recursion principle to cover all the special cases. Our most general recursion principle involves *recursion over well-founded relations* [recall Definition 8.12(*i*)]. Some of the more special recursion principles, for example, 13.2, will seem more intuitively clear.

In this section we shall give a general recursion principle and also various special cases which are useful in practice. In later sections we will apply the principles as needed; the general principle itself is used in Sec. 15, for example.

Theorem 13.1 *(General recursion principle) Let R be a well-founded relation such that for all $x \in Fld\ R$, $\{y : yRx\}$ is a set, and let F be a function with domain $Fld\ R \times V$. Then there is a unique function G such that $Dmn\ G = Fld\ R$ and for all $x \in Fld\ R$,*

$$Gx = F(x, G\restriction\{y : yRx\}).$$

Proof Let

(1) $M = \{h : h$ is a function, $Dmn\ h \subseteq Fld\ R$, and $\forall x \in Dmn\ h$ $(hx = F(x, h\restriction\{y : yRx\}) \wedge \{y : yRx\} \subseteq Dmn\ h)\}$.

We may intuitively think of members of M as approximations of the desired function G, defined only on "R-initial segments" of $Fld\ R$. Let $G = \bigcup M$. To check that G is a function, it is enough, by virtue of Theorem 5.9, to prove the following statement:

(2) For all $f, g \in M$, $f\restriction(Dmn\ f \cap Dmn\ g) = g\restriction(Dmn\ f \cap Dmn\ g)$.

To prove (2), let $f, g \in M$. Let $A = \{x : x \in Dmn f \cap Dmn g$ and $fx \neq gx\}$. We want to show that A is empty; suppose, on the contrary, that $A \neq 0$. By the definition of well-foundedness—Definition 8.12(i)—choose $x \in A$ such that $A \cap \{y : yRx\} = 0$. Now, by (1), $\{y : yRx\} \subseteq Dmn\ f \cap Dmn\ g$. Also, for any y with yRx we have $fy = gy$, by the choice of x. Hence by (1),

$$fx = F(x, f\restriction\{y : yRx\}) = F(x, g\restriction\{y : yRx\}) = gx,$$

which contradicts $x \in A$. Thus $A = 0$ after all, (2) is established, and G is a function.

Clearly $Dmn\ G \subseteq Fld\ R$. To establish the converse inclusion, we need two more easy properties of M.

(3) If N is a nonempty subclass of M, then $\bigcap N \in M$.

Clearly, by 5.15, $\bigcap N$ is a set, $\bigcap N$ is a function, and $Dmn \bigcap N \subseteq Fld\ R$. Suppose that $x \in Dmn \bigcap N$. Then $x \in Dmn\ h$ for each $h \in N$, so that, by (1), $\{y : yRx\} \subseteq Dmn\ h$ for each $h \in N$, and so $\{y : yRx\} \subseteq Dmn \bigcap N$, by (2). Let h be any element of N. Then $x \in Dmn\ h$, and $(\bigcap N)(x) = hx = F(x, h\restriction\{y : yRx\}) = F(x, (\bigcap N)\restriction\{y : yRx\})$. This establishes (3).

(4) If N is a subset of M, then $\bigcup N \in M$.

Indeed, $\bigcup N$ is then a set, by 5.3, and it is a function, by (2) and 5.9. Suppose that $x \in Dmn \bigcup N$. Then there is an $h \in N$ such that $x \in Dmn\ h$. By (1), $\{y : yRx\} \subseteq Dmn\ h \subseteq Dmn \bigcup N$. Also by (1),

$$(\bigcup N)x = hx = F(x, h\restriction\{y : yRx\}) = F(x, (\bigcup N)\restriction\{y : yRx\}).$$

This establishes (4).

Now let us return to the main part of the proof. We need to show that $Fld\ R \subseteq Dmn\ G$. Suppose that this is not the case, and by the definition of well-foundedness choose $x \in Fld\ R \sim Dmn\ G$ such that $(Fld\ R \sim Dmn\ G) \cap \{y : yRx\} = 0$. If yRx, then $y \in Dmn\ G$ and hence $\{h : h \in M \wedge y \in Dmn\ h\} \neq 0$; further, by (3), $\bigcap \{h : h \in M \wedge y \in Dmn\ h\} \in M$. Now we define H to be the function with domain $\{y : yRx\}$ such that for each y such that yRx,

$$(5) \quad Hy = \bigcap \{h : h \in M \wedge y \in Dmn\ h\}.$$

Now $Dmn\ H = \{y : yRx\}$ is a set by the hypothesis of the theorem, so that, by the axiom of substitution (1.35), $Rng\ H$ is a set. By (4), $\bigcup Rng\ H \in M$. Let $k = \bigcup Rng\ H$. From (5) we easily infer that $\{y : yRx\} \subseteq Dmn\ k$. Now let

$$h = k \cup \{(x, F(x, k\restriction\{y : yRx\}))\}.$$

Clearly then $h \in M$. But $x \in Dmn\ h$, so that $x \in Dmn\ G$, which contradicts our choice of x. This contradiction establishes that $Dmn\ G = Fld\ R$.

If $x \in Fld\ R$, then $x \in Dmn\ G$ and so $x \in Dmn\ h$ for some $h \in M$, so that

$$Gx = hx = F(x, h\restriction\{y : yRx\}) = F(x, G\restriction\{y : yRx\}).$$

Thus we have finished the proof that G has the desired properties.

As to uniqueness, suppose H also satisfies the conditions of the theorem. Let $A = \{x : x \in Fld\ R \wedge Gx \neq Hx\}$. Assume that $A \neq 0$; choose $x \in A$ such that $A \cap \{y : yRx\} = 0$. Then $Gy = Hy$ whenever yRx, and hence

$$Gx = F(x, G\restriction\{y : yRx\}) = F(x, H\restriction\{y : yRx\}) = Hx,$$

which contradicts $x \in A$. Hence $A = 0$, so that $G = H$. This completes the proof.

We want to give at once our most limited recursion principle. We give two proofs for this principle: one based on 13.1, the other more involved, in which a simplified version of the proof of 13.1 appears. The second proof should be useful in seeing the idea of the proof of 13.1 clearly.

Theorem 13.2 *(The iteration principle) Let A be any set, $a \in A$, and f a function mapping A into A. Then there is a unique function g mapping ω into A such that $g0 = a$ and $g\S m = fgm$ for all $m \in \omega$.*

First proof Let $R = \{(m,n) : m < n\}$. Clearly R is a well-founded relation such that $\{y : yRx\}$ is a set for each $x \in Fld\ R$; indeed, the first fact is true since $\leq$ is a well-ordering, and the second follows since

Fld $R = \omega$ and $m = \{n : n < m\}$ for each $m \in \omega$. Let F be the function with domain $\omega \times V$ such that for any $m \in \omega$ and $x \in V$,

$$F(m,x) = \begin{cases} a & \text{if } x = 0, \\ f[x(\bigcup Dmn\ x)] & \text{if } x(\bigcup Dmn\ x) \in A, \\ 0 & \text{in any other case.} \end{cases}$$

To understand the second part of this definition, recall the functional notation from 4.2: $x(\bigcup Dmn\ x)$ is the unique y such that $(\bigcup Dmn\ x, y) \in x$ if there is such a unique y. In case such a unique y does not exist, we are not really interested in the result, as will be seen.

Now let G be as in Theorem 13.1, and let $g = G\restriction\omega$ [g is a set by 4.15(*iii*).] Thus g maps ω into V. By induction on m, we show that $gm \in A$ for all $m \in \omega$:

$$g0 = G0 = F(0,G\restriction 0) = F(0,0) = a \in A;$$

assuming that $gm \in A$, we have

$$g\mathcal{S}m = G\mathcal{S}m = F(\mathcal{S}m, G\restriction\{y : y < \mathcal{S}m, y \in \omega\}) = F(\mathcal{S}m, G\restriction\mathcal{S}m) = fGm = fgm \in A$$

Here we have used the fact that $\bigcup\mathcal{S}m = m$ for any $m \in \omega$ [see 9.11(*v*)]; this, of course, motivated the strange-looking definition of F. Thus $g : \omega \to A$, and the above argument also shows that g has the properties desired in the theorem.

For uniqueness of g, assuming that h also satisfies the conditions of the theorem, it is easily shown by induction that $gm = hm$ for every $m \in \omega$. Alternatively, one can argue using the uniqueness of G.

Second proof Let

(1) $M = \{h : h$ is a function, $Dmn\ h \in \omega$, $Rng\ h \subseteq A$, and for all $m \in Dmn\ h$, $hm = a$ if $m = 0$ and $hm = fhn$ if $m = \mathcal{S}n\}$.

Let $g = \bigcup M$ (since $M \subseteq \bigcup_{m\in\omega} {}^mA$, M is a set). Now

(2) For all $h,k \in M$, $h\restriction(Dmn\ h \cap Dmn\ k) = k\restriction(Dmn\ h \cap Dmn\ k)$.

To prove (2), let $h,k \in M$ be given, but suppose that the conclusion fails. Let m be the least integer $n \in \omega$ such $kn \neq hn$. Then $m \neq 0$, since $h0 = a = k0$, by (1). Hence there is a $p \in \omega$ such that $m = \mathcal{S}p$. Thus, using (1),

$$hm = fhp = fkp = km,$$

a contradiction. Thus (2) holds, so that, by 5.9, g is a function. Now we prove, by induction, that for every $m \in \omega$,

(3) $m \in Dmn\ g$ and $gm = a$ if $m = 0$, and $gm = fgn$ if $m = \mathcal{S}n$ for some n.

Suppose that (3) holds for all $m < p$. If $p = 0$, note that $\{(0,a)\} \in M$, and hence (3) holds for p. If $p = \mathcal{S}n$ for some n, then $g \restriction p \cup \{(p,fgn)\} \in M$, and so $p \in Dmn\ g$ and $gp = fgn$, and (3) holds for p. Hence, by complete induction, (3) holds for all $m \in \omega$. This completes the proof of the existence of g. Uniqueness is proved as in the first proof.

We can now derive other useful recursion principles quickly. In the next two principles, as well as in the iteration principle, for simplicity the given functions have only one argument rather than two as in the general recursion theorem.

Theorem 13.3 (*General recursion principle for ordinals*) *Let A be an ordinal, and let F be a function with domain V. Then there is a unique function G with domain A such that, for every $\alpha \in A$, $G\alpha = F(G \restriction \alpha)$.*

Proof Let $R = \{(\alpha,\beta) : \alpha < \beta \in A\}$. Let F' have domain $Fld\ R \times V$, with $F'(x,y) = Fy$ for any $x \in Fld\ R$, $y \in V$. Then the hypotheses of 13.1 hold, so that there is a function G such that $Dmn\ G = Fld\ R$ and $\forall x \in Fld\ R[Gx = F'(x,G \restriction x)]$. Thus $\forall x \in Fld\ R[Gx = F(G \restriction x)]$. If $1 \in A$, then $Fld\ R = A$ and G is as desired. If $1 \notin A$, then $A = 0$ or $A = 1$ and the existence of G is trivial. The uniqueness of G is easily shown, either directly or by appealing to 13.1.

Theorem 13.4 (*Usual recursion principle for ordinals*) *Let A be a nonzero ordinal, B a class. Suppose that $a \in B$, F is a function mapping B into B, and G is a function mapping C into B, where $C = \{f : f \in {}^{\alpha}B \text{ for some } \alpha \in A\}$. Then there is a unique function H mapping A into B such that*

(*i*) $H0 = a$.

(*ii*) $H\mathcal{S}\alpha = FH\alpha$ *for every* α *for which* $\mathcal{S}\alpha \in A$.

(*iii*) $H\beta = G(H \restriction \beta)$ *for every limit ordinal* $\beta \in A$.

Proof We want to apply 13.3, and to this end we define a certain function F° with domain V. Namely, for any set x we let

$$F^{\circ}x = \begin{cases} a & \text{if } x = 0; \\ F(x(\bigcup Dmn\ x)) & \text{if } x \text{ is a function whose domain is a successor ordinal in } A \text{ and } Rng\ x \subseteq B; \\ Gx & \text{if } x \text{ is a function whose domain is a limit ordinal} \in A \text{ and } Rng\ x \subseteq B; \\ 0 & \text{in any other case.} \end{cases}$$

Note that such a function F° does exist. The only question that might occur is in the second case. If x is a function whose domain is $\mathcal{S}\alpha \in A$, and $Rng\ x \subseteq B$, then $Dmn\ x = \mathcal{S}\alpha$, $\bigcup Dmn\ x = \alpha$, by 9.11(v), $x(\bigcup Dmn\ x) \in B$, and hence $x(\bigcup Dmn\ x) \in Dmn\ F$.

Now, by 13.3, let H be a function with domain A such that for every $\alpha \in A$, $H\alpha = F^{\circ}(H\restriction\alpha)$. We now claim that H maps A into B. If this is not true, choose $\alpha \in A$ minimal such that $H\alpha \notin B$. Since $H0 = F^{\circ}(H\restriction 0) = F^{\circ}0 = a \in B$, we have $\alpha \neq 0$. If $\alpha = S\beta$ for some β, then $Rng\,(H\restriction\alpha) \subseteq B$ and so $H\alpha = F^{\circ}(H\restriction\alpha) = F(H\restriction\alpha)\beta = FH\beta \in B$, contradicting the choice of α. Finally, if α is a limit ordinal, then $H\alpha = F^{\circ}(H\restriction\alpha) = G(H\restriction\alpha) \in B$, again a contradiction. Thus such an α does not exist, so that $H : A \to B$. A repetition of the preceding argument shows that H has the desired properties also. The uniqueness proof is easy.

Note that for $A = \omega$, condition (*iii*) of Theorem 13.4 drops out, and we obtain the iteration principle again.

We now want to give versions of the recursion principles in which parameters appear. This is frequently desirable in applications.

Theorem 13.5 (*General recursion principle, with a parameter*) *Let R be a well-founded relation such that, for all $x \in Fld\ R$, $\{y : yRx\}$ is a set, and let F be a function with domain $V \times Fld\ R \times V$. Then there is a unique function G with domain $V \times Fld\ R$ such that, for every $x \in V$ and $y \in Fld\ R$,*

$$G(x,y) = F(x,y,G\restriction\{(x,z) : zRy\}).$$

Proof Let $S = \{((x,y),(x,z)) : yRz\}$. Then S is well-founded: Suppose that $0 \neq A \subseteq Fld\ S$. Let $B = \{2^{nd}\,a : a \in A\}$. (Recall, from 5.13, that $2^{nd}\,(x,y) = y$ for any $x,y \in V$.) Then $0 \neq B \subseteq Fld\ R$, and so we choose $y \in B$ such that $\{u : uRy\} \cap B = 0$. There is an $x \in V$ such that $(x,y) \in A$. Then $\{b : bS(x,y)\} \cap A = 0$. Indeed, if $bS(x,y)$ and $b \in A$, then $(2^{nd}\,b)Ry$ and $2^{nd}\,b \in B$, which is impossible. Thus S is well-founded.

Also, $\{a : aSb\}$ is a set for any $b \in Fld\ S$. Indeed, let $b = (x,y)$ with $x \in V$, $y \in Fld\ R$. Then $\{a : aSb\} = \{x\} \times \{u : uRy\}$, so that the hypothesis of the theorem implies that $\{a : aSb\}$ is a set.

Note that $Fld\ S = V \times Fld\ R$. Hence F maps $Fld\ S \times V$ into V. Now we can apply 13.1. Let G be the function with domain $Fld\ S$ such that, for all $a \in Fld\ S$, $Ga = F(a,G\restriction\{b : bSa\})$. Thus, for any $x \in V$ and $y \in Fld\ R$,

$$\begin{aligned} G(x,y) &= F(x,y,G\restriction\{b : bS(x,y)\}) \\ &= F(x,y,G\restriction\{(x,z) : zRy\}), \end{aligned}$$

as desired. The uniqueness of G is easily established.

Theorem 13.6 (*General recursion principle for ordinals, with a parameter*) *Let A be an ordinal, and let F be a function with domain $V \times A \times V$. Then*

there is a unique function G with domain $V \times A$ such that, for all x and all $\alpha \in A$,

$$G(x,\alpha) = F(x,\alpha,G\restriction\{(x,\beta) : \beta < \alpha\}).$$

We leave the proof as an exercise (cf. the proof of 13.3).

Theorem 13.7 (*Usual recursion principle for ordinals, with a parameter*) *Let A be an ordinal, B a class. Suppose that F is a function mapping B into B, G is a function mapping $B \times A \times B$ into B, and H is a function mapping $B \times A \times V$ into B. Then there is a unique function K mapping $B \times A$ into B such that, for all $x \in B$,*
(*i*) $K(x,0) = Fx$.
(*ii*) $K(x,\mathcal{S}\alpha) = G(x,\alpha,K(x,\alpha))$ *for every α for which $\mathcal{S}\alpha \in A$.*
(*iii*) $K(x,\beta) = H(x,\beta,K\restriction\{(x,\gamma) : \gamma < \beta\})$ *for every limit ordinal $\beta \in A$.*

Proof Since the proof is analogous to the proof of 13.4, we will sketch it only. The idea is to apply 13.6. To this end we define a function F° with domain $V \times A \times V$. If $x,y \in V$ and $\alpha \in A$, we set

$$F^\circ(x,\alpha,y) = \begin{cases} Fx & \text{if } \alpha = 0 \text{ and } x \in B; \\ G(x,\bigcup\alpha,y(x,\bigcup\alpha)) & \text{if } x \in B,\ \alpha \text{ is a successor ordinal, and } y \in {}^{\{x\}\times\alpha}B; \\ H(x,\alpha,y) & \text{if } x \in B,\ \alpha \text{ is a limit ordinal, and } y \in V; \\ 0 & \text{otherwise.} \end{cases}$$

The remainder of the proof is straightforward.

Theorem 13.8 (*Primitive recursion*) *Let B be a set, f a function mapping B into B, and g a function mapping $B \times \omega \times B$ into B. Then there is a unique function h mapping $B \times \omega$ into B such that, for all $x \in B$ and all m,*
(*i*) $h(x,0) = fx$.
(*ii*) $h(x,\mathcal{S}m) = g(x,m,h(x,m))$.

Theorem 13.8 is obtained from 13.7 by specializing to $A = \omega$.

We now give some important applications of recursion.

Theorem 13.9 (*Fixed-point theorem for normal functions*) *Let μ be an Ord-termed normal function. Then for every β there is an $\alpha > \beta$ such that $\mu\alpha = \alpha$.*

Proof We define a function ν by iteration. Let $\nu 0 = \mu\mathcal{S}\beta$ and $\nu\mathcal{S}m = \mu\nu m$. Also, let $\nu\omega = \bigcup_{m\in\omega} \nu m$. Since for any m, by 12.2, $\nu m \leq \mu\nu m =$

$\nu \S m$, ν is a half-normal function, by 12.5. Hence so is $\mu \circ \nu$, by 12.11, and hence

$$\nu\omega \leq \mu\nu\omega = \bigcup_{m \in \omega} \mu\nu m = \bigcup_{m \in \omega} \nu \S m \lneqq \nu\omega,$$

so that $\nu\omega$ is the desired α (note that $\beta \leq \mu\beta < \mu \S \beta = \nu 0 \leq \nu\omega$, so that $\beta < \nu\omega$).

The construction in the proof of 13.9 gives the least $\alpha > \beta$ such that $\mu\alpha = \alpha$. In fact, assume the notation of the proof, and let γ be any ordinal $> \beta$ such that $\mu\gamma = \gamma$. By induction, it is easily seen that $\nu m \leq \gamma$ for every $m \in \omega$; hence $\nu\omega \leq \gamma$, as desired.

From 13.9 it follows that the class A of fixed points of any *Ord*-termed normal function μ is a proper class. Indeed, if A is a set, then $\bigcup A \in Ord$, and 13.9 yields a fixed point $\alpha > \bigcup A$. Thus $\alpha \leq \bigcup A < \alpha$, a contradiction.

The next theorems show that ordinals fully represent well-orderings, as indicated at the beginning of this chapter. These theorems, then, really give the essence of what ordinals are; any reasonable notion of ordinal would have to satisfy them.

Theorem 13.10 *If $\preceq$ is a well-ordering, $\preceq$ a set, then there is a unique α and a unique function f such that f is an isomorphism from $\{(\beta,\gamma) : \beta \leq \gamma < \alpha\}$ onto $\preceq$.*

Proof We apply the general recursion theorem for ordinals (13.3). By 2.6(i), choose $a \notin Fld\ (\preceq)$. For each set h let $Fh =$ $\preceq$-least element of $Fld\ (\preceq) \sim Rng\ h$, if $Fld\ (\preceq) \sim Rng\ h \neq 0$, and $Fh = a$ otherwise. By 13.3, let G be the function with domain *Ord* such that, for every α, $G\alpha = F(G\restriction\alpha)$. Now

(1) If $\alpha < \beta$ and $G\alpha = a$, then $G\beta = a$.

Indeed, if $G\alpha = a$, then $F(G\restriction\alpha) = G\alpha = a$, so that $Fld\ (\preceq) \sim Rng\ (G\restriction\alpha) = 0$. But $G\restriction\alpha \subseteq G\restriction\beta$, so that $Fld\ (\preceq) \sim Rng\ (G\restriction\beta) = 0$ also; it follows that $G\beta = F(G\restriction\beta) = a$. This establishes (1).

(2) If $\alpha < \beta$ and $G\beta \neq a$, then $G\alpha \prec G\beta$.

Indeed, by (1), we then have $G\alpha \neq a$ also. Hence $G\alpha = F(G\restriction\alpha) =$ $\preceq$-least element of $Fld\ (\preceq) \sim Rng\ (G\restriction\alpha)$, and $G\beta = F(G\restriction\beta) =$ $\preceq$-least element of $Fld\ (\preceq) \sim Rng\ (G\restriction\beta)$. Since $G\restriction\alpha \subseteq G\restriction\beta$, we have $Fld\ (\preceq) \sim Rng\ (G\restriction\beta) \subseteq Fld\ (\preceq) \sim Rng\ (G\restriction\alpha)$ and hence $G\beta \in Fld\ (\preceq) \sim Rng\ (G\restriction\alpha)$. Thus $G\alpha \preceq G\beta$ since $G\alpha$ is the $\preceq$-least element of $Fld\ (\preceq) \sim Rng\ (G\restriction\alpha)$. On the other hand, $G\alpha \in Rng\ (G\restriction\beta)$, so that the fact that $G\beta \in Fld\ (\preceq) \sim Rng\ (G\restriction\beta)$ implies that $G\alpha \neq G\beta$. This establishes (2). Next,

(3) There is an α such that $G\alpha = a$.

For, otherwise, by (2), G would be a one-one mapping of Ord into $Fld\ (\preceq)$; since $Fld\ (\preceq)$ is a set by 3.14, Ord would be a set, by 4.11, contradicting 9.8. Hence (3) holds. Letting α be the least ordinal such that $G\alpha = a$, it is then clear, using 8.11, that $f = G\restriction\alpha$ satisfies the conditions of the theorem. As to uniqueness, suppose that β and g also satisfy the conditions. Then $g^{-1}\circ f$ is a strictly increasing α-termed sequence of ordinals, and $Rng\ (g^{-1}\circ f) \subseteq \beta$, so that, by 12.3, $\alpha \leq \beta$. Similarly $\beta \leq \alpha$, so that $\alpha = \beta$. Then, by 12.4, $g^{-1}\circ f = f^{-1}\circ g = I\restriction\alpha$, so that $f = g$.

Theorem 13.11 *For any well-ordering $\preceq$, let $\tau(\preceq)$ be the unique ordinal α given by Theorem* 13.10. *Then $\preceq$ is isomorphic to $\preceq'$ iff $\tau(\preceq) = \tau(\preceq')$. Every ordinal α has the form $\tau(\preceq)$ for some well-ordering $\preceq$, namely, for $\preceq = \{(\beta,\gamma) : \beta \leq \gamma < \alpha\}$.*

Proof Let $\preceq$ and $\preceq'$ be two well-orderings, with f, g isomorphisms from $\{(\beta,\gamma) : \beta \leq \gamma < \tau(\preceq)\}$ and $\{(\beta,\gamma) : \beta \leq \gamma < \tau(\preceq')\}$ onto $\preceq$ and $\preceq'$ respectively. If h is an isomorphism from $\preceq$ onto $\preceq'$, $h\circ f$ is an isomorphism from $\{(\beta,\gamma) : \beta \leq \gamma < \tau(\preceq)\}$ onto $\preceq'$, and so, by 13.10, $\tau(\preceq) = \tau(\preceq')$. On the other hand, if $\tau(\preceq) = \tau(\preceq')$, then $g\circ f^{-1}$ is an isomorphism from $\preceq$ onto $\preceq'$.

The last statement of the theorem is obvious.

Thus isomorphism restricted to well-orderings is an equivalence relation, and the ordinals pick out exactly one representative from each equivalence class. Furthermore, there is at most one isomorphism between two well-orderings. If δ is an ordinal and $\Gamma \subseteq \delta$, then $\{(\alpha,\beta) : \alpha \leq \beta,\ \alpha \in \Gamma,\ \beta \in \Gamma\}$ is isomorphic to $\{(\alpha,\beta) : \alpha \leq \beta \in \gamma\}$ for some $\gamma \leq \delta$ (see Theorem 12.3). If R and S are well-orderings, then either (*i*) there is a $b \in Fld\ S$ such that R is isomorphic to $\{(x,y) : xSySb \wedge y \neq b\}$ or (*ii*) R is isomorphic to S or (*iii*) there is an $a \in Fld\ R$ such that $\{(x,y) : xRyRa \wedge y \neq a\}$ is isomorphic to S.

We conclude this section with a discussion of the Peano postulates for natural numbers.

Definition 13.12 (*i*) *If A is a set, $a \in A$, and f is a function mapping A into A, we say that (A,f,a) is a* ***model for the Peano postulates*** *provided that the following three conditions hold:*

P1 $fx \neq a$, *for all* $x \in A$.

P2 *For all* $x,y \in A$, *if* $fx = fy$, *then* $x = y$.

P3 *For every subset B of A, if $a \in B$ and if $\forall x(x \in B \Rightarrow fx \in B)$, then $B = A$.*

(*ii*) *If (A,f,a) and (B,g,b) are models for the Peano postulates, then (A,f,a) is* ***isomorphic to*** *(B,g,b) provided that there is a one-one function h mapping A onto B such that $ha = b$ and $\forall x \in A\ (hfx = ghx)$.*

The notion of isomorphic models for the Peano postulates extends the notion of isomorphism of relations introduced in 8.3.

Theorem 13.13 $(\omega,f,0)$ *is a model for the Peano postulates, where* $fm = \mathcal{S}m$ *for every* m.

Proof $(P1)$ is obvious, $(P2)$ follows from 9.13(ix), and $(P3)$ follows from the definition of ω (11.1).

Theorem 13.14 *Any two models for the Peano postulates are isomorphic.*

Proof Clearly it suffices to show that, if (B,g,b) is a model for the Peano postulates, then $(\omega,f,0)$ is isomorphic to (B,g,b), where f is defined as in 13.13. By the iteration principle (13.2), let h be the function mapping ω into B such that $h0 = b$ and $h\mathcal{S}m = ghm$ for all m. We have only to show that h is one-one and has range B. Suppose that h is not one-one. Let m be the least integer such that there is an integer $n > m$ with $hm = hn$. Thus $n = \mathcal{S}p$ for some p, and so $hn = h\mathcal{S}p = ghp$. Now $m \neq 0$; otherwise, $ghp = hn = hm = h0 = b$, contradicting $(P1)$ for (B,g,b). Thus there is a q such that $m = \mathcal{S}q$. Now $hm = h\mathcal{S}q = ghq$, so that $ghq = ghp$. By $(P2)$ for (B,g,b), $hq = hp$. But, from 9.13(x), we know that $q < p$, so the fact that $q < m$ contradicts the choice of m.

To show that h has range B, we apply $(P3)$. Since $h0 = b$, $b \in Rng\ h$. Suppose that $x \in Rng\ h$, say $x = hm$. Then $h\mathcal{S}m = ghm = gx$, and so $gx \in Rng\ h$. Thus, by $(P3)$ for (B,g,b), $B = Rng\ h$. This completes the proof.

Remark 13.15 The general recursion principle (13.1) is due to Montague, Scott, Tarski 1956, who proved an even more general result. For a detailed discussion of the iteration principle and its proof, and of the Peano postulates, see Henkin 1960.

Theorem 13.14 gives rise to a seeming paradox when combined with a result of Gödel 1931. By Gödel's result, set theory (if consistent) has two distinct models; a certain number-theoretic statement holds in one, but not in the other. But, by 13.14, any two systems of integers are isomorphic, and hence any statement that holds of one holds of the other also. Thus we conclude that 13.14 applies only to a definite and fixed universe of sets. It is only when several possible universes of sets are compared that different properties of integers are possible. More formally, 13.14 may be viewed as simply an expression derivable from our set-theoretical axioms, although Gödel's result states the existence of a number-theoretical expression φ such that neither φ nor $\neg\varphi$ can be derived from our axioms. Theorem 13.14 is a statement about sets, functions, etc., while Gödel's theorem is a metamathematical result about our set

theory. It is a mixing of language and metalanguage that leads to the paradox.

EXERCISES

13.16 Prove 13.6.

13.17 Fill in all details in the proof of 13.7.

13.18 Show that there is a function μ mapping *Ord* into *Ord* such that for any α, $\mu 0 = 0$, $\mu \mathcal{S} \alpha = \mathcal{S}\mathcal{S}\mu\alpha$, and $\mu\alpha = \bigcup_{\beta<\alpha} \mu\beta$ for α a limit ordinal.

13.19 Show that there is a function f mapping $\omega \times \omega$ into ω such that the following conditions hold for all m, n:

(*a*) $f(0,n) = \mathcal{S}n$;
(*b*) $f(\mathcal{S}m,0) = f(m,1)$;
(*c*) $f(\mathcal{S}m,\mathcal{S}n) = f(m,f(\mathcal{S}m,n))$.

Hint: Let $R = \{((m,n),(p,q)) : m < p \text{ or } (m = p \wedge n < q)\}$.

14 ORDINAL ARITHMETIC

In this section we introduce ordinal addition, multiplication, and exponentiation and prove basic facts about them. As indicated in Sec. 13, these operations will be defined recursively, which means that the recursion principles in Sec. 13 will be applied to prove the existence and uniqueness of the operations. We begin with addition.

Theorem 14.1 *There is a unique function K mapping $Ord \times Ord$ into Ord such that for all α, β, γ,*

(*i*) $K(\alpha,0) = \alpha$.
(*ii*) $K(\alpha,\mathcal{S}\beta) = \mathcal{S}(K(\alpha,\beta))$.
(*iii*) $K(\alpha,\gamma) = \bigcup_{\delta<\gamma} K(\alpha,\delta)$ *if* $\gamma = \bigcup\gamma \neq 0$.

Proof We apply 13.7, the usual recursion principle for ordinals, with a parameter. Let $F\alpha = \alpha$ for all α. Let $G(\alpha,\beta,\gamma) = \mathcal{S}\gamma$ for all α, β, γ. For any α, β and any $f \in V$ let $H(\alpha,\beta,f) = \bigcup_{x \in Dmn\, f} fx$ if f is a function with $Rng\, f \subseteq Ord$, and let $H(\alpha,\beta,f) = 0$ otherwise. Then choose K mapping $Ord \times Ord$ into Ord in accordance with 13.7. For any α, β, γ with $\gamma = \bigcup\gamma \neq 0$ we then have

$$K(\alpha,0) = F\alpha = \alpha.$$
$$K(\alpha,\mathcal{S}\beta) = G(\alpha,\beta,K(\alpha,\beta)) = \mathcal{S}(K(\alpha,\beta)).$$
$$K(\alpha,\gamma) = H(\alpha,\gamma,K\restriction\{(\alpha,\delta) : \delta < \gamma\}) = \bigcup_{\delta<\gamma} K(\alpha,\delta).$$

This completes the proof (uniqueness is easy).

Definition 14.2 $\dot{+}$ *is the unique function K of Theorem* 14.1. *We write* $\alpha \dot{+} \beta$ *instead of* $\dot{+}(\alpha,\beta)$.

We reserve the ordinary $+$ for cardinal addition (see Sec. 20). Note that $\alpha \dot{+} 1 = \mathcal{S}(\alpha \dot{+} 0) = \mathcal{S}\alpha$, so that we can now write the more suggestive $\alpha \dot{+} 1$ for $\mathcal{S}\alpha$.

Definition 14.2 and Theorem 14.1 can be reformulated as follows.

Theorem 14.3 *For any α, β, γ with $\gamma = \bigcup\gamma \neq 0$,*
(i) $\alpha \dot{+} 0 = \alpha$.
(ii) $\alpha \dot{+} \mathcal{S}\beta = \mathcal{S}(\alpha \dot{+} \beta)$.
(iii) $\alpha \dot{+} \gamma = \bigcup_{\delta<\gamma} (\alpha \dot{+} \delta)$.

Theorem 14.4 *For every α, $\langle \alpha \dot{+} \beta : \beta \in Ord\rangle$ is a normal function.*

Proof By 12.6.

This theorem is extremely useful in elementary ordinal arithmetic.

It will help the intuitive picture of the addition of ordinals if we give an equivalent definition. For this purpose it is helpful to develop further our intuitive comments at the beginning of the chapter to the effect that ordinals are associated with well-orderings. Recall that, for any ordinal γ, $\{(\alpha,\beta) : \alpha \leq \beta < \gamma\}$ is a well-ordering. In the other direction 13.10 tells us that every well-ordering is isomorphic to an ordering of this special form. Thus $\dot{+}$ corresponds to a certain operation on well-orderings, which we give in the following.

Theorem 14.5 *For any α, β, let $R = \{((0,\gamma),(1,\delta)) : \gamma \in \alpha, \delta \in \beta\} \cup \{((0,\gamma),(0,\delta)) : \gamma \leq \delta < \alpha\} \cup \{((1,\gamma),(1,\delta)) : \gamma \leq \delta < \beta\}$. Then R is a well-ordering, R is a set, and there is a unique isomorphism F of R with $\{(\gamma,\delta) : \gamma \leq \delta < \alpha \dot{+} \beta\}$.*

Thus $\alpha \dot{+} \beta$ is ordered by first putting all elements of α in their natural order and then adjoining the elements of β in their natural order.

Proof It is straightforward to check that R is a well-ordering. Let $F = \{((0,\gamma),\gamma) : \gamma \in \alpha\} \cup \{((1,\gamma),\alpha \dot{+} \gamma) : \gamma \in \beta\}$. Clearly $\alpha \subseteq Rng\, F \subseteq \alpha \dot{+} \beta$, using 14.4 for the last inclusion. To show that $Rng\, F = \alpha \dot{+} \beta$, it suffices to show that any $\delta \in (\alpha \dot{+} \beta) \sim \alpha$ is in the range of F. Since $\delta \notin \alpha$, i.e., $\delta \not< \alpha$, we thus have $\alpha \leq \delta < \alpha \dot{+} \beta$. Applying 12.13 to the normal function $\langle \alpha \dot{+} \gamma : \gamma < \mathcal{S}\beta\rangle$ (cf. 14.4), we obtain a $\gamma < \mathcal{S}\beta$ such that $\alpha \dot{+} \gamma \leq \delta < \alpha \dot{+} \mathcal{S}\gamma$. Since $\alpha \dot{+} \mathcal{S}\gamma = \mathcal{S}(\alpha \dot{+} \gamma)$, it follows from 9.13(*v*) that $\alpha \dot{+} \gamma = \delta$, as desired. Therefore, $Rng\, F = \alpha \dot{+} \beta$. Using the normality of $\langle \alpha \dot{+} \gamma : \gamma < \beta\rangle$ again, we easily see that xRy iff $Fx \leq Fy$ and that F is one-one. This shows the existence of F. Uniqueness follows from 13.10.

Some commonly used properties of ordinal addition are as follows.

Theorem 14.6 (*i*) $\alpha \dot{+} (\beta \dot{+} \gamma) = (\alpha \dot{+} \beta) \dot{+} \gamma$.
(*ii*) $\alpha < \beta$ *iff there is a* $\gamma > 0$ *such that* $\alpha \dot{+} \gamma = \beta$.
(*iii*) *If* $\alpha < \beta$, *then* $\alpha \dot{+} \gamma \leq \beta \dot{+} \gamma$.
(*iv*) $S(\alpha) = \alpha \dot{+} 1$.
(*v*) $0 \dot{+} \alpha = \alpha$.
(*vi*) $\beta \leq \alpha \dot{+} \beta$.
(*vii*) *If* $\beta < \gamma$, *then* $\alpha \dot{+} \beta < \alpha \dot{+} \gamma$.

Proof (*i*) By transfinite induction on γ:

$$\alpha \dot{+} (\beta \dot{+} 0) = \alpha \dot{+} \beta = (\alpha \dot{+} \beta) \dot{+} 0;$$
$$\alpha \dot{+} (\beta \dot{+} S\gamma) = \alpha \dot{+} S(\beta \dot{+} \gamma) = S(\alpha \dot{+} (\beta \dot{+} \gamma))$$
$$= S((\alpha \dot{+} \beta) \dot{+} \gamma) = (\alpha \dot{+} \beta) \dot{+} S\gamma;$$

for $\gamma = \bigcup \gamma \neq 0$, we need to use 12.12. Let $\mu = \langle \alpha \dot{+} \delta : \delta \in Ord \rangle$ and $\nu = \langle \beta \dot{+} \delta : \delta \in Ord \rangle$. Then

$$\begin{aligned} \alpha \dot{+} (\beta \dot{+} \gamma) &= \mu\nu\gamma = \textstyle\bigcup_{\delta<\gamma} \mu\nu\delta && \text{by 12.12,} \\ &= \textstyle\bigcup_{\delta<\gamma} [\alpha \dot{+} (\beta \dot{+} \delta)] \\ &= \textstyle\bigcup_{\delta<\gamma} [(\alpha \dot{+} \beta) \dot{+} \delta] && \text{by induction assumption,} \\ &= (\alpha \dot{+} \beta) \dot{+} \gamma && \text{by 14.4.} \end{aligned}$$

To prove (*ii*), note that $\alpha \dot{+} 0 = \alpha < \beta$, although $\beta < S\beta \leq \alpha \dot{+} S\beta$, by 12.2. Applying 12.13, we obtain a γ such that $\alpha \dot{+} \gamma \leq \beta < \alpha \dot{+} S\gamma$. Since $\alpha \dot{+} S\gamma = S(\alpha \dot{+} \gamma)$, it follows that $\alpha \dot{+} \gamma = \beta$. (*iii*) is easily shown by transfinite induction on γ, and (*v*) by transfinite induction on α. For (*iv*), $\alpha \dot{+} 1 = \alpha \dot{+} S0 = S(\alpha \dot{+} 0) = S\alpha$. (*vi*) and (*vii*) are immediate consequences of 14.4 and 12.2.

Note that we can now prove that $2 \dot{+} 2 = 4$; until now we could not even formulate this fact. Indeed, $2 \dot{+} 2 = 2 \dot{+} SS0 = S(2 \dot{+} S0) = SS(2 \dot{+} 0) = SS2 = 4$. In Sec. 19, where we discuss finite sets, we recapture several other elementary facts about integers. See also the discussion in Sec. 15 of Cantor normal form. Observe, to answer an obvious question, that $\alpha \dot{+} \beta \neq \beta \dot{+} \alpha$ in general. In fact, $\omega \dot{+} 1 = S\omega > \omega$. On the other hand, $1 \dot{+} \omega = \omega$, as follows from Theorem 14.9.

Theorem 14.7 $m \dot{+} n \in \omega$.

Proof By induction on n.

Theorem 14.8 $m \dot{+} n = n \dot{+} m$.

Proof By induction on m:

$0 \dot{+} n = n = n \dot{+} 0$ by 14.6(*v*).

Assuming $\forall n(m \dot{+} n = n \dot{+} m)$, we show $\forall n(\mathcal{S}m \dot{+} n = n \dot{+} \mathcal{S}m)$ by induction on n:

$$\begin{aligned} \mathcal{S}m \dot{+} 0 &= \mathcal{S}m = 0 \dot{+} \mathcal{S}m \qquad \text{by } 14.6(v), \\ \mathcal{S}m \dot{+} \mathcal{S}n &= \mathcal{S}(\mathcal{S}m \dot{+} n) = \mathcal{S}(n \dot{+} \mathcal{S}m) \\ &= \mathcal{S}\mathcal{S}(n \dot{+} m) = \mathcal{S}\mathcal{S}(m \dot{+} n) = \mathcal{S}(m \dot{+} \mathcal{S}n) \\ &= \mathcal{S}(\mathcal{S}n \dot{+} m) = \mathcal{S}n \dot{+} \mathcal{S}m. \end{aligned}$$

Theorem 14.9 $\omega \leq \alpha$ *iff* $1 \dot{+} \alpha = \alpha$.

Proof First we show that $1 \dot{+} \omega = \omega$. Indeed, $\omega \leq 1 \dot{+} \omega$, by $14.6(vi)$, while

$$1 \dot{+} \omega = \bigcup_{m \in \omega} (1 \dot{+} m) \subseteq \omega,$$

by 14.7. Thus $1 \dot{+} \omega = \omega$. If $\omega \leq \alpha$, then $\alpha = \omega \dot{+} \beta$ for some β, using $14.6(ii)$, and hence $1 \dot{+} \alpha = 1 \dot{+} (\omega \dot{+} \beta) = (1 \dot{+} \omega) \dot{+} \beta = \omega \dot{+} \beta = \alpha$. If, on the other hand, $\alpha < \omega$, then, by 14.8, $1 \dot{+} \alpha = \alpha \dot{+} 1 \neq \alpha$. This completes the proof.

We now turn to ordinal multiplication.

Definition 14.10 $\bullet$ *is the unique function mapping* $Ord \times Ord$ *into* Ord *such that for any* α, β, γ *with* $\gamma = \bigcup \gamma \neq 0$,
(i) $\alpha \bullet 0 = 0$.
(ii) $\alpha \bullet \mathcal{S}\beta = (\alpha \bullet \beta) \dot{+} \alpha$.
(iii) $\alpha \bullet \gamma = \bigcup_{\delta < \gamma} (\alpha \bullet \delta)$.

Again, we reserve the usual multiplication symbol $\cdot$ for cardinal multiplication (see Sec. 21).

Theorem 14.11 *If* $\alpha \neq 0$, *then* $\langle \alpha \bullet \beta : \beta \in Ord \rangle$ *is a normal function.*

Proof It is sufficient to observe that $\alpha \bullet \mathcal{S}\beta = \alpha \bullet \beta \dot{+} \alpha > \alpha \bullet \beta$, by 14.4, and to apply 12.6.

It follows from 14.11 that $\langle \alpha \bullet \beta : \beta \in Ord \rangle$ is always half-normal. The following theorem expresses ordinal multiplication in terms of the well-orderings $\{(\gamma,\delta) : \gamma \leq \delta < \alpha\}$ and $\{(\gamma,\delta) : \leq \delta < \beta\}$.

Theorem 14.12 *Let* $\preceq = \{((\gamma,\delta),(\varepsilon,\zeta)) : \gamma,\varepsilon \in \alpha;\delta,\zeta \in \beta;$ *and either* $\delta < \zeta$, *or else* $\delta = \zeta$ *and* $\gamma \leq \varepsilon\}$. *Then* $\preceq$ *is a well-ordering, and there is a unique isomorphism* F *of* $\preceq$ *with* $\{(\gamma,\delta) : \gamma \leq \delta < \alpha \bullet \beta\}$.

Thus $\alpha \cdot \beta$ is ordered by *substituting* a copy of α for each element of β.

Proof We may assume that $\alpha, \beta \neq 0$. It is straightforward to check that $\leq$ is a well-ordering. For any $\gamma \in \alpha, \delta \in \beta$, let $F(\gamma, \delta) = (\alpha \cdot \delta) \dot{+} \gamma$. Then $(\alpha \cdot \delta) \dot{+} \gamma < (\alpha \cdot \delta) \dot{+} \alpha$, (by 14.4) $= \alpha \cdot \mathcal{S}\delta \leq \alpha \cdot \beta$ (by 14.11). Thus $F(\gamma, \delta) \in \alpha \cdot \beta$. F maps onto $\alpha \cdot \beta$: Suppose that $\varepsilon \in \alpha \cdot \beta$. We have $\alpha \cdot 0 = 0 \leq \varepsilon$, and $\varepsilon < \alpha \cdot \beta$. Hence, by 12.13, there is a δ such that $\alpha \cdot \delta \leq \varepsilon < \alpha \cdot \mathcal{S}\delta$. Clearly $\delta < \beta$. Now $(\alpha \cdot \delta) \dot{+} 0 \leq \varepsilon$ and $(\alpha \cdot \delta) \dot{+} \alpha = \alpha \cdot \mathcal{S}\delta > \varepsilon$, so that, again by 12.13, there is a γ such that $(\alpha \cdot \delta) \dot{+} \gamma \leq \varepsilon < (\alpha \cdot \delta) \dot{+} \mathcal{S}\gamma$. Clearly $\gamma < \alpha$ and $F(\gamma, \delta) = (\alpha \cdot \delta) \dot{+} \gamma = \varepsilon$. Thus F maps onto $\alpha \cdot \beta$. Next, if $(\gamma, \delta) \prec (\varepsilon, \zeta)$, then either $\delta < \zeta$, and hence $F(\gamma, \delta) = (\alpha \cdot \delta) \dot{+} \gamma < (\alpha \cdot \delta) \dot{+} \alpha = \alpha \cdot \mathcal{S}\delta \leq \alpha \cdot \zeta \leq (\alpha \cdot \zeta) \dot{+} \varepsilon = F(\varepsilon, \zeta)$, or $\delta = \zeta$ and $\gamma < \varepsilon$, in which case $F(\gamma, \delta) = (\alpha \cdot \delta) \dot{+} \gamma < (\alpha \cdot \delta) \dot{+} \varepsilon = F(\varepsilon, \zeta)$. By 8.11, the proof of the existence of F is complete. Uniqueness follows from 13.10.

We now give basic properties of ordinal multiplication:

Theorem 14.13 (*i*) $\alpha \cdot (\beta \dot{+} \gamma) = (\alpha \cdot \beta) \dot{+} (\alpha \cdot \gamma)$.
(*ii*) $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$.
(*iii*) $0 \cdot \alpha = \alpha \cdot 0 = 0$.
(*iv*) $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$.
(*v*) *If* $\alpha \neq 0$, *then* $\beta \leq \alpha \cdot \beta$.
(*vi*) *If* $\alpha \neq 0$ *and* $\beta < \gamma$, *then* $\alpha \cdot \beta < \alpha \cdot \gamma$.
(*vii*) *If* $\alpha \neq 0$ *and* $\beta > 1$, *then* $\alpha < \alpha \cdot \beta$.
(*viii*) *If* $\alpha < \beta$, *then* $\alpha \cdot \gamma \leq \beta \cdot \gamma$.
(*ix*) $\alpha \cdot 2 = \alpha \dot{+} \alpha$.
(*x*) *If* $\alpha, \beta > 1$, *then* $\alpha \dot{+} \beta \leq \alpha \cdot \beta$.
(*xi*) *If* $\alpha, \beta \neq 0$, *then* $\alpha \cdot \beta \neq 0$.

Proof (*i*) and (*ii*) each involve a straightforward transfinite induction on γ, similar to the proof of the associative law for addition, 14.6(*i*). In proving (*ii*), it is helpful to use (*i*). $\forall \alpha (0 \cdot \alpha = 0)$ is also easily shown by induction on α; $\forall \alpha (\alpha \cdot 0 = 0)$ is known. For (*iv*), note first that $\alpha \cdot 1 = \alpha \cdot \mathcal{S}0 = (\alpha \cdot 0) \dot{+} \alpha = 0 \dot{+} \alpha = \alpha$, using 14.6(*v*). The fact that $1 \cdot \alpha = \alpha$ for all α is shown by an easy transfinite induction on α. (*v*) and (*vi*) are both immediate consequences of 14.11, also using 12.2 for (*v*). (*vii*) follows immediately from (*vi*) and (*iv*). (*viii*) is easily seen by transfinite induction, using 14.6(*iii*). As to (*ix*), we have $\alpha \cdot 2 = \alpha \cdot \mathcal{S}1 = (\alpha \cdot 1) \dot{+} \alpha = \alpha \dot{+} \alpha$, using (*iv*). We prove (*x*) by transfinite induction on β, assuming $\alpha > 1$. $\alpha \dot{+} 2 \leq \alpha \dot{+} \alpha = \alpha \cdot 2$, and, assuming that $\beta > 1$,

$$\begin{aligned} \alpha \dot{+} \mathcal{S}\beta &= \mathcal{S}(\alpha \dot{+} \beta) \leq \mathcal{S}(\alpha \cdot \beta) \qquad \text{using } 9.13(x), \\ &= (\alpha \cdot \beta) \dot{+} 1 < (\alpha \cdot \beta) \dot{+} \alpha = \alpha \cdot \mathcal{S}\beta \end{aligned}$$

Finally, if β is a limit ordinal, then

$$\begin{aligned}\alpha \dot{+} \beta &= \bigcup_{\gamma<\beta} (\alpha \dot{+} \gamma) \\ &= \bigcup_{1<\gamma<\beta} (\alpha \dot{+} \gamma) \leq \bigcup_{1<\gamma<\beta} (\alpha \cdot \gamma) \\ &= \bigcup_{\gamma<\beta} (\alpha \cdot \gamma) \\ &= \alpha \cdot \beta,\end{aligned}$$

since $\langle \alpha \dot{+} \gamma : \gamma \in Ord \rangle$ is strictly increasing. Thus (x) holds. Finally (xi) follows immediately from (v).

Simple arithmetic using multiplication can now be carried out. For example, $2 \cdot 3 = 2 \cdot \mathcal{S}2 = (2 \cdot 2) \dot{+} 2 = (2 \dot{+} 2) \dot{+} 2 = 4 \dot{+} 2 = \mathcal{S}(4 \dot{+} 1) = \mathcal{S}\mathcal{S}4 = 6$. Again, the commutative law for multiplication fails. This results from the fact that $2 \cdot \omega = \omega$, although $\omega \cdot 2 = \omega + \omega > \omega$, by 14.13$(ix)$. The first fact, $2 \cdot \omega = \omega$, follows from 14.15.

Theorem 14.14 (i) $m \cdot n \in \omega$.
(ii) $m \cdot n = n \cdot m$.

Proof (i) is easily seen by induction on n, using 14.7. We prove (ii) by induction on n. The case $n = 0$ is given by 14.13(iii). Now assume, inductively, that $\forall m\ (m \cdot n = n \cdot m)$. We prove $\forall m\ (m \cdot \mathcal{S}n = \mathcal{S}(n) \cdot m)$ by induction on m. The case $m = 0$ is again given by 14.13(iii). Finally,

$$\begin{aligned}\mathcal{S}m \cdot \mathcal{S}n &= \mathcal{S}m \cdot n \dot{+} \mathcal{S}m = n \cdot \mathcal{S}m \dot{+} \mathcal{S}m \\ &= (n \cdot m \dot{+} n) \dot{+} \mathcal{S}m \\ &= (m \cdot n \dot{+} n) \dot{+} \mathcal{S}m \\ &= (m \cdot n \dot{+} \mathcal{S}m) \dot{+} n && \text{by 14.6}(i), \text{ 14.8}, \\ &= [(m \cdot n \dot{+} m) \dot{+} 1] \dot{+} n && \text{by 14.6}(iv), \text{ 14.6}(i), \\ &= (m \cdot \mathcal{S}n \dot{+} 1) \dot{+} n \\ &= (\mathcal{S}n \cdot m \dot{+} 1) \dot{+} n \\ &= \mathcal{S}n \cdot m \dot{+} \mathcal{S}n && \text{by 14.6}(i), \text{ 14.6}(iv), \\ &= \mathcal{S}n \cdot \mathcal{S}m.\end{aligned}$$

Theorem 14.15 *If* $m \neq 0$, *then* $m \cdot \omega = \omega$.

Proof By 14.13(v), $\omega \leq m \cdot \omega$. On the other hand,

$$m \cdot \omega = \bigcup_{n \in \omega} (m \cdot n) \subseteq \omega.$$

Thus $m \cdot \omega = \omega$.

Note that $m \cdot \alpha$ is not equal to α for every $\alpha \geq \omega$; for example, $2 \cdot (\omega \dot{+} 1) = \omega \dot{+} 2 \neq \omega \dot{+} 1$.

The next theorem, the *division algorithm*, specializes to the ordinary division algorithm in the case of natural numbers. It plays a central role in the advanced development of ordinal arithmetic.

Theorem 14.16 *(Division algorithm) If α and β are given, with $\beta \neq 0$, then there exist unique γ and δ such that $\alpha = \beta \cdot \gamma \dotplus \delta$ and $\gamma \leq \alpha$, $\delta < \beta$.*

Proof We have $\beta \cdot 0 = 0 \leq \alpha$ and $\beta \cdot \S\alpha > \beta \cdot \alpha \geq \alpha$, using 14.13(*vi*) and (*v*). Hence, by 14.11 and 12.13, there is a unique γ such that $\beta \cdot \gamma \leq \alpha < \beta \cdot \S\gamma$. Since $\beta \cdot \gamma \dotplus 0 = \beta \cdot \gamma \leq \alpha$ and $\beta \cdot \gamma \dotplus \beta = \beta \cdot \S\gamma > \alpha$, we may apply 12.13 in conjunction with 14.4 this time to obtain a unique δ such that $\beta \cdot \gamma \dotplus \delta \leq \alpha < \beta \cdot \gamma \dotplus \S\delta$. Since $\beta \cdot \gamma \dotplus \S\delta = \S(\beta \cdot \gamma \dotplus \delta)$, it follows that $\alpha = \beta \cdot \gamma \dotplus \delta$. Now $\alpha < \gamma$ implies that $\alpha < \gamma \leq \beta \cdot \gamma$, by 14.13(*v*), and this contradicts our assumption above; hence $\gamma \leq \alpha$. Similarly, $\delta < \beta$. This proves existence. If γ' and δ' also satisfy the conditions, then $\beta \cdot \gamma' \leq \alpha < \beta \cdot \gamma' \dotplus \beta = \beta \cdot \S\gamma'$, and hence $\gamma = \gamma'$, by the uniqueness of γ. Also, then, $\beta \cdot \gamma \dotplus \delta' = \alpha < \beta \cdot \gamma \dotplus \S\delta'$, so that $\delta = \delta'$, by the uniqueness of δ. This completes the proof.

Theorem 14.17 *The following three conditions are equivalent:*
(*i*) *α is a limit ordinal.*
(*ii*) *$\alpha = \omega \cdot \beta$ for some $\beta \neq 0$.*
(*iii*) *For every $m \in \omega \sim 1$, $m \cdot \alpha = \alpha$, and $\alpha \neq 0$.*

Proof (*i*) $\Rightarrow$ (*ii*) By the division algorithm, choose β, m such that $\alpha = \omega \cdot \beta \dotplus m$. If $m \neq 0$, then $m = \S n$ for some n, and $\alpha = \S(\omega \cdot \beta \dotplus n)$, which is impossible. Thus $m = 0$ and $\alpha = \omega \cdot \beta$.

(*ii*) $\Rightarrow$ (*iii*) $m \cdot \alpha = m \cdot (\omega \cdot \beta) = (m \cdot \omega) \cdot \beta = \omega \cdot \beta$, by 14.15.

(*iii*) $\Rightarrow$ (*i*) Suppose that (*iii*) holds but (*i*) fails; say $\alpha = \S\beta$. Then $\alpha = 2 \cdot \alpha = 2 \cdot \S\beta = 2 \cdot \beta \dotplus 2 \geq \beta \dotplus 2$; thus $\S\S\beta \leq \S\beta$, which is impossible.

We now turn to the last basic operation on ordinals, *exponentiation*.

Definition 14.18 *Let be the unique function mapping $Ord \times Ord$ into Ord such that, for any α, β, γ with $\gamma = \bigcup\gamma \neq 0$,*
(*i*) $\alpha^{\cdot 0} = 1$.
(*ii*) $\alpha^{\cdot \S\beta} = \alpha^{\cdot \beta} \cdot \alpha$.
(*iii*) $\alpha^{\cdot \gamma} = \bigcup_{\delta < \gamma} \alpha^{\cdot \delta}$.

We write exponentiation without the dot only for cardinal exponentiation, introduced in Sec. 22.

Theorem 14.19 *If $\alpha > 1$, then $\langle \alpha^{\cdot \beta} : \beta \in Ord \rangle$ is a normal function.*

Proof Note first that $\alpha^{\cdot \beta} \neq 0$ for all β (transfinite induction on β). Thus $\alpha^{\cdot \S\beta} = \alpha^{\cdot \beta} \cdot \alpha > \alpha^{\cdot \beta} \cdot 1 = \alpha^{\cdot \beta}$, using 14.13(*vi*). Now 12.6 yields the desired result.

Again, it is clear that $\langle \alpha^{\cdot\beta} : \beta \in Ord \rangle$ is also half-normal if $\alpha \neq 0$. We will not express the well-ordering of $\alpha^{\cdot\beta}$ explicitly in terms of α and β, as we did for $\alpha \dot{+} \beta$ and $\alpha \bullet \beta$ (14.5 and 14.12), because of the rather complicated expression involved; see Bachmann 1967 or Sierpinski 1965. The simplest properties of exponentiation are given in the following.

Theorem 14.20 (*i*) $0^{\cdot\alpha} = 1$ *if* $\alpha = 0$ *or if* α *is a limit ordinal.*
(*ii*) $0^{\cdot\alpha} = 0$ *if* α *is a successor ordinal.*
(*iii*) $1^{\cdot\alpha} = 1$.
(*iv*) $\alpha^{\cdot 0} = 1$.
(*v*) $\alpha^{\cdot 1} = \alpha$.
(*vi*) $\alpha^{\cdot 2} = \alpha \bullet \alpha$.
(*vii*) *If* $\alpha > 1$ *and* $\beta > 1$, *then* $\alpha < \alpha^{\cdot\beta}$.
(*viii*) *If* $\alpha > 1$, *then* $\beta \leq \alpha^{\cdot\beta}$.
(*ix*) *If* $\alpha > 1$ *and* $\beta < \gamma$, *then* $\alpha^{\cdot\beta} < \alpha^{\cdot\gamma}$.
(*x*) *If* $\alpha < \beta$, *then* $\alpha^{\cdot\gamma} \leq \beta^{\cdot\gamma}$.
(*xi*) $\alpha^{\cdot(\beta \dot{+} \gamma)} = \alpha^{\cdot\beta} \bullet \alpha^{\cdot\gamma}$ *if* $\alpha \neq 0$.
(*xii*) $(\alpha^{\cdot\beta})^{\cdot\gamma} = \alpha^{\cdot(\beta \bullet \gamma)}$ *if* $\alpha \neq 0$.
(*xiii*) *If* $1 < \alpha, \beta$, *then* $1 < \alpha^{\cdot\beta}$.
(*xiv*) *If* $1 < \alpha, \beta$, *then* $\alpha \bullet \beta \leq \alpha^{\cdot\beta}$.

Proof (*i*) and (*ii*) are easily established together by transfinite induction; (*iii*) also follows by an easy transfinite induction. Further, (*iv*), (*v*), and (*vi*) are all straightforward. (*viii*) and (*ix*) follow from 14.19, and (*vii*) from (*v*) and (*ix*). An easy transfinite induction suffices to establish (*x*). The inductive proofs of (*xi*) and (*xii*) are similar to the proof of 14.6(*i*), but since some steps are more involved, we will sketch the proof of (*xi*). We have

$$\begin{aligned} \alpha^{\cdot(\beta \dot{+} 0)} &= \alpha^{\cdot\beta} = \alpha.^{\beta} \bullet 1 = \alpha^{\cdot\beta} \bullet \alpha^{\cdot 0} \qquad \text{by } (iv) \text{ and } 14.13(iv), \\ \alpha^{\cdot(\beta \dot{+} \S\gamma)} &= \alpha^{\cdot\S(\beta \dot{+} \gamma)} = \alpha^{\cdot(\beta \dot{+} \gamma)} \bullet \alpha = (\alpha^{\cdot\beta} \bullet \alpha^{\cdot\gamma}) \bullet \alpha \\ &= \alpha^{\cdot\beta} \bullet (\alpha^{\cdot\gamma} \bullet \alpha) = \alpha^{\cdot\beta} \bullet \alpha^{\cdot\S\gamma}. \end{aligned}$$

Now assume that $\gamma = \bigcup \gamma \neq 0$. Let $\mu = \langle \beta \dot{+} \delta : \delta \in Ord \rangle$, $\nu = \langle \alpha^{\cdot\delta} : \delta \in Ord \rangle$, and $\xi = \langle \alpha^{\cdot\beta} \bullet \delta : \delta \in Ord \rangle$. By (*iii*), we may assume that $\alpha > 1$, so that ν is normal. Then, by (*iv*) and (*viii*), ξ is normal; and without any special conditions we know that μ is normal. Hence

$$\begin{aligned} \alpha^{\cdot(\beta \dot{+} \gamma)} &= \nu\mu\gamma = \bigcup_{\delta<\gamma} \nu\mu\delta \\ &= \bigcup_{\delta<\gamma} \alpha^{\cdot(\beta \dot{+} \delta)} = \bigcup_{\delta<\gamma} (\alpha^{\cdot\beta} \bullet \alpha^{\cdot\delta}) \\ &= \bigcup_{\delta<\gamma} \xi\nu\delta = \xi\nu\gamma = \alpha^{\cdot\beta} \bullet \alpha^{\cdot\gamma}. \end{aligned}$$

Next, (*xiii*) is true since $\alpha^{\cdot 1} = \alpha > 1$ and $\langle \alpha^{\cdot\beta} : \beta \in Ord \rangle$ is strictly increasing. Finally, (*xiv*) can be established by an easy induction on β.

Theorem 14.21 (*i*) $m^{\cdot n} \in \omega$.
(*ii*) $(l \cdot m)^{\cdot n} = l^{\cdot n} \cdot m^{\cdot n}$.

Proof Both statements are easily seen by induction on n.

Remark 14.22 For more detailed information on ordinal arithmetic see Bachmann 1967, Sierpinski 1965, and Tarski 1956.

EXERCISES

14.23 Prove 14.6(*iii*).
14.24 Prove 14.6(*v*).
14.25 For every $\beta \neq 0$ the function $\langle \alpha \dot{+} \beta : \alpha \in Ord \rangle$ is not limiting. *Hint:* There is a least $\alpha \geq \omega$ such that $\alpha \dot{+} \beta > \beta$; this α is a limit ordinal, and the function is "not limiting at α."
14.26 Show that $m \dot{+} \omega = \omega$.
14.27 Justify Definition 14.10.
14.28 Show that not every limit ordinal has the form $\alpha \cdot \omega$.
14.29 Justify Definition 14.18.
14.30 Show that there is a function $*$ such that, for all α, β, γ, with $\gamma = \bigcup \gamma \neq 0$,

(1) $\alpha^{*0} = \alpha$,
(2) $\alpha^{*\mathcal{S}\beta} = \alpha^{\cdot(\alpha * \beta)}$,
(3) $\alpha^{*\gamma} = \bigcup_{\delta < \gamma} \alpha^{*\delta}$.

Prove

(*a*) If $\alpha > 1$, then $\langle \alpha^{*\beta} : \beta \in Ord \rangle$ is a half-normal function.
(*b*) If $m > 1$, then $\alpha \geq \omega$ iff $m^{*\alpha} = \omega$.
(*c*) If $\alpha > 1$ and $\beta \geq \omega$, then $\alpha^{*\beta} = \alpha^{*\omega}$.

14.31 Show that $(2 \cdot 2)^{\cdot \omega} \neq 2^{\cdot \omega} \cdot 2^{\cdot \omega}$.
14.32 If $\delta \geq \omega$, $m \in \omega$, and $n \in \omega \sim 1$, then $(\delta \dot{+} m) \cdot n = \delta \cdot n \dot{+} m$.
14.33 If $\delta \neq 0$, then $\delta \cdot \omega = \mathcal{S}\delta \cdot \omega$.
14.34 For every $\beta > 1$ the function $\langle \alpha \cdot \beta : \alpha \in Ord \rangle$ is not limiting.
14.35 If δ is a limit ordinal, $n \in \omega$, and $m \in \omega \sim 2$, then $(\delta \dot{+} n)^{\cdot m} < \delta^m \cdot 2$.
14.36 If δ is a limit ordinal, then $(\delta \dot{+} n)^{\cdot \omega} = \delta^{\cdot \omega}$.
14.37 For every $\beta > 1$ the function $\langle \alpha^{\cdot \beta} : \alpha \in Ord \rangle$ is not limiting.

15 SPECIAL TOPICS

In this section we will give a brief survey of several somewhat advanced topics in the theory of ordinal numbers. We begin with a rather technical lemma and a theorem useful in what follows.

Lemma 15.1 *Suppose that* $\alpha = \gamma^{\cdot \delta_0} \cdot \varepsilon_0 \dot{+} \zeta_0$ *with* $\delta_0 \leq \alpha$, $0 < \varepsilon_0 < \gamma$, *and* $\zeta_0 < \gamma^{\cdot \delta_0}$ *and that* $\beta = \gamma^{\cdot \delta_1} \cdot \varepsilon_1 \dot{+} \zeta_1$ *with* $\delta_1 \leq \beta$, $0 < \varepsilon_1 < \gamma$, *and* $\zeta_1 < \gamma^{\cdot \delta_1}$.

Then $\alpha < \beta$ *iff one of the following three conditions holds:*
(*i*) $\delta_0 < \delta_1$.
(*ii*) $\delta_0 = \delta_1$ *and* $\varepsilon_0 < \varepsilon_1$.
(*iii*) $\delta_0 = \delta_1$, $\varepsilon_0 = \varepsilon_1$, *and* $\zeta_0 < \zeta_1$.

Proof If $\delta_0 < \delta_1$, then

$$\begin{aligned}\alpha &= \gamma^{\cdot\delta_0}\cdot\varepsilon_0 \dot{+} \zeta_0 < \gamma^{\cdot\delta_0}\cdot\varepsilon_0 \dot{+} \gamma^{\cdot\delta_0} \\ &= \gamma^{\cdot\delta_0}\cdot(\varepsilon_0 \dot{+} 1) \leq \gamma^{\cdot\delta_0}\cdot\gamma \\ &= \gamma^{\cdot(\delta_0\dot{+}1)} \leq \gamma^{\cdot\delta_1} \leq \gamma^{\cdot\delta_1}\cdot\varepsilon_1 \leq \beta.\end{aligned}$$

If $\delta_0 = \delta_1$ and $\varepsilon_0 < \varepsilon_1$, then

$$\begin{aligned}\alpha &= \gamma^{\cdot\delta_0}\cdot\varepsilon_0 \dot{+} \zeta_0 < \gamma^{\cdot\delta_0}\cdot\varepsilon_0 \dot{+} \gamma^{\cdot\delta_0} \\ &= \gamma^{\cdot\delta_0}\cdot(\varepsilon_0 \dot{+} 1) \leq \gamma^{\cdot\delta_0}\cdot\varepsilon_1 \leq \beta.\end{aligned}$$

Finally, if $\delta_0 = \delta_1$, $\varepsilon_0 = \varepsilon_1$, and $\zeta_0 < \zeta_1$, then obviously $\alpha < \beta$.

If, conversely, (*i*) to (*iii*) all fail, then either (1) $\delta_0 = \delta_1$, $\varepsilon_0 = \varepsilon_1$, and $\zeta_0 = \zeta_1$, so that $\alpha = \beta$ and $\alpha \not< \beta$, or (2) one of (*i*) to (*iii*) holds with 0 and 1 interchanged, so that $\beta < \alpha$, by the first part of the proof, and hence $\alpha \not< \beta$.

Lemma 15.1 is used in proving the following supplement to the division algorithm.

Theorem 15.2 *If* $\alpha > 0$ *and* $\beta > 1$, *then there exist unique* γ, δ, ε *such that* $\alpha = \beta^{\cdot\gamma}\cdot\delta \dot{+} \varepsilon$, $\gamma \leq \alpha$, $0 < \delta < \beta$, *and* $\varepsilon < \beta^{\cdot\gamma}$.

Proof We have $\beta^{\cdot 0} = 1 \leq \alpha$, and $\beta^{\cdot \mathcal{S}\alpha} \geq \mathcal{S}\alpha > \alpha$. Hence, by 12.13, choose γ such that $\beta^{\cdot\gamma} \leq \alpha < \beta^{\cdot\mathcal{S}\gamma}$. By the division algorithm, 14.16, choose δ, ε such that $\alpha = \beta^{\cdot\gamma}\cdot\delta \dot{+} \varepsilon$, with $\delta \leq \alpha$ and $\varepsilon < \beta^{\cdot\gamma}$. Clearly $\gamma \leq \alpha$ (see the first sentence of this proof). If $\delta = 0$, then $\alpha = \varepsilon < \beta^{\cdot\gamma}$, a contradiction. Finally, $\beta \leq \delta$ implies $\alpha \geq \beta^{\cdot\gamma}\cdot\delta \geq \beta^{\cdot\gamma}\cdot\beta = \beta^{\cdot\mathcal{S}\gamma}$, a contradiction; hence $\delta < \beta$. This proves existence, and uniqueness follows by 15.1.

Theorem 15.2 will be applied later in the decimal representations of ordinals. Now we find it useful in discussing certain peculiarities of ordinal arithmetic.

The integers obey cancellation laws with respect to the arithmetic operations: $m \dot{+} n = m \dot{+} p \Rightarrow n = p$, $m\cdot n = m\cdot p \wedge m \neq 0 \Rightarrow n = p$, and $m^{\cdot n} = m^{\cdot p} \wedge m \geq 2 \Rightarrow n = p$ (these facts are easily established by induction). For ordinals in general these facts no longer hold. There are even ordinals $\alpha \geq \omega$ such that $\beta \dot{+} \alpha = \alpha$ for all $\beta < \alpha$ (hence $1 \dot{+} \alpha = 2 \dot{+} \alpha$); we say then that α *additively absorbs* β. We want to give a fairly complete picture of which ordinals absorb others, with

respect to addition, multiplication, and exponentiation. Beginning with addition, we first exhibit many ordinals absorptive with respect to addition.

Theorem 15.3 *If* $\alpha < \omega^{\cdot\beta}$, *then* $\alpha \dot{+} \omega^{\cdot\beta} = \omega^{\cdot\beta}$.

Proof First we show

(1) If $\gamma < \beta$, then $\omega^{\cdot\gamma} \dot{+} \omega^{\cdot\beta} = \omega^{\cdot\beta}$.

Indeed, by 14.6(*ii*), choose δ such that $\beta = \gamma \dot{+} \delta$. Then

$$\begin{aligned} \omega^{\cdot\gamma} \dot{+} \omega^{\cdot\beta} &= \omega^{\cdot\gamma} \dot{+} \omega^{\cdot\gamma \cdot \omega^{\cdot\delta}} && \text{by 14.20(}xi\text{),} \\ &= \omega^{\cdot\gamma \cdot (1 \dot{+} \omega^{\cdot\delta})} && \text{by 14.13(}i\text{),} \\ &= \omega^{\cdot\gamma \cdot \omega^{\cdot\delta}} && \text{by 14.9,} \\ &= \omega^{\cdot\beta} && \text{by 14.20(}xi\text{).} \end{aligned}$$

Here we know that $\omega^{\cdot\delta} \geq \omega$ since $\delta \geq 1$ and $\omega^{\cdot 1} = \omega$, by 14.20(*v*). By an easy induction on m, we now have

(2) If $\gamma < \beta$, then $\omega^{\cdot\gamma \cdot m} \dot{+} \omega^{\cdot\beta} = \omega^{\cdot\beta}$ for all m.

Now we turn to the main part of the proof; hence suppose that $\alpha < \omega^{\cdot\beta}$. The case $\beta = 0$ is trivial, as is that of $\alpha < \omega$, by an obvious induction using 14.9. Hence assume that $\omega \leq \alpha$. Write $\alpha = \omega^{\cdot\gamma \cdot m} \dot{+} \delta$, by 15.2, with $m \neq 0$, $\delta < \omega^{\cdot\gamma}$. Then clearly $\gamma < \beta$. Hence

$$\begin{aligned} \omega^{\cdot\beta} \leq \alpha \dot{+} \omega^{\cdot\beta} &= (\omega^{\cdot\gamma \cdot m} \dot{+} \delta) \dot{+} \omega^{\cdot\beta} \leq \omega^{\cdot\gamma \cdot (m \dot{+} 1)} \dot{+} \omega^{\cdot\beta} \\ &&& \text{by 14.6(}iii\text{),} \\ &= \omega^{\cdot\beta} && \text{by (2).} \end{aligned}$$

This completes the proof.

Theorem 15.3 suggests the following.

Definition 15.4 α *is a* γ-***number*** *iff* $\beta \dot{+} \alpha = \alpha$ *for all* $\beta < \alpha$.

Thus α is a γ-number provided that it is a fixed point of every normal function $\langle \beta + \gamma : \gamma \in Ord \rangle$ with $\beta < \alpha$ (cf. 13.9). By 15.3, $\omega^{\cdot\beta}$ is always a γ-number, and obviously so are 0 and 1. These are the only ones:

Theorem 15.5 *The following three conditions are equivalent.*
(*i*) α *is a* γ-*number.*
(*ii*) *For all* $\beta, \gamma < \alpha$, $\beta \dot{+} \gamma < \alpha$.
(*iii*) $\alpha = 0$, *or* $\alpha = \omega^{\cdot\beta}$ *for some* β.

Proof (*i*) $\Rightarrow$ (*ii*) We have $\beta \dot{+} \gamma < \beta \dot{+} \alpha = \alpha$.

(*ii*) $\Rightarrow$ (*iii*) Assume that $\alpha \neq 0, 1$. Then, by (*ii*), $\omega \leq \alpha$. By 15.2, choose β, m, γ such that $\alpha = \omega^{\cdot\beta \cdot m} \dot{+} \gamma$ with $m \neq 0$ and $\gamma < \omega^{\cdot\beta}$. If

$\gamma \neq 0$, then $\omega^{\cdot\beta}\cdot m < \alpha$ and $\gamma < \omega^{\cdot\beta} \leq \omega^{\cdot\beta}\cdot m < \alpha$, and $\omega^{\cdot\beta}\cdot m \dot{+} \gamma = \alpha$, contradicting (*ii*). Thus $\gamma = 0$. If $m > 1$, then $m = n \dot{+} 1$ for some n, and $\omega^{\cdot\beta}\cdot n \dot{+} \omega^{\cdot\beta} = \alpha$, with $\omega^{\cdot\beta}\cdot n < \alpha$ and $\omega^{\cdot\beta} < \alpha$, again contradicting (*ii*). Thus $\alpha = \omega^{\cdot\beta}$.

(*iii*) $\Rightarrow$ (*i*) By 15.3.

Next we deal with absorption for multiplication.

Definition 15.6 *α is a δ-number iff for all β with $0 < \beta < \alpha$ we have $\beta\cdot\alpha = \alpha$.*

Theorem 15.7 *The following conditions are equivalent.*

(*i*) *α is a δ-number.*

(*ii*) *For all $\beta,\gamma < \alpha$, $\beta\cdot\gamma < \alpha$.*

(*iii*) *$\alpha \in \{0,1,2\}$, or $\alpha = \omega^{\cdot\omega^{\cdot\beta}}$ for some β.*

Proof (*i*) $\Rightarrow$ (*ii*) Assuming that β, $\gamma < \alpha$, we have $\beta \neq 0$ and $\beta\cdot\gamma < \beta\cdot\alpha = \alpha$ or $\beta = 0$ and $\beta\cdot\gamma = 0 < \alpha$.

(*ii*) $\Rightarrow$ (*iii*) Assume that $\alpha \notin \{0,1,2\}$. Then $\alpha \notin \omega$, since otherwise $\alpha = m \dot{+} 1$ for some $m \geq 2$ and then $m\cdot m \geq \alpha$. This fact, that $m\cdot m \geq m \dot{+} 1$ for every $m \geq 2$, is easily established.

Hence $\omega \leq \alpha$. Next

(1) α is a γ-number.

Indeed, suppose that β, $\gamma < \alpha$; we want to show that $\beta \dot{+} \gamma < \alpha$. We may assume that $1 < \beta$, γ. Then $\beta \dot{+} \gamma \leq \beta\cdot\gamma < \alpha$, using 14.13(*x*). Thus (1) holds. Hence $\alpha = \omega^{\cdot\gamma}$ for some γ. If δ, $\varepsilon < \gamma$, then $\omega^{\cdot\delta}$, $\omega^{\cdot\varepsilon} < \omega^{\cdot\gamma}$, so that $\omega^{\cdot\delta}\cdot\omega^{\cdot\varepsilon} < \omega^{\cdot\gamma}$; that is, $\omega^{\cdot(\delta\dot{+}\epsilon)} < \omega^{\cdot\gamma}$. This implies that $\delta \dot{+} \varepsilon < \gamma$. Hence γ is a γ-number. Say $\gamma = \omega^{\cdot\beta}$; then $\alpha = \omega^{\cdot\omega^{\cdot\beta}}$.

(*iii*) $\Rightarrow$ (*i*) It suffices to show that $\omega^{\cdot\omega^{\cdot\beta}}$ is a δ-number. Assume that $0 < \gamma < \omega^{\cdot\omega^{\cdot\beta}}$. If $\gamma < \omega$, then $\gamma\cdot\omega^{\cdot\omega^{\cdot\beta}} = \omega^{\cdot\omega^{\cdot\beta}}$, by 14.17(*iii*)—$\omega^{\cdot\omega^{\cdot\beta}}$ is easily seen to be a limit ordinal by transfinite induction on β. Thus assume that $\omega \leq \gamma$. Since $\gamma < \omega^{\cdot\omega^{\cdot\beta}}$, this implies that $\beta \neq 0$, and hence $\omega^{\cdot\beta}$ is a limit ordinal (this last statement is easily seen by induction on β). By 15.2, choose δ, m, ε such that $m \neq 0$, $\varepsilon < \omega^{\cdot\delta}$, and $\gamma = \omega^{\cdot\delta}\cdot m \dot{+} \varepsilon$. Then $\delta < \omega^{\cdot\beta}$, hence $\delta \dot{+} 1 < \omega^{\cdot\beta}$, and

$$\begin{aligned}
\omega^{\cdot\omega^{\cdot\beta}} \leq \gamma\cdot\omega^{\cdot\omega^{\cdot\beta}} &= (\omega^{\cdot\delta}\cdot m \dot{+} \varepsilon)\cdot\omega^{\cdot\omega^{\cdot\beta}} \\
&\leq (\omega^{\cdot\delta}\cdot m \dot{+} \omega^{\cdot\delta})\cdot\omega^{\cdot\omega^{\cdot\beta}} \\
&= \omega^{\cdot\delta}\cdot(m \dot{+} 1)\cdot\omega^{\cdot\omega^{\cdot\beta}} \\
&\leq \omega^{\cdot(\delta\dot{+}1)}\cdot\omega^{\cdot\omega^{\cdot\beta}} \\
&= \omega^{\cdot((\delta\dot{+}1)\dot{+}\omega^{\cdot\beta})} = \omega^{\cdot\omega^{\cdot\beta}} \qquad \text{by 15.3.}
\end{aligned}$$

This completes the proof.

Finally, we can ask about fixed points for exponentiation.

Definition 15.8 *α is an ε-number iff $\beta^{\cdot\alpha} = \alpha$ for all β such that $1 < \beta < \alpha$.*

We do not have a precise arithmetical expression for ε-numbers like 15.5(*iii*) for γ-numbers and 15.7(*iii*) for δ-numbers. We can give an analog of 15.5(*ii*) and 15.7(*ii*), but first we need the following.

Theorem 15.9 *Every ε-number is a δ-number.*

Proof Assume that α is an ε-number and that $0 < \beta < \alpha$. If $\beta = 1$, then $\beta \cdot \alpha = \alpha$. If $1 < \beta$, then, by 14.20(*xiv*), $\alpha = 1 \cdot \alpha \leq \beta \cdot \alpha \leq \beta^{\cdot\alpha} = \alpha$, so that $\beta \cdot \alpha = \alpha$. Thus α is a δ-number.

Theorem 15.10 *The following conditions are equivalent.*
(*i*) *α is an ε-number.*
(*ii*) *$\alpha = 1$, or for all $\beta, \gamma < \alpha$, $\beta^{\cdot\gamma} < \alpha$.*

Proof (*i*) $\Rightarrow$ (*ii*) Clearly we may assume that $\alpha \neq 0, 1, 2$. If $\beta = 0$, then $\beta^{\cdot\gamma} \leq 1 < \alpha$; if $\beta = 1$, then $\beta^{\cdot\gamma} = 1 < \alpha$; and if $1 < \beta$, then $\beta^{\cdot\gamma} < \beta^{\cdot\alpha} = \alpha$, so that (*ii*) holds.

(*ii*) $\Rightarrow$ (*i*) Again it is clear that we may assume that $\alpha \neq 0, 1, 2$. Note from (*ii*) and 14.20(*xiv*) that

(1) α is a δ-number.

Suppose that $1 < \beta < \alpha$. If $\alpha = \S\gamma$ for some γ, then $\alpha \leq \beta^{\cdot\alpha} = \beta^{\cdot\S\gamma} = \beta^{\cdot\gamma} \cdot \beta < \alpha$, by (*ii*) and (1), which is impossible. Thus α is a limit ordinal, and so $\alpha \leq \beta^{\cdot\alpha} = \bigcup_{\gamma<\alpha} \beta^{\cdot\gamma} \leq \alpha$, by (*ii*), so that $\alpha = \beta^{\cdot\alpha}$, as desired.

Theorem 15.11 *0, 1, 2, and ω are ε-numbers, and these are the only ε-numbers $\leq \omega$.*

Theorem 15.11 is easily shown, using 15.9 and 15.7(*iii*). The following theorem enables one to manufacture larger and larger ε-numbers. In Exercise 21.28 we shall see further that every infinite cardinal number is, as an ordinal, an ε-number (and hence also a δ-number and a γ-number).

Theorem 15.12 *For each $\alpha > 1$, the least ε-number $> \alpha$ is the ordinal $\nu\omega$, where ν is recursively defined as follows: $\nu 0 = \alpha$, $\nu(m \dotplus 1) = \nu m^{\cdot\nu m}$ for $m \in \omega$, and $\nu\omega = \bigcup_{m\in\omega} \nu m$.*

Proof First, let β be any ε-number $> \alpha$. It is easily shown, by induction on m, that $\nu m < \beta$ for every $m \in \omega$, making use of 15.10(*ii*). Hence $\nu\omega \leq \beta$. Second, we have to show that $\nu\omega$ is itself an ε-number. Suppose that $\gamma, \delta < \nu\omega$. Then there exist $m, n \in \omega$ such that $\gamma < \nu m$ and $\delta < \nu n$. Say $m \leq n$. Thus

$$\gamma^{\cdot\delta} \leq \nu m^{\cdot\nu n} \leq \nu n^{\cdot\nu n} = \nu(n \dotplus 1) < \nu(n \dotplus 2) \leq \nu\omega.$$

This completes the proof.

If we apply 15.12 to $\alpha = \omega$, then we get the first ε-number $> \omega$. In this case $\nu 0 = \omega$, $\nu 1 = \omega^{\cdot\omega}$, and

$$\begin{aligned}\nu 2 &= (\omega^{\cdot\omega})^{\cdot\omega^{\cdot\omega}} = \omega^{\cdot(\omega\cdot(\omega^{\cdot\omega}))} \qquad \text{by } 14.20(xii)\\ &= \omega^{\cdot(\omega^{\cdot\omega})},\end{aligned}$$

since $\omega^{\cdot\omega}$ is a δ-number. In general, one can show by induction that $\nu m = \omega^{\cdot\omega^{\cdot^{\cdot^{\cdot\omega}}}}$, where $m + 1$ ω's occur on the right-hand side. Hence the first ε-number $> \omega$ is naturally written as $\omega^{\cdot\omega^{\cdot^{\cdot}}}$, and it is customary to call this number ε_0. In many mathematical situations transfinite induction up to ε_0 is called for; we may mention, for example, Gentzen's proof of the consistency of arithmetic (see Gentzen 1936). The reason that ε_0 plays such a central role in arguments by transfinite induction is connected with the representation of ordinals in *Cantor normal form*. We will now discuss this normal form and then indicate informally why ε_0 is a "critical number" for transfinite inductions.

Theorem 15.15 shows that any ordinal can be expressed in "decimal form" with respect to any ordinal $\beta > 1$ as base; Cantor normal form corresponds to the choice $\beta = \omega$, and for natural numbers with $\beta = 10$ we get the usual school form. To formulate this theorem, we need the notion of a finite sum of ordinals.

Definition 15.13 *If $\mu \in {}^m Ord$ for some m, and n is any integer, we define $\dot{\sum}(\mu,n)$ by recursion as follows:*

$$\dot{\sum}(\mu,0) = 0;$$

$$\dot{\sum}(\mu, n \dot{+} 1) = \begin{cases} \dot{\sum}(\mu,n) \dot{+} \mu n & \text{if } m > n; \\ 0 & \text{otherwise.} \end{cases}$$

We write $\dot{\sum}_{i<n} \mu i$ instead of $\dot{\sum}(\mu,n)$.

Theorem 15.14 (*Associative law*) *If $\mu \in {}^m Ord$, and $n < m$, then $\dot{\sum}_{i<m} \mu i = \dot{\sum}_{i<n} \mu i \dot{+} \dot{\sum}_{n<i<m} \mu i$.*

Note that $\dot{\sum}_{n\le i<m} \mu i$ is to be interpreted as $\dot{\sum}_{i<p} \nu i$, where p is the unique integer such that $n \dot{+} p = m$ [see 14.6(*ii*)] and $\nu = \langle \mu(n \dot{+} i) : i < p\rangle$. Theorem 15.14 is then easily shown, by induction on p. Theorem 15.14 is not the most general associative law (see Exercise 15.31 at the end of the section), but it is sufficient for most purposes.

Now we can state the base-expansion theorem.

Theorem 15.15 (*Base-expansion theorem*) *Let* $\beta > 1$. *Then for any ordinal* α *there exist unique* m, γ, δ *satisfying the following conditions.*
(*i*) $\gamma, \delta \in {}^{m}Ord$.
(*ii*) $\alpha = \dot{\sum}_{i<m} \beta^{\cdot\gamma_i \cdot \delta_i}$.
(*iii*) *For all* $i < m$, $\alpha \geq \gamma_i$, *and if* $i \dot{+} 1 < m$, then $\gamma_i > \gamma_{i\dot{+}1}$.
(*iv*) *For all* $i < m$, $0 < \delta_i < \beta$.

Proof If $\alpha = 0$, let $m = \gamma = \delta = 0$. Now assume that $\alpha > 0$. We define γ°, δ°, ε by recursion. Let $\varepsilon_0 = \alpha$. If ε_m has been defined, apply Theorem 15.2, in case $\varepsilon_m \neq 0$, to obtain γ°_m, δ°_m, $\varepsilon_{m\dot{+}1}$ such that

(1) If $\varepsilon_m \neq 0$; then $\varepsilon_m = \beta^{\cdot\gamma_m^\circ \cdot \delta^\circ_m} \dot{+} \varepsilon_{m\dot{+}1}$, with $\gamma^\circ_m \leq \varepsilon_m$, $0 < \delta^\circ_m < \beta$, and $\varepsilon_{m\dot{+}1} < \beta^{\cdot\gamma_m^\circ}$; if $\varepsilon_m = 0$, then $\varepsilon_{m\dot{+}1} = \gamma^\circ_m = \delta^\circ_m = 0$.

From (1) it follows that, if $\varepsilon_m > 0$, then $\varepsilon_m > \varepsilon_{m\dot{+}1}$. Hence, by 11.6, $\varepsilon_m = 0$ for some m. Let m be chosen minimal with this property. Let $\gamma = \gamma^\circ \restriction m$, $\delta = \delta^\circ \restriction m$. Thus (*i*) holds; by (1), (*iv*) and the first part of (*iii*) hold. Since $\varepsilon_i > \varepsilon_{i\dot{+}1}$ for $i \dot{+} 1 < m$, (1) and 15.1(*i*) imply that $\gamma_{i\dot{+}1} \leq \gamma_i$. If $\gamma_{i\dot{+}1} = \gamma_i$, then

$$\varepsilon_{i\dot{+}1} = \beta^{\cdot\gamma_i \cdot \delta_{i\dot{+}1}} \dot{+} \varepsilon_{i\dot{+}2} \geq \beta^{\cdot\gamma_i} > \varepsilon_{i\dot{+}1} \qquad \text{by (1)},$$

a contradiction. Thus $\gamma_{i\dot{+}1} < \gamma_i$, so that (*iii*) holds. Now, by induction on i, it is easily seen that

(2) For all $i \leq m$, $\varepsilon_0 = \dot{\sum}_{j<i} \beta^{\cdot\gamma_j \cdot \delta_j} \dot{+} \varepsilon_i$.

With $i = m$ in (2) we obtain (*ii*). This finishes the existence proof. As to uniqueness, suppose that $\hat{m}$, $\hat{\gamma}$, $\hat{\delta}$ also satisfy the conditions of the theorem. By induction on p, one easily proves

(3) For all i with $i \dot{+} 1 < \hat{m}$, if $(i \dot{+} 1) \dot{+} p = \hat{m}$, then

$$\dot{\sum}_{i+1<j<\hat{m}} (\hat{\beta}^{\cdot\hat{\gamma}_i \cdot \hat{\delta}_j}) < \hat{\beta}^{\cdot\hat{\gamma}_i}.$$

Now one can use (3) and 15.2 to prove, by induction on i, that for all $i < m$, $\hat{m} > i$ and $\gamma_i = \hat{\gamma}_i$, $\delta_i = \hat{\delta}_i$. Uniqueness now easily follows. This completes the proof.

If we take $\alpha \in \omega$ and $\beta = 10$ in 15.15, we get the usual decimal expansion of α. Here it is customary to make a certain sequence μ correspond to α. The domain of μ is $n = (\bigcup_{i<m} \gamma_i) \dot{+} 1$, and for each $j < n$, μj is 0 if $j \notin Rng\ \gamma$ and $\mu\gamma_i = \delta_i$ for every $i < m$. Thus $309 = 3 \cdot 10^{\cdot 2} \dot{+} 9 \cdot 10^0$; in this case, Theorem 15.15 gives $m = 2$, $\gamma = \langle 2,0 \rangle$, and $\delta = \langle 3,9 \rangle$. In this way we recapture an important part of elementary mathematics in a rigorous framework. It is not too difficult to prove

the usual rules for "carrying" in addition and for multiplying; we leave this for an exercise.

By 15.15, any ordinal can be expressed to base 10. Thus $\omega = 10^{\cdot\omega}$, $\omega \dot{+} 13 = (10^{\cdot\omega} \dot{+} 10) \dot{+} 3$, $\omega \cdot 2 = 10^{\cdot\omega} \cdot 2$. For infinite ordinals it is more interesting to use base ω; in this case, the representation given in Theorem 15.15 is called the *Cantor normal form* of α:

$$\alpha = \dot{\sum}_{i<m} (\omega^{\cdot\gamma_i} \cdot n_i),$$

with $\alpha \geq \gamma_0 > \gamma_1 > \cdots > \gamma_{m-1}$ and $n_i \neq 0$ for all $i < m$. One can again express the exponents $\gamma_0, \ldots, \gamma_{m-1}$ in Cantor normal form, and express their exponents in Cantor normal form, and so forth. An important property of ε_0, the first ε-number $> \omega$, is that for $\alpha < \varepsilon_0$ this process must stop after finitely many steps. Thus with each ordinal $< \varepsilon_0$ one can associate a certain configuration of natural numbers, so that, in a sense, up to ε_0 we have not gone beyond the integers. In particular, transfinite induction up to ε_0 does not, loosely speaking, transcend ordinary induction over integers.

Turning to another topic, we will now discuss the important notion of the *rank* of sets. Roughly speaking, we want to assign an ordinal ρx to each set x in such a way that the magnitude of ρx measures the complexity of x. Our only criterion for complexity is that x is more complex than each of its members.

Definition 15.16 *ρ is the unique function with domain* V *such that for any set x,*

$$\rho x = \bigcap \{\alpha : \rho y < \alpha \text{ for each } y \in x\}.$$

We call ρx the ***rank*** *of the set x.*

Thus ρx is the least ordinal $> \rho y$ for each $y \in x$. Definition 15.16 is easily justified, using the general recursion theorem, since $\{(x,y) : x \in y\}$ is a well-founded relation. Note that 0 has rank 0, 1 has rank 1, $\{1\}$ has rank 2. Some useful properties of the rank function are the following.

Theorem 15.17 *(i) If $x \in y$, then $\rho x < \rho y$.*
(ii) If $x \subseteq y$, then $\rho x \leq \rho y$.
(iii) $\rho(x \cup y) = \rho x \cup \rho y$.
(iv) $\rho(\bigcup x) = \bigcup \{\rho y : y \in x\} \leq \rho x$.
(v) $\rho\{x\} = \rho x \dot{+} 1$.
(vi) $\rho(Sx) = \rho x \dot{+} 1$.
(vii) $\rho\alpha = \alpha$.

Proof Conditions *(i)*, *(ii)*, and *(v)* are obvious, and *(vii)* is easily estab-

lished by transfinite induction on α. For (*iii*), the inequality $\rho x \cup \rho y \leq \rho(x \cup y)$ follows by (*ii*). If $z \in x \cup y$, then either $z \in x$ and hence $\rho z < \rho x \leq \rho x \cup \rho y$, or $z \in y$ and hence $\rho z < \rho y \leq \rho x \cup \rho y$. Therefore $\rho z < \rho x \cup \rho y$ whenever $z \in x \cup y$, so that the inequality $\rho(x \cup y) \leq \rho x \cup \rho y$ also holds, and (*iii*) is established. To prove (*iv*), first note that for any $z \in \bigcup x$ we have $z \in y \in x$ for some y, and hence, by (*i*), $\rho z < \rho y \leq \bigcup\{\rho y : y \in x\}$. Hence $\rho z < \bigcup\{\rho y : y \in x\}$ for each $z \in \bigcup x$, so that $\rho(\bigcup x) \leq \bigcup\{\rho y : y \in x\}$. Obviously $\bigcup\{\rho y : y \in x\} \leq \rho x$. Finally, if $z \in x$, then $z \subseteq \bigcup x$ and hence $\rho z \leq \rho(\bigcup x)$. Therefore, $\bigcup\{\rho y : y \in x\} = \rho(\bigcup x)$. As to (*vi*), since $x \in Sx$, we have $\rho x \dot{+} 1 \leq \rho(Sx)$. On the other hand, if $y \in Sx$, then $y \subseteq x$ and hence, by (*ii*), $\rho y \leq \rho x < \rho x \dot{+} 1$; it follows that $\rho(Sx) \leq \rho x \dot{+} 1$. This completes the proof.

Theorem 15.18 *For every* α, $\{x : \rho x < \alpha\}$ *is a set, and*

$$\{x : \rho x < \alpha\} = \bigcup_{\beta<\alpha} S\{x : \rho x < \beta\}.$$

Proof We proceed by induction on α. Suppose that the result is known for all $\beta < \alpha$. Thus, in particular, $S\{x : \rho x < \beta\}$ is a set for each $\beta < \alpha$, and $\bigcup_{\beta<\alpha} S\{x : \rho x < \beta\}$ is a set. Hence it is enough to prove the equation in 15.18. If $y \in \bigcup_{\beta<\alpha} S\{x : \rho x < \beta\}$, then there is a $\beta < \alpha$ such that $y \in S\{x : \rho x < \beta\}$; that is, $y \subseteq \{x : \rho x < \beta\}$. Therefore, $\rho x < \beta$ for each $x \in y$, and it follows that $\rho y \leq \beta < \alpha$. This establishes one desired inclusion. Now suppose that $\rho y < \alpha$, and let $\beta = \rho y$. For any $x \in y$ we have $\rho x < \rho y = \beta$, by 15.17(*i*); hence $y \subseteq \{x : \rho x < \beta\}$, and hence $y \in S\{x : \rho x < \beta\}$ and so $y \in \bigcup_{\beta<\alpha} S\{x : \rho x < \beta\}$. This establishes the other inclusion, and the proof is complete.

Definition 15.19 *For every* α, $M_\alpha = \{x : \rho x < \alpha\}$.

Theorem 15.20 (*i*) $M_\alpha = \bigcup_{\beta<\alpha} SM_\beta$.
(*ii*) $\rho(M_\alpha) = \alpha$.
(*iii*) *If* $x \in M_\alpha$, *then* $x \subseteq M_\alpha$.
(*iv*) *If* $\alpha < \beta$, *then* $M_\alpha \subset M_\beta$ *and* $M_\alpha \in M_\beta$.
(*v*) $SM_\alpha = M_{\alpha+1}$.
(*vi*) *If* α *is a limit ordinal, then* $M_\alpha = \bigcup_{\beta<\alpha} M_\beta$.
(*vii*) $V = \bigcup_{\alpha \in Ord} M_\alpha$.

Proof (*i*), (*iii*), and (*vii*) are easily established, using the previous results on ρ. As to (*ii*), if $x \in M_\alpha$, then $\rho x < \alpha$, by the definition of M_α; hence $\rho(M_\alpha) \leq \alpha$, by the definition of ρ. If $\beta < \alpha$, then $\beta = \rho\beta < \rho\alpha = \alpha$, by 15.17(*vii*), and so $\beta \in M_\alpha$. Thus $\alpha \subseteq M_\alpha$, so that $\alpha = \rho\alpha \leq \rho(M_\alpha)$, by 15.17(*ii*). Therefore (*ii*) holds. (*iv*) follows immediately from (*ii*). By (*iv*), $\bigcup_{\beta<\alpha+1} SM_\beta = SM_\alpha$, so that (*v*) follows, using (*i*). (*vi*) is immediate from (*v*) and (*i*), and this completes the proof.

Theorem 15.20 has a fundamental significance for the philosophy of set theory. The following facts follow from 15.20. We have $M_0 = 0$, $M_1 = S0$, $M_2 = SM_1, \ldots, M_\omega = \bigcup_{m\in\omega} M_m$, $M_{\omega+1} = SM_\omega, \ldots$. Every set appears in some M_α. Thus we may say that all sets are built up by a simple inductive procedure from the empty set using just two operations: forming the set of all subsets, and taking the union of preceding steps.

One of the main uses of the rank function ρ is in justifying definitions by abstraction. As indicated in Sec. 7, one frequently wishes to "identify" certain elements of some class K. Rigorously, an equivalence relation R with field K is given, and we want to "identify" equivalent elements of K. If K is a set, K/R does the trick. If K is a proper class, this no longer works. Using the rank function, however, we can also take care of this case.

Definition 15.21 *If R is an equivalence relation with field K, then for any $x \in K$ let $\tau_R x = \{y : xRy \text{ and } \rho y = \bigcap\{\rho z : zRx\}\}$. $\tau_R x$ is called the R-**type** of x.*

Theorem 15.22 *Under the assumptions of* 15.21, *$\tau_R x$ is a set, and $\tau_R x \neq 0$. Furthermore, for any $y,z \in K$ we have yRz iff $\tau_R y = \tau_R z$.*

Proof Let $\alpha = \bigcap\{\rho z : zRx\}$. Thus $\rho y = \alpha$ for each $y \in \tau_R x$, so that $\tau_R x \subseteq M_{\alpha+1}$. $M_{\alpha+1}$ is a set, by Theorem 15.18, so that $\tau_R x$ is a set. Now α is the least element of $\{\rho z : zRx\}$. Hence there is a z such that $\alpha = \rho z$ and zRx. Thus $z \in \tau_R x$, and so $\tau_R x \neq 0$.

Now assume that yRz. Then, for any u, uRy iff uRz, since R is an equivalence relation. It follows that $\bigcap\{\rho u : uRy\} = \bigcap\{\rho u : uRz\}$ and $\tau_R y = \tau_R z$. On the other hand, suppose that $\tau_R y = \tau_R z$. Choose $u \in \tau_R y$. Then uRy and uRz, so that yRz, as desired.

Theorem 15.22 expresses the essential properties of R-types. As a special case we obtain *order types*. Let R consist of all pairs (S,T) such that $S,T \in V$ and S and T are isomorphic simple orderings (see 8.10). It is easily checked that R is an equivalence relation. For this R we call R-types *order types*. We could have taken ordinals to be the order types of well-orderings and developed the whole theory of ordinals on this basis.

Remark 15.23 For advanced ordinal arithmetic see Bachmann 1967, Sierpinski 1965, and Tarski 1956. Definition 15.21 is essentially due to Dana Scott.

EXERCISES

Prove the following.

15.24 For any $\alpha > 0$, $\alpha \cdot \omega$ is the smallest γ-number greater than α.

15.25 If α is a γ-number, then $\alpha^{\cdot\beta}$ is a γ-number.

15.26 If β is a limit number, then $\alpha^{\cdot\beta}$ is a γ-number.

15.27 For any $\alpha > 1$, $\alpha^{\cdot\omega}$ is the smallest δ-number greater than α.

15.28 If α is a δ-number and β is a γ-number, then $\alpha^{\cdot\beta}$ is a δ-number.

15.29 For any $\alpha > 2$, α is an ε-number iff $\alpha = 2^{\cdot\alpha}$.

15.30 Justify Definition 15.13 by using an appropriate recursion theorem.

15.31 Suppose that $\mu \in {}^{m}Ord$, $\nu \in {}^{n}m$, $m > 1$, $n > 1$, ν is strictly increasing, $\nu 0 = 0$, and $\nu(n-1) = m - 1$. Then

$$\dot{\sum}_{i<m} \mu_i = \left[\dot{\sum}_{i<n-2} \dot{\sum}_{\nu i \leq j < \nu(i\dot{+}1)} \mu_j\right] \dot{+} \dot{\sum}_{\nu(n-2) \leq j \leq \nu(n-1)} \mu_j$$

15.32 Justify the recursive definition given in the proof of 15.15.

15.33 Prove the usual rules for "carrying" in addition of natural numbers, and the usual rules of multiplying.

15.34 Write the following in Cantor normal form:

(*a*) $[(\omega \cdot 3 \dot{+} \omega^{\cdot 7}) \cdot \omega^{\cdot 3}] \cdot \omega^{\cdot 2}$.

(*b*) $[\omega^{\cdot 8} \dot{+} \omega^{\cdot(\omega \cdot 3)}] \cdot (\omega \dot{+} 1)$.

15.35 Give convenient rules for adding, multiplying, and exponentiating numbers in Cantor normal form.

15.36 Show that addition, $\oplus$, of order types can be defined so that if R is of type α and S of type β, then $\alpha \oplus \beta$ is the order type of

$$\{((a,0),(b,0)) : aRb\} \cup \{((a,0),(x,1)) : a \in Fld\ R, x \in Fld\ S\} \cup \{((x,1),(y,1)) : xSy\}.$$

3
The Axiom of Choice

In this short chapter we give various equivalent forms of the axiom of choice and show some typical applications.

16 EQUIVALENTS OF THE AXIOM OF CHOICE

In Axiom 1.36, we gave the relational axiom of choice. This postulate is a very strong form of the axiom of choice that is frequently convenient to have in our system of set theory. However, for most applications a somewhat weaker axiom suffices; this weaker form has more intuitive appeal, and is more widely known, than Axiom 1.36. We will refer to it as the *axiom of choice*, although it is not really an axiom of our system and is not equivalent to 1.36, which we will always refer to as the *relational axiom of choice.*

Axiom of choice For any set A of nonempty sets, there is a function F with domain A such that $Fx \in x$ for every $x \in A$.

The function F is called a *choice function* for A. Note that, speaking intuitively, the existence of F implies a simultaneous choice of elements for each member of A, and A may well have infinitely many members. The axiom does not imply that F can be constructed in any practical sense. Historically, considerations of this kind have led to controversy about the advisability of using the axiom of choice, but, as stated earlier, it seems to be a generally accepted principle now. For further discussion see Fraenkel, Bar-Hillel 1958.

The following theorem comes as no surprise.

Theorem 16.1 *The axiom of choice holds.*

Proof Let A be any set of nonempty sets. Set $R = \{(x,y) : y \in x \in A\}$. Since each member of A is nonempty, $Dmn\ R = A$. By the relational axiom of choice, let F be a function such that $F \subseteq R$ and $Dmn\ F = Dmn\ R$. Thus F is a choice function for A.

We now give what we consider to be the most useful principles equivalent to the axiom of choice.

Multiplicative principle If A is a function with domain $I \in V$ and if $A_i \neq 0$ for every $i \in I$, then $\bigtimes_{i \in I} A_i \neq 0$.

Zermelo's principle If P is a partition of a set A, then there is a $B \subseteq A$ such that $B \cap M$ has exactly one element for every $M \in P$.

This formulation of Zermelo's principle only apparently involves the notion of a cardinal number (which we have not yet introduced), namely, the cardinal number 1. The last phrase in the statement of the principle can be restated as follows:

There is an $x \in B \cap M$ such that, for all $y \in B \cap M$, $x = y$.

Counting principle For every set A there is an ordinal α and a one-one function F that maps α onto A.

Well-ordering principle For every set A there is a well-ordering with field A.

Zorn's lemma If a set A is partially ordered by a relation R, and every subset of A simply ordered by R has an R-upper bound in A, then A has an R-maximal element.

Zorn's lemma should really be called *Kuratowski's lemma.* Note that the hypothesis of Zorn's lemma implies that $A \neq 0$, since the empty subset of A must have an R-upper bound in A. Any element of A is an R-upper bound of the empty set. Thus Zorn's lemma can be equivalently formulated as follows:

If a nonempty set A is partially ordered by a relation R, and every nonempty subset of A simply ordered by R has an R-upper bound in A, then A has an R-maximal element.

Also recall that A partially ordered by a relation R means that $(A \times A) \cap R$ is a partial ordering. Clearly we may assume that R itself is a partial ordering, and that $Fld\ R = A$.

Maximality principle If A is a set of sets, and every subset of A simply ordered by inclusion has an $\subseteq$-upper bound in A, then A has an $\subseteq$-maximal element.

We note at once that the maximality principle is simply a special case of Zorn's lemma, and that frequently only the maximality principle is applied when Zorn's lemma is cited in a proof.

Kuratowski's principle If R is a partial ordering, $R \in V$, and S is a simple ordering such that $S \subseteq R$, then there is an $\subseteq$-maximal simple ordering T such that $S \subseteq T \subseteq R$.

Trichotomy principle For any two sets A and B, there is a one-one mapping either from A into B or from B into A.

The trichotomy principle does not, on the face of it, involve three alternatives. However, if it holds, then one of the following three conditions holds, for any sets A and B:

1 There is a one-one function mapping A into B, but none mapping B into A.
2 There is a one-one function mapping B into A, but none mapping A into B.
3 There is a one-one function mapping A into B, and also one mapping B into A.

The trichotomy principle has played an important role in the development of the theory of cardinal numbers. Condition 3 is equivalent to the assertion that there is a one-one function mapping A onto B, and this equivalence can be proved without using the axiom of choice (see Chap. 4).

Mapping principle For any two nonempty sets A and B, there is a function mapping either A onto B or B onto A.

Theorem 16.2 *All the preceding principles are equivalent to the axiom of choice (this we prove, of course, within the set theory based on all the axioms except the relational axiom of choice). They are all valid under the assumption of* 1.36, *the relational axiom of choice.*

Proof The second assertion of the theorem is an obvious consequence of the first part and Theorem 16.1. For the first assertion, we treat the multiplicative principle and Zermelo's principle separately, since both are easily seen to be equivalent to the axiom of choice. For each implication we assume the notation in the definition of the principle to be established.

Axiom of choice $\Rightarrow$ multiplicative principle $\{A_i : i \in I\}$ is a set of nonempty sets (it is a set by the substitution axiom), so, by the choice principle, let F be a choice function for it. For each $i \in I$ let $Gi = FA_i$. Then clearly $G \in \mathsf{P}_{i \in I} A_i$.

Multiplicative principle $\Rightarrow$ Zermelo's principle Let $C = I\restriction P$. Since $P \subseteq SA$, P is a set; thus C is a function whose domain is a set. By the multiplicative principle, choose $f \in \mathsf{P}_{M \in P} C_M = \mathsf{P}_{M \in P} M$. Clearly $Rng\, f \subseteq A$ and $Rng\, f \cap M$ has exactly one element for every $M \in P$.

Zermelo's principle $\Rightarrow$ axiom of choice Let $K = \{(x,y) : y \in x \in A\}$, and let $P = \{\{(x,y) : y \in x\} : x \in A\}$. Thus P is a partition of K. Choose $F \subseteq K$ such that $F \cap M$ has exactly one element for every $M \in P$. It is then easily checked that F is a choice function for A.

We prove the equivalence of all the other principles with the axiom of choice by one big circle of implications.

Axiom of choice $\Rightarrow$ counting principle Let f be a choice function for $SA \sim \{0\}$. By 2.6(i), choose $a \notin A$. By recursion, there is a function H with domain Ord such that for every α,

$$H\alpha = \begin{cases} f(A \sim H^*\alpha) & \text{if } A \sim H^*\alpha \neq 0, \\ a & \text{otherwise.} \end{cases}$$

(Compare the proof of Theorem 13.10.) Now

(1) If $\alpha < \beta$ and $H\alpha = a$, then $H\beta = a$.

Indeed, $H^*\alpha \subseteq H^*\beta$, so that, if $H\alpha = a$, then $A \sim H^*\alpha = 0$ and so

$A \sim H^*\beta = 0$ and $H\beta = a$. Also

(2) If $\alpha < \beta$ and $H\beta \neq a$, then $H\alpha \neq H\beta$.

For $H\alpha \in H^*\beta$ and $H\beta \in A \sim H^*\beta$, so that $H\alpha \neq H\beta$.

By (2), there is an α such that $H\alpha = a$, since otherwise H would be a one-one function from *Ord* into A and so *Ord* would be a set, by 4.11, contradicting 9.8. Choosing α minimal such that $H\alpha = a$, it is clear, by (2), that $H\restriction\alpha$ is a one-one function mapping α onto A, as desired.

Counting principle $\Rightarrow$ well-ordering principle Choose α, f such that f is a one-one function mapping α onto A. Let $R = \{(a,b) : a,b \in A \text{ and } f^{-1}a \leq f^{-1}b\}$. It is easily checked that R is a well-ordering with field A.

Well-ordering principle $\Rightarrow$ Zorn's lemma We may assume that $Fld\ R = A$. Let S be a well-ordering with field A. By 2.6(i), choose $d \notin A$. By recursion, there is a function F with domain *Ord* such that, for any α,

$$F\alpha = \begin{cases} S\text{-least element of } \{a : bRa \in A \\ \text{and } b \neq a, \text{ for every } b \in F^*\alpha\} & \text{if this set is nonempty,} \\ d & \text{otherwise.} \end{cases}$$

(Again compare this with the proof of 13.10.) Clearly we have

(3) If $\alpha < \beta$ and $F\beta \neq d$, then $(F\alpha)R(F\beta)$ and $F\alpha \neq F\beta$.

From (3) we see, by a familiar argument, that there is an α such that $F\alpha = d$. Choosing such an α minimal, we infer that

(4) $\{a : bRa \in A \text{ and } b \neq a, \text{ for every } b \in F^*\alpha\} = 0$.

From (3), however, we see that $F^*\alpha$ is simply ordered by R. Hence, by the hypothesis of Zorn's lemma, choose an R-upper bound $a \in A$ of $F^*\alpha$. Then a is an R-maximal element of A. For, if aRb and $a \neq b$, we would have cRb and $c \neq b$, for every $c \in F^*\alpha$, since cRa for every such c, and this would contradict (4).

Zorn's lemma $\Rightarrow$ maximality principle Obvious.

Maximality principle $\Rightarrow$ Kuratowski's principle Let $A = \{T : T \text{ is a simple ordering and } S \subseteq T \subseteq R\}$. Every subset B of A simply ordered by inclusion has an $\subseteq$-upper bound in A, namely, $\bigcup B$. Hence, by the maximality principle, let T be an $\subseteq$-maximal element of A. Obviously T satisfies the conclusion of Kuratowski's principle.

Kuratowski's principle $\Rightarrow$ trichotomy principle Let $R = \{(f,g) : f \text{ and } g \text{ are one-one functions, } Dmn\ f \subseteq Dmn\ g \subseteq A, Rng\ f \subseteq Rng\ g \subseteq B, \text{ and}$

$f \subseteq g\}$. Clearly R is a partial ordering. By Kuratowski's principle, with $S = 0$, let T be an $\subseteq$-maximal simple ordering such that $T \subseteq R$. By Theorem 8.9(i), $\bigcup Fld\ T \in Fld\ R$. Let $f = \bigcup Fld\ T$. If $Dmn\ f \neq A$ and $Rng\ f \neq B$, choose $a \in A \sim Dmn\ f$ and $b \in B \sim Rng\ f$, and let $g = f \cup \{(a,b)\}$. Clearly, then, fRg and $f \neq g$. Let $T' = T \cup \{(h,g) : h \in Fld\ T\} \cup \{(g,g)\}$. Then T' is a simple ordering, and $T \subset T' \subseteq R$, a contradiction. Therefore $Dmn\ f = A$ or $Rng\ f = B$. In the first case, f is a one-one mapping from A into B. In the second case, f^{-1} is a one-one mapping from B into A.

Trichotomy principle $\Rightarrow$ mapping principle By the trichotomy principle suppose, by symmetry, that there is a one-one function f mapping A into B. Choose $a \in A$. Let $g = f^{-1} \cup \{(b,a) : b \in B \sim Rng\ f\}$. Clearly g maps B onto A.

Mapping principle $\Rightarrow$ axiom of choice Let $M = \{R : R$ is a well-ordering with field $\subseteq S \bigcup A\}$. Then M is a set, since $M \subseteq S(S \bigcup A \times S \bigcup A)$. For each $R \in M$, let FR be the unique α such that R is isomorphic to $\{(\beta,\gamma) : \beta \leq \gamma < \alpha\}$, and let GR be the unique isomorphism from the latter set onto R. FR and GR exist by Theorem 13.10. Let $\Gamma = Rng\ F$. Thus Γ is a set. Furthermore

(5) $\Gamma = \{\alpha :$ there is a one-one function f mapping α into $S \bigcup A\}$.

Indeed, if $\alpha \in \Gamma$, say $\alpha = FR$, then GR is a one-one function mapping α into $S \bigcup A$. On the other hand, if f is a one-one function mapping α into $S \bigcup A$, let $R = \{(f\beta,f\gamma) : \beta \leq \gamma < \alpha\}$; then $R \in M$, $FR = \alpha$, and $GR = f$, so that $\alpha \in \Gamma$. This establishes (5). Let $\beta = \bigcup \Gamma \dotplus 1$. Then there is no one-one function mapping β into $S \bigcup A$; otherwise, by (5), $\beta \in \Gamma$ and hence $\beta \subseteq \bigcup \Gamma$, $\beta \leq \bigcup \Gamma < \beta$, a contradiction. Thus, indeed,

(6) There is no one-one function mapping β into $S \bigcup A$.

Furthermore

(7) There is no function mapping $\bigcup A$ onto β.

Indeed, if f maps $\bigcup A$ onto β, then $\langle f^{-1*}\{\alpha\} : \alpha < \beta\rangle$ maps β one-one into $S \bigcup A$, contradicting (6). Therefore (7) holds. We now apply the mapping principle to infer from (7) that there is a function g mapping β onto $\bigcup A$. Let h be the function with domain A such that $hx = g(\bigcap g^{-1*}\, x)$ for each $x \in A$. It is easily checked that h is a choice function for A.

This completes the proof of Theorem 16.2.

In connection with the multiplicative principle one should note the following theorem, which is easily established by induction on m, without using the axiom of choice.

Theorem 16.3 *For every m, if A is a function with domain m and $A_i \neq 0$ for every $i < m$, then $\mathsf{P}_{i \in m} A_i \neq 0$.*

Thus a choice can always be made from finitely many nonempty sets without using a principle equivalent to the axiom of choice.

Remark 16.4 Many equivalents of the axiom of choice are known; for a comprehensive account see Rubin, Rubin 1963. Our derivation of Zorn's lemma from the well-ordering principle is, we think, quite simple and natural. Most authors use a much more complicated proof that does not rely on ordinals but really contains, as a consequence of this, many of the steps in the proof of the recursion theorem. See, for example, Halmos 1960, pp. 62 to 65.

It is appropriate here to repeat the fact mentioned in Remark 1.37, that the axiom of choice is known to be independent of the other axioms of set theory; see Cohen 1963 to 1964.

EXERCISES

16.5 Give as direct a proof as possible for the following implications:

(*a*) Axiom of choice $\Rightarrow$ trichotomy principle.

(*b*) Zorn's lemma $\Rightarrow$ axiom of choice.

(*c*) Zorn's lemma $\Rightarrow$ well-ordering principle.

16.6 Show that each of the following statements is equivalent to the axiom of choice (without using the relational axiom of choice).

(*a*) If A is a set and $R \in V$ is a relation, then there is a maximal $B \subseteq A$ such that $B \times B \subseteq R$.

(*b*) If A is a set, then there is a maximal $B \subseteq A$ such that for all $x,y \in B$, $x \cap y = 0$ or $x = y$.

(*c*) If A is a set, then there is a maximal $B \subseteq A$ such that for all $x,y \in B$, $x \cap y \neq 0$.

17 APPLICATIONS OF THE AXIOM OF CHOICE

The axiom of choice is applied in many situations in mathematics, and many of these applications are essential, in that the other axioms do not suffice to obtain the desired result. In this section we will illustrate such applications, restricting ourselves to a few cases in which it is possible to introduce the necessary mathematical notions rapidly. The applications we give are illustrative only, and will not be used in succeeding sections. However, in Chap. 4 the axiom of choice will be applied several times in an essential way. For those with a broad background the following list of results for which the axiom of choice is needed may be of interest:

1 There is a set of real numbers that is not Lebesgue-measurable.
2 A product of compact topological spaces is compact.
3 Any vector space has a basis.
4 A countable union of countable sets is countable.
5 The Hahn-Banach extension theorem for linear functionals.
6 A divisible subgroup of an Abelian group is a direct summand.
7 In every ring with identity there is a maximal ideal.
8 Every Boolean algebra is isomorphic to a field of sets.
9 The completeness theorem for first-order logic.
10 The Banach-Tarski paradox.
11 Every partial ordering can be extended to a simple ordering.

Among the applications we give in this section are items 3 and 7. Item 11 is given as an exercise.

We begin with some simple applications in set theory itself.

Theorem 17.1 *(Principle of dependent choice) Let R be a relation with field A, $A \in V$ and suppose that for every $a \in A$ there is a $b \in A$ such that aRb. Further, suppose that $a_0 \in A$. Then there is a function f with domain ω such that $f_0 = a_0$ and $(fm)Rf(m \dot{+} 1)$ for every $m \in \omega$.*

Proof Let g be a choice function for $SA \sim \{0\}$. By recursion, there is a function f with domain ω such that $f_0 = a_0$ and $f(m \dot{+} 1) = g\{y : (fm)Ry\}$ for every m. Obviously f is the desired function.

Theorem 17.2 *If A is a function with domain ω and $A_i \subset A_{i+1}$ for every i, then there is a one-one function f mapping ω into $\bigcup_{i<\omega} A_i$ such that $Rng f \not\subseteq A_i$ for every $i < \omega$.*

Proof Let g be a choice function for $S(\bigcup_{i<\omega} A_i) \sim \{0\}$. For each $i < \omega$ let $f_i = g(A_{i+1} \sim A_i)$. Clearly f is the desired function.

Theorem 17.3 *For any partial ordering $R \in V$ there exist S, T satisfying the following conditions.*

(i) $R = S \cup T$, $S \cap T = I{\restriction}Fld\ R$, *and S and T are partial orderings each with field = Fld R.*

(ii) *There is no function f mapping ω into Fld S such that* $\forall m,n(m < n \Rightarrow (fm)S(fn) \wedge fm \neq fn)$.

(iii) *There is no function g mapping ω into Fld T such that* $\forall m,n(m < n \Rightarrow (fn)T(fm) \wedge fm \neq fn)$.

Proof Let U be a well-ordering with $Fld\ U = Fld\ R$. Define $S = \{(x,y) : xRy \text{ and } yUx\}$ and $T = \{(x,y) : xRy \text{ and } xUy\}$. The desired properties are now easily checked.

Theorem 17.4 *If $R \in V$ is a transitive relation that is reflexive on its field, then there is a partial ordering $S \subseteq R$ such that for every $x \in Fld\ R$ there is a $y \in Fld\ S$ such that xRy and yRx and S are isomorphic to the relation $\leq$ of Theorem 8.5.*

Proof Let $T = \{(x,y) : xRy \text{ and } yRx\}$. It is easily checked that T is an equivalence relation with field $Fld\ R$. Let $P = Fld\ R/T$. Thus P is a partition; by Zermelo's axiom, choose $A \subseteq Fld\ R$ such that $A \cap M$ has exactly one element for every $M \in P$. Let $S = R \cap (A \times A)$. The desired properties of S are easily checked.

We now turn to the two algebraic theorems 3 and 7. We will introduce only enough of the notions involved to prove these results conveniently.

Definition 17.5 *A **ring with identity** is a six-termed sequence $\mathfrak{A} = \langle A, +, \cdot, -, 0, 1\rangle$ such that $A \in V$, $A \neq 0$ (here 0 denotes the empty set, but not elsewhere in this definition), $+$ and $\cdot$ map $A \times A$ into A, $-$ maps A into A, $0,1 \in A$, and the following conditions hold for all $a,b,c \in A$.*

(*i*) $a + (b + c) = (a + b) + c$.
(*ii*) $a + b = b + a$.
(*iii*) $a + 0 = a$.
(*iv*) $a + -a = 0$.
(*v*) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
(*vi*) $a \cdot (b + c) = a \cdot b + a \cdot c$.
(*vii*) $(b + c) \cdot a = b \cdot a + c \cdot a$.
(*viii*) $a \cdot 1 = 1 \cdot a = a$.

In Definition 17.5 we use standard notation, which conflicts with our set-theoretical notation. Thus, in 17.5, 1 is used for an element that is not, in general, the integer 1, and 0 is used in two different senses. $-$ is used differently from its previous use for the predecessor of an ordinal, and $+$ and $\cdot$ do not mean the operations on cardinal numbers which will be introduced in Chap. 4. We believe this "abuse of notation" will not lead to any confusion.

Lemma 17.6 *If $\mathfrak{A}$ is a ring with identity, with notation as in Definition 17.5, then $b \cdot 0 = 0 \cdot b = 0$ for any $b \in A$.*

Proof We have

$$b \cdot 0 + b \cdot 0 = b \cdot (0 + 0) = b \cdot 0.$$

Hence

$$\begin{aligned} b \cdot 0 = b \cdot 0 + 0 &= b \cdot 0 + (b \cdot 0 + -(b \cdot 0)) \\ &= (b \cdot 0 + b \cdot 0) + -(b \cdot 0) = b \cdot 0 + -(b \cdot 0) = 0. \end{aligned}$$

This establishes part of our assertion, and the other part is similarly proved.

Definition 17.7 *Let $\mathfrak{A}$ be a ring with identity, with notation as in Definition 17.5. By an* ***ideal*** *in $\mathfrak{A}$ we mean a nonempty subset I of A satisfying the following two conditions.*
(i) For all $a,b \in I$ we also have $a + -b \in I$.
(ii) For all $a \in I$ and $b \in A$ we have $a \cdot b \in I$ and $b \cdot a \in I$.
The ideal I is said to be a ***maximal ideal*** *if $I \neq A$ and there is no ideal J such that $I \subset J \subset A$.*

Theorem 17.8 *In any ring with identity having at least two elements there is a maximal ideal.*

Proof Let $\mathfrak{A} = \{I : I \text{ is an ideal in } \mathfrak{A} \text{ and } I \neq A\}$. Now

(1) $\{0\} \in \mathfrak{A}$.

For clearly $\{0\} \neq 0$. Since $0 + -0 = 0$, 17.7(*i*) holds. Lemma 17.6 assures us that 17.7(*ii*) holds. Thus $\{0\}$ is an ideal. By the hypothesis that A has at least two elements, $\{0\} \neq A$. Thus, indeed, (1) holds.

Suppose $\mathfrak{B}$ is a subset of $\mathfrak{A}$ simply ordered by inclusion. If $\mathfrak{B} = 0$, then, by (1), $\{0\}$ is an $\subseteq$-upper bound of $\mathfrak{B}$ in $\mathfrak{A}$. Assume that $\mathfrak{B} \neq 0$. Then, we claim, $\bigcup\mathfrak{B}$ is an $\subseteq$-upper bound of $\mathfrak{B}$ in $\mathfrak{A}$. Obviously $I \subseteq \bigcup\mathfrak{B}$ for every $I \in \mathfrak{B}$, so that it is simply a matter of showing that $\bigcup\mathfrak{B} \in \mathfrak{A}$. Since $\mathfrak{B} \neq 0$, choose $I \in \mathfrak{B}$; since $I \neq 0$, by 17.7, choose $a \in I$. Then $a \in \bigcup\mathfrak{B}$, so that $\bigcup\mathfrak{B} \neq 0$. Now suppose that $a,b \in \bigcup\mathfrak{B}$. Then there exist $J,K \in \mathfrak{B}$ with $a \in J$ and $b \in K$. Since $\mathfrak{B}$ is simply ordered by inclusion, we may by symmetry suppose that $J \subseteq K$. Thus $a,b \in K$. By 17.7(*i*) for K, $a + -b \in K$. Hence $a + -b \in \bigcup\mathfrak{B}$, and this shows that 17.7(*i*) holds for $\bigcup\mathfrak{B}$. Next, suppose that $a \in \bigcup\mathfrak{B}$ and $b \in A$. Say $a \in J \in \mathfrak{B}$. Then $a \cdot b \in J$ and $b \cdot a \in J$, so that $a \cdot b \in \bigcup\mathfrak{B}$ and $b \cdot a \in \bigcup\mathfrak{B}$, so 17.7(*ii*) holds for $\bigcup\mathfrak{B}$. Thus $\bigcup\mathfrak{B}$ is an ideal in $\mathfrak{A}$. Suppose that $\bigcup\mathfrak{B} = A$. Then $1 \in \bigcup\mathfrak{B}$; say $1 \in J \in \mathfrak{B}$. For any $a \in A$ we then have $a = a \cdot 1 \in J$, so that $J = A$. Thus $J \notin \mathfrak{A}$; but this contradicts $J \in \mathfrak{B} \subseteq \mathfrak{A}$. This proves, finally, that $\bigcup\mathfrak{B}$ is an $\subseteq$-upper bound of $\mathfrak{B}$ in $\mathfrak{A}$. The hypotheses of the maximality principle are now met, so that we conclude that $\mathfrak{A}$ has an $\subseteq$-maximal element J. Clearly J is a maximal ideal in $\mathfrak{A}$.

Definition 17.9 *Let $\mathfrak{A}$ be a ring with identity, using the notation of Definition 17.5. $\mathfrak{A}$ is a* ***field*** *if A has at least two elements, $a \cdot b = b \cdot a$ for all $a,b \in A$, and for every $a \in A$ with $a \neq 0$ there is a $b \in A$ such that $a \cdot b = 1$.*

Let $\mathfrak{A}$ be a field. A ***vector space over*** *$\mathfrak{A}$ is a five-termed sequence $\mathfrak{B} = \langle B,+',\cdot',-',0'\rangle$ such that $B \in V$, $B \neq 0$, $+'$ is a mapping of $B \times B$*

into B, $\cdot'$ is a mapping of $A \times B$ into B, $-'$ maps B into B, $0' \in B$, and the following conditions hold for any $a,b \in A$ and $x,y,z \in B$.

(*i*) $x +' (y +' z) = (x +' y) +' z$.
(*ii*) $x +' y = y +' x$.
(*iii*) $x +' 0' = x$.
(*iv*) $x +' -' x = 0$.
(*v*) $(a + b)\cdot' x = a\cdot' x +' b\cdot' x$.
(*vi*) $a\cdot'(x +' y) = a\cdot' x +' a\cdot' y$.
(*vii*) $1\cdot' x = x$.
(*viii*) $(a\cdot b)\cdot' x = a\cdot'(b\cdot' x)$.

Lemma 17.10 *Let $\mathfrak{A}$ be a field, and $\mathfrak{B}$ a vector space over $\mathfrak{A}$, using the notation of Definitions 17.5 and 17.9. Then for any elements $a,b \in A$ and $x \in B$ we have*

(*i*) $-a\cdot b = -(a\cdot b)$.
(*ii*) $-a\cdot' x = -'(a\cdot' x)$.
(*iii*) $0\cdot' x = 0'$.
(*iv*) $a\cdot' 0' = 0'$.

Proof (*i*) We have, using Lemma 17.6,

$$a\cdot b + -a\cdot b = (a + -a)\cdot b = 0\cdot b = 0.$$

Hence

$$\begin{aligned} -a\cdot b &= -a\cdot b + 0 = -a\cdot b + (a\cdot b + -(a\cdot b)) \\ &= (-a\cdot b + a\cdot b) + -(a\cdot b) = (a\cdot b + -a\cdot b) + -(a\cdot b) \\ &= 0 + -(a\cdot b) = -(a\cdot b) + 0 = -(a\cdot b). \end{aligned}$$

Thus (*i*) holds. (*iii*) and (*iv*) are established analogously to the proof of 17.6, and then (*ii*) is proved much like (*i*).

Definition 17.11 *Let $\mathfrak{A}$ and $\mathfrak{B}$ be as in Definitions 17.5 and 17.9.*

(*i*) *If $m \in \omega \sim 1$ and $x \in {}^mB$, we define the expression $x_0 +' \cdots +' x_i$ for $i \leq m - 1$ by recursion.*

$$x_0 +' \cdots +' x_0 = x_0,$$
$$x_0 +' \cdots +' x_{i+1} = (x_0 +' \cdots +' x_i) +' x_{i+1} \quad \text{for } i + 1 \leq m - 1.$$

(*ii*) *A subset C of B is* ***independent*** *if for every $m \in \omega \sim 1$, every one-one $x \in {}^mC$, and every $a \in {}^mA$, the equation*

$$a_0\cdot' x_0 +' \cdots +' a_{m-1}\cdot' x_{m-1} = 0'$$

implies that $a_i = 0$ for every $i < m$.

(*iii*) *A subset C of B is a* ***basis*** *for $\mathfrak{B}$ if C is independent and for every $y \in B$*

there exist $m \in \omega \sim 1$, *a one-one* $x \in {}^mC$, *and* $a \in {}^mA$ *such that*

$$y = a_0 \cdot' x_0 +' \cdots +' a_{m-1} \cdot' x_{m-1}.$$

We will assume without proof some simple properties of the general sum $x_0 +' \cdots +' x_{m-1}$, such as the associative law, and generalizations of 17.9(*vi*).

Theorem 17.12 *Every vector space with at least two elements has a basis.*

Proof We assume the notation of Definitions 17.5 and 17.9. Let $\mathfrak{A} = \{C : C \text{ is an independent subset of } \mathfrak{B}\}$. In order to apply the maximality principle, suppose that $\mathfrak{B}$ is a subset of $\mathfrak{A}$ simply ordered by inclusion. Clearly $\bigcup\mathfrak{B}$ is an $\subseteq$-upper bound for $\mathfrak{B}$, so that we simply need to show that $\bigcup\mathfrak{B} \in \mathfrak{A}$, i.e., that $\bigcup\mathfrak{B}$ is independent. To do so, we need the following statement:

(1) For any $m \in \omega \sim 1$ and any $C \in {}^m\mathfrak{B}$, there is an $i < m$ such that for every $j < m$, $C_j \subseteq C_i$.

This statement is easily shown by induction on m. Now, to prove that $\bigcup\mathfrak{B}$ is independent, assume that $m \in \omega \sim 1$, x is one-one and $x \in {}^m\bigcup\mathfrak{B}$, $a \in {}^mA$, and

(2) $a_0 \cdot' x_0 +' \cdots +' a_{m-1} \cdot' x_{m-1} = 0'$.

Then there is a $C \in {}^m\mathfrak{B}$ such that $x_i \in C_i$ for every $i < m$; indeed, C can be obtained even without applying the axiom of choice, by appealing to 16.3. By (1), choose $i < m$ such that for every $j < m$, $C_j \subseteq C_i$. Then $x \in {}^mC_i$, and since C_i is independent, (2) yields $a_j = 0$ for every $j < m$. Thus, indeed, $\bigcup\mathfrak{B}$ is independent.

Hence we may apply the maximality principle to obtain an $\subseteq$-maximal element C of $\mathfrak{A}$. To show that C is a basis for $\mathfrak{B}$, let y be any element of B; we need to find $m \in \omega \sim 1$, a one-one $x \in {}^mC$, and $a \in {}^mA$ such that

(3) $y = a_0 \cdot' x_0 +' \cdots +' a_{m-1} \cdot' x_{m-1}$.

If $y \in C$, we may choose $m = 1$, $x = \{(0,y)\}$, and $a = \{(0,1)\}$; obviously (3) then holds. Suppose that $y \notin C$. Then $C \subset C \cup \{y\}$, so that $C \cup \{y\}$ is not independent. Therefore there exist $m \in \omega \sim 1$, a one-one $x \in {}^m(C \cup \{y\})$, and $a \in {}^mA$ such that

(4) $a_0 \cdot' x_0 +' \cdots +' a_{m-1} \cdot' x_{m-1} = 0'$

and $a_j \neq 0$ for a certain $j < m$. If $y \notin Rng\ x$, or if $x_i = y$ and $a_i = 0$, we easily conclude that C is not independent, a contradiction. Hence there is an i such that $x_i = y$ and $a_i \neq 0$, and without loss of generality

we may assume that $i = m - 1$. Choose b such that $a_{m-1} \cdot b = 1$. Then, multiplying both sides of (4) by $-b$,

$$(-b \cdot a_0) \cdot' x_0 +' \cdots +' (-b \cdot a_{m-1}) \cdot' x_{m-1} = 0'.$$

But $(-b \cdot a_{m-1}) \cdot' x_{m-1} = -(b \cdot a_{m-1}) \cdot' x_{m-1} = -1 \cdot' x_{m-1} = -'(1 \cdot' x_{m-1}) = -' x_{m-1} = -' y$. We easily infer that either $m = 1$ and $y = 0'$, or $m > 1$ and

$$y = (-b \cdot a_0) \cdot' x_0 +' \cdots +' (-b \cdot a_{m-2}) \cdot' x_{m-2};$$

in either case the proof is finished.

EXERCISES

17.13 For every partial ordering $R \in V$ there is a simple ordering S such that $R \subseteq S$ and $Fld\ R = Fld\ S$.

17.14 Suppose that A is a set such that for every m there does not exist a function mapping m onto A. Show that there is a one-one function mapping ω into A.

17.15 Let $\mathfrak{A}$ be a ring with identity, with notation as in Definition 17.5. A subset B of A is a *subring* of $\mathfrak{A}$ if $B \neq 0$, $a + -b \in B$ whenever $a,b \in B$, and $a \cdot b \in B$ whenever $a,b \in B$. Show that for any subring B of $\mathfrak{A}$ the set $\{C : C \text{ is a subring of } A,\ B \cap C = \{0\}\}$ has an $\subseteq$-maximal member.

17.16 If $R \in V$ is a simple ordering, then there is a set $A \subseteq Fld\ R$ well-ordered by R such that for every $x \in Fld\ R$ there is a $y \in A$ such that xRy.

17.17 (Extending Exercise 17.13). For every partial $R \in V$ we have $R = \bigcap \{S : R \subseteq S,\ S \text{ is a simple ordering, } Fld\ R = Fld\ S\}$.

4
Cardinals

We now begin the most important topic in this book, namely, the theory of counting. We generalize the usual finite cardinals m to the infinite case in order to have criteria for comparing the size of any sets (abstracted from any other aspects). We begin with the basic definitions and simplest properties of cardinals. Then we introduce the usual operations of cardinal arithmetic. In Sec. 23 we discuss singular and regular cardinals, which play an important role in abstract set theory. Section 24 is devoted to some applications.

18 CARDINALS: BASIC DEFINITIONS

Definition 18.1 (*i*) *A is **equipotent** with B iff there is a one-one function mapping A onto B.*

(*ii*) *A is a cardinal number, or simply a **cardinal**, if $A \in Ord$ and A is not equipotent with any $\alpha \in A$.*

(*iii*) *$Card = \{\alpha : \alpha$ is a cardinal$\}$.*

(*iv*) *Lowercase German letters* $\mathfrak{m},\mathfrak{n},\mathfrak{p},\mathfrak{q}$, . . . *are used for cardinals.*

Equipotence between sets is clearly an equivalence relation, and in speaking about *counting*, or about the "size" of sets, one wants to identify sets equivalent under this relation. It is not hard to see that every equivalence class except that of the empty set is actually a proper class. Indeed, for any set A the relation $B_\alpha = \{(\alpha,x) : x \in A\}$ is clearly equipotent with A, for any α. Thus for $A \neq 0$ $\langle B_\alpha : \alpha \in Ord\rangle$ maps Ord one-one into the equivalence class of A, and so, by Theorem 4.11, that equivalence class is a proper class. Thus equivalence classes under equipotence are not suitable for the above identifying process since we cannot combine them into bigger collections (compare with the discussion following 8.10). So we pick one representative from each equivalence class under equipotence, and use these representatives to describe the size of sets. The representatives are the cardinal numbers, and the intuitive facts just mentioned are formally the following statements: (1) distinct cardinal numbers are nonequipotent (18.2); (2) every set is equipotent with some cardinal (18.3).

Immediately from 18.1 we obtain the following.

Theorem 18.2 *If* $\mathfrak{m} \neq \mathfrak{n}$, *then* $\mathfrak{m}$ *and* $\mathfrak{n}$ *are not equipotent.*

Theorem 18.3 *For any set* A *there is a unique cardinal* $\mathfrak{m}$ *such that* $\mathfrak{m}$ *and* A *are equipotent.*

Proof By the counting principle, there is an ordinal α equipotent with A. The least such α is clearly a cardinal, $\mathfrak{m}$. The uniqueness of $\mathfrak{m}$ follows from 18.2.

Clearly Theorem 18.3 is actually equivalent to the counting principle and hence to the axiom of choice. It is also possible to get around the difficulty mentioned prior to 18.2—that equipotence equivalence classes are proper classes—without using the axiom of choice, by applying the method described in Sec. 15, which we will refer to as *Scott's definition* (see Montague, Scott, Tarski 1956). This means, of course, giving a different definition of *cardinal*. The theorems of this chapter that depend on 18.3, including those formulated using the following Definition 18.4, depend upon the axiom of choice. Many theorems in this chapter would not depend upon the axiom of choice using Scott's definition. For important theorems we will indicate its dependence or not on the axiom of choice under Scott's definition.

Theorem 18.3 justifies the following important definition.

Definition 18.4 *For any set* A, *let* $|A|$ *be the unique* $\mathfrak{m}$ *equipotent with* A. $|A|$ *is called the* ***power,*** *or* ***cardinality,*** *or* ***number of elements*** *of* A.

Alternative notations for $|A|$ are $\bar{\bar{A}}$ and $\#(A)$. $|A|$ is a uniquely determined representative of A under the equipotence equivalence relation. As such, $|A|$ has properties analogous to the properties of equivalence classes given in 7.4:

Theorem 18.5 *(i)* *x is equipotent with y iff $|x| = |y|$.*
(ii) *If x and y are both equipotent with z, then x and y are equipotent, and $|x| = |y|$.*
(iii) *x is equipotent with $|x|$.*
(iv) *If x is equipotent with $|y|$, then x is equipotent with y.*

Since cardinals are defined as special kinds of ordinals, it is also useful to state some general properties of $|\alpha|$.

Theorem 18.6 *(i)* $|\alpha| \leq \alpha$.
(ii) *If α and A are equipotent, then $|A| \leq \alpha$.*
(iii) *α is a cardinal iff $|\alpha| = \alpha$.*
(iv) *If $\omega \leq \alpha$, then $|\alpha \dot{+} 1| = |\alpha|$.*

Proof Only *(iv)* is nontrivial. Assuming that $\omega \leq \alpha$, the function $\langle i \dot{+} 1 : i \in \omega \rangle \cup [I\restriction(\alpha \sim \omega)] \cup \{(\alpha,0)\}$ clearly is one-one and maps $\alpha \dot{+} 1$ onto α.

We now give two important theorems comparing the size of sets. These theorems will be used very frequently in what follows, frequently without citation.

Theorem 18.7 *If $A \subseteq B \in V$, then $|A| \leq |B|$.*

Proof Let $\mathfrak{m} = |B|$, and let f be a one-one function mapping B onto $\mathfrak{m}$. Let $R = \{(\alpha,\beta) : \alpha \leq \beta < m \text{ and } \alpha,\beta \in f^*A\}$. Then R is a well-ordering. By 13.10, there exist α and g such that g is a one-one function mapping α onto $Fld\ R$ and for all $\beta,\gamma \in \alpha$, $\beta \leq \gamma$ iff $(g\beta)R(g\gamma)$. Now g is a strictly increasing α-termed sequence of ordinals, and $Rng\ g = Fld\ R = f^*A \subseteq \mathfrak{m}$, so that, by Theorem 12.3, $\alpha \leq \mathfrak{m}$. But $f^{-1}{\circ}g$ is a one-one function mapping α onto A; i.e., α and A are equipotent. By 18.6(*ii*), $|A| \leq \alpha \leq \mathfrak{m}$, as desired.

Theorem 18.8 *The following three conditions are equivalent for any sets A, B.*
(i) $|A| \leq |B|$.
(ii) *There is a one-one function mapping A into B.*
(iii) *$A = 0$, or there is a function mapping B onto A.*

Proof $(i) \Rightarrow (ii)$. If f is a one-one mapping of A onto $|A|$ and g is a one-one mapping of $|B|$ onto B, then $g \circ f$ is a one-one mapping of A into B.

$(ii) \Rightarrow (iii)$. Assume that $A \neq 0$. Let f be a one-one mapping of A into B. Let $a \in A$, and let $g = f^{-1} \cup \langle a : b \in B \sim Rng\ f \rangle$. Then g is a function mapping B onto A.

$(iii) \Rightarrow (i)$. Assume that $A \neq 0$. Let f map B onto A. By Theorem 4.14(i), let g map A into B with $f \circ g = I \restriction A$. Then g is one-one. Therefore $|A| = |g^*A| \leq |B|$, using 18.5 and 18.7.

Corollary 18.9 *(Cantor-Bernstein theorem) If A and B are sets such that A is equipotent with a subset of B and B is equipotent with a subset of A, then A and B are equipotent.*

Corollary 18.9 follows immediately from the equivalence of 18.8(i) and (ii) (recall that $\leq$ is a simple ordering). However, a direct proof of 18.9 not involving the axiom of choice (as this one does, if traced backwards) is instructive because of the methods involved in the proof. Let g map A one-one into B, and let h map B one-one into A. Let $R = \{(X,Y) : X \subseteq Y \subseteq A\}$. For any $X \subseteq A$ let $FX = A \sim h^*(B \sim g^*X)$. Then for any $X, Y \in SA$,

$$\begin{aligned} X \subseteq Y &\Rightarrow g^*X \subseteq g^*Y \Rightarrow B \sim g^*Y \subseteq B \sim g^*X \\ &\Rightarrow h^*(B \sim g^*Y) \subseteq h^*(B \sim g^*X) \Rightarrow FX \subseteq FY. \end{aligned}$$

Thus all the hypotheses of Theorem 8.7 are met, and we conclude that $FX = X$ for some $X \in SA$. Now $A \sim X = A \sim FX = h^*(B \sim g^*X)$. Thus $h^{-1} \restriction (A \sim X)$ maps $A \sim X$ onto $B \sim g^*X$. Also, $g \restriction X$ maps X onto g^*X. By Theorem 4.6 $g \restriction X \cup h^{-1} \restriction (A \sim X)$ is a one-one function mapping A onto B.

We have not yet exhibited even one cardinal. In the remainder of this section we will show that many ordinals are cardinals. In fact, *Card* is a proper class.

Theorem 18.10 $\omega \subseteq Card$.

Proof We prove $\forall m(m \in Card)$ by induction on m. The case $m = 0$ follows by a vacuous implication—see 18.1(ii). Now assume that $m \in Card$. If $m \dotplus 1$ is not a cardinal, then there is an $n < m \dotplus 1$ such that n is equipotent with $m \dotplus 1$; that is, there is a one-one function f mapping $m \dotplus 1$ onto n. Obviously $n \neq 0$. Let g be the permutation of n such that $gfm = n - 1$ and $g(n - 1) = fm$, and $gi = i$ if $i \neq n - 1, fm$ (thus $g = I \restriction n$ if $fm = n - 1$). [Recall from 4.9(v) that a *permutation* of A is a one-one map of A onto itself.] Then $g \circ f$ is a one-one map of $m \dotplus 1$ onto n such that $gfm = n - 1$, so that $(g \circ f) \restriction m$ is a one-one map

of m onto $n - 1$. By the induction assumption, $m = n - 1$. Hence $m \dotplus 1 = (n - 1) \dotplus 1 = n$, a contradiction. This completes the proof.

Theorem 18.11 $\omega \in Card$.

Proof By 18.6(*i*), $|\omega| \leq \omega$. For any $m \in \omega$ we have $m \subseteq \omega$, and hence $m = |m| \leq |\omega|$, by 18.10 and 18.7. Hence $|\omega| = \omega$, so that ω is a cardinal.

Theorem 18.12 *For any set* A, $|A| < |SA|$.

Proof The function $\langle \{a\} : a \in A \rangle$ is a one-one map from A into SA, so that, by 18.8, $|A| \leq |SA|$. Equality is impossible because of Theorem 6.8.

Corollary 18.13 *For any* α *there is an* $\mathfrak{m}$ *such that* $\alpha < \mathfrak{m}$.

Corollary 18.13 indicates that there are many cardinals. To formulate this fact more precisely, we need one more result.

Theorem 18.14 *If* Γ *is a set of cardinals, then* $\bigcup\Gamma$ *is a cardinal.*

Proof Suppose, on the contrary, that $|\bigcup\Gamma| < \bigcup\Gamma$ [compare with 18.6(*iii*)]. Thus $|\bigcup\Gamma| \in \mathfrak{n} \in \Gamma$ for some $\mathfrak{n}$. But then $\mathfrak{n} \subseteq \bigcup\Gamma$, and so $\mathfrak{n} = |\mathfrak{n}| \leq |\bigcup\Gamma|$, a contradiction.

Definition 18.15 (*i*) *For any ordinal* α, α^+ *is the least cardinal* $> \alpha$.
(*ii*) *The function* $\aleph$ *is defined by transfinite recursion:*

$\aleph_0 = \omega$.
$\aleph_{\alpha+1} = (\aleph_\alpha)^+$.
$\aleph_\gamma = \bigcup_{\beta<\gamma} \aleph_\beta$ *if* $\gamma = \bigcup\gamma \neq 0$.

Theorem 18.16 $\aleph$ *is a normal function and maps Ord onto Card* $\sim \omega$. *Hence Card is a proper class.*

Proof $\aleph$ is clearly a normal function and maps *Ord* into *Card* $\sim \omega$. Let $\mathfrak{m} \in Card \sim \omega$. By Theorem 12.2, $\mathfrak{m} \leq \aleph_\mathfrak{m} < \aleph_\mathfrak{m}{}^+$. Let α be the least ordinal such that $\mathfrak{m} < \aleph_\alpha$. Clearly α is not a limit ordinal and $\alpha \neq 0$. Thus $\alpha = \S\beta$ for some β. Then, by the minimality of α, $\aleph_\beta \leq \mathfrak{m} < \aleph_{\S\beta}$. Then 18.15(*i*) gives $\mathfrak{m} = \aleph_\beta$, so that $\mathfrak{m} \in Rng\ \aleph$. Thus $\aleph$ maps onto *Card* $\sim \omega$, as desired.

We will state informally some easy consequences of 18.15. Thus $\mathfrak{m} < \mathfrak{n}$ iff $\mathfrak{m}^+ \leq \mathfrak{n}$, and $\mathfrak{m} < \mathfrak{n}^+$ iff $\mathfrak{m} \leq \mathfrak{n}$. By 18.12, $\mathfrak{m}^+ \leq |S\mathfrak{m}|$; the possibility of equality here is a conjecture, the *generalized continuum hypothesis*, which is discussed further in Sec. 22. Since $\aleph$ is a normal function, by 13.9 it has fixed points; these are ordinals α such that

$\aleph_\alpha = \alpha$ (and such an α is actually a cardinal). Looking at the proof of 13.9, we see that the first fixed point of $\aleph$ is the least upper bound of the sequence

$$\aleph_0, \aleph_{\aleph_0}, \aleph_{\aleph_{\aleph_0}}, \ldots .$$

These cardinals, and indeed all cardinals greater than $\aleph_1$, say, may seem too large to be of genuine mathematical importance; but it turns out that many important mathematical questions, for example, in abstract measure theory or in the theory of Abelian groups, essentially involve very large cardinals.

EXERCISES

Prove the following statements.

18.17 For any set A, $|A| \leq |A \times 2|$.
18.18 If R is an equivalence relation on a set A, then $|A/R| \leq |A|$.
18.19 If A and B are sets such that $|A| \leq |B|$, then $|SA| \leq |SB|$.
18.20 If $Y \subseteq X \in V$ and Y is equipotent with $Y \cup Z$, then X is equipotent with $X \cup Z$.
18.21 For a given set A, with $|A| \geq 3$, let $B = SA \sim \{\{a\} : a \in A\}$. Then $|A| < |B|$.
18.22 $|\omega \dotplus m| = \omega$ for every m.
18.23 $|\omega \sim m| = \omega$ for every m.
18.24 $\mathfrak{m}^+ = \{\alpha : |\alpha| \leq \mathfrak{m}\}$ for every $\mathfrak{m}$.
18.25 If $\Gamma \subseteq Ord$ and Γ is a proper class such that $\bigcup \Delta \in \Gamma$ whenever $\Delta \subseteq \Gamma$ and $\Delta \in V$, then there is a normal function mapping Ord onto Γ.

19 FINITE AND INFINITE SETS

In this section we will prove various basic theorems about finite and infinite sets, although some theorems of great importance about them will be postponed until later sections.

Definition 19.1 *Let A be a set. Then A is **finite** if $|A| < \aleph_0$; **infinite** if $|A| \geq \aleph_0$; **denumerable** if $|A| = \aleph_0$; **countable** if $|A| \leq \aleph_0$; and **uncountable** if $|A| > \aleph_0$.*

We begin with some easy consequences of this definition.

Theorem 19.2 *If $A \subseteq B$, or if there is a function mapping B onto A, then A is finite if B is finite, and A is countable if B is countable.*

Theorem 19.3 (*i*) α *is finite iff* $\alpha < \omega$.
(*ii*) *If* $\mathfrak{m}$ *is infinite, then* $\mathfrak{m}$ *is a limit ordinal.*

Thus ω consists of all finite ordinals, and finite ordinals and finite cardinals are the same things. With regard to 19.3(*ii*), we will see later that not every limit ordinal is a cardinal; by Exercise 21.28, every cardinal is even an ϵ-number. By 18.6(*iv*), $\omega \dotplus m$ is denumerable for every $m \in \omega$; $S\omega$ is a simple example of an uncountable set, by 18.12.

Theorem 19.4 (*i*) $|\{x\}| = 1$.
(*ii*) *If* $x \neq y$, $|\{x,y\}| = 2$.

Proof (*i*) $\{(0,x)\}$ is a one-one function mapping 1 onto $\{x\}$.
(*ii*) $\{(0,x),(1,y)\}$ is a one-one function mapping 2 onto $\{x,y\}$.

We now turn to some "deeper" facts about finite and infinite sets.

Theorem 19.5 *If* A *is a finite set and* $B \subset A$, *then* $|B| < |A|$.

Proof Let $m = |A|$, and let f be a one-one function mapping A onto m. Choose $b \in A \sim B$. Let g be the permutation of m such that $gfb = m - 1$, $g(m - 1) = fb$, and $gi = i$ if $i \neq fb, m - 1$. Thus $g^*f^*(A \sim \{b\}) \subseteq m - 1$, so that $(g \circ f)\restriction B$ is a one-one function mapping B into $m - 1$ (note that $B \subseteq A \sim \{b\}$). Hence, by 18.8, $|B| \leq |m - 1| = m - 1 < m = |A|$. This completes the proof.

Theorem 19.5 gives a characteristic property of finite sets, as the following theorem of Dedekind shows.

Theorem 19.6 *A set* A *is infinite iff* A *is equipotent with a proper subset of itself.*

Proof $\Leftarrow$ By Theorem 19.5.

$\Rightarrow$ Assume that A is infinite; that is, $\aleph_0 \leq |A|$; that is, $\omega \leq |A|$. By 18.8, since $|\omega| = \omega$, there is a one-one function f mapping ω into A. Let $g = \langle i \dotplus 1 : i \in \omega \rangle$ and $h = (f \circ g \circ f^{-1}) \cup I\restriction(A \sim Rng\ f)$. It is easily checked that h is a one-one function mapping A onto its proper subset $A \sim \{f0\}$, as desired.

Theorem 19.6 essentially depends on the axiom of choice, even under Scott's definition (compare with the remark following 18.3). Note that "A infinite" should be redefined as "there is no one-one correspondence between A and a natural number m" for a definition of "infinite" not depending on the axiom of choice.

The last two theorems of this section give slightly more subtle results about finite sets; they are frequently useful.

Theorem 19.7 *If A and B are finite sets, $|A| = |B|$, and f maps A into B, then f maps A onto B iff f is one-one.*

Proof $\Leftarrow$ Assume that f is one-one. If f does not map A onto B, then $f^*A \subset B$, and so, by 19.5, $|A| = |f^*A| < |B|$, a contradiction.

$\Rightarrow$ Let g be a mapping of B into A such that $f \circ g = I \restriction B$, by Theorem 4.14($i$). Then g is one-one, and hence, by the first part of this proof, g maps B onto A. Thus $f = f \circ g \circ g^{-1} = (I \restriction B) \circ g^{-1} = g^{-1}$, so that f is one-one.

Theorem 19.8 *If $\preceq$ is a finite partial ordering, then $\prec$ is well-founded and there is a $\preceq$-maximal element. In particular, if $\preceq$ is a finite simple ordering, then $\preceq$ is a well-ordering and there is a $\preceq$-greatest element.*

Proof Suppose that $\preceq$ is a finite partial ordering and $\prec$ is not well-founded. Then there is a nonempty subset A of $Fld\ (\prec)$ such that $A \cap \{y : y \prec x\} \neq 0$ for every $x \in A$. Choose $a \in A$. Let f be a choice function for $\mathbf{S}Fld\ (\preceq) \sim \{0\}$. We define a function g mapping ω into $Fld\ (\preceq)$ by

$$g0 = a$$
$$g(m \dot{+} 1) = f\{b : b \in A, b \prec gm\}$$

for any $m \in \omega$. The above assumption on A implies that $\{b : b \prec gm\}$ is nonempty for each m. By induction on n, it is easily seen that, for all n and m, if $m < n$, then $gn \prec gm$. Hence g is a one-one mapping of ω into $Fld\ (\preceq)$, so that $\langle (g(m \dot{+} 1), gm) : m \in \omega \rangle$ is a one-one mapping of ω into $\prec$. Hence $|\prec| \geq \omega$, by Theorem 18.8, contradicting the assumption that $\preceq$ is finite. Thus $\prec$ is well-founded after all.

The proof that there is a $\preceq$-maximal element is entirely analogous, and will be omitted; and the second part of the theorem follows easily from the first.

Remark 19.9 A very readable discussion of finite sets is Tarski 1924. Among important facts about finite and infinite sets that we will prove in later sections are: a finite union of finite sets is finite (20.10); the pigeon-hole principle (20.11); a countable union of countable sets is countable (21.15); and a finite cartesian product of finite sets is finite (21.21).

EXERCISES

Prove the following statements.

19.10 The following four statements are equivalent:
(*a*) A is finite.
(*b*) A belongs to every class K such that $0 \in K$ and $B \cup \{x\} \in K$ whenever $B \in K$ and x is a set.
(*c*) Every nonempty set of subsets of A has an $\subseteq$-minimal element.
(*d*) Every proper subset of A is finite.

19.11 Let $\mathfrak{A} \neq 0$ be a set satisfying the following condition:

$$\forall A[A \in \mathfrak{A} \Leftrightarrow \forall B(B \text{ finite} \wedge B \subseteq A \Rightarrow B \in \mathfrak{A})].$$

Then $\mathfrak{A}$ has an $\subseteq$-maximal member.

19.12 α is finite iff $\geq \cap (\alpha \times \alpha)$ is a well-ordering.

20 CARDINAL ADDITION

The first of the arithmetical operations on cardinals, addition, will be discussed in this section. For ordinals, a binary operation of addition suffices for most purposes, although an infinitary operation could have been defined, but for cardinals a general infinitary operation is needed.

Throughout the next three sections I and J are arbitrary sets. By a *system of sets* we mean a function whose domain is a set.

Definition 20.1 *Let* $\mathfrak{m} = \langle \mathfrak{m}_i : i \in I \rangle$ *be a system of cardinals (any function with range* $\subseteq$ *Card and domain a set). The* ***cardinal sum*** *of* $\mathfrak{m}$, *denoted by* $\sum_{i \in I} \mathfrak{m}_i$, *is the cardinality of the set*

$$\bigcup_{i \in I} \{(i,\alpha) : \alpha \in \mathfrak{m}_i\}.$$

$\sum_{i \in 2} \mathfrak{m}_i$ *is denoted by* $\mathfrak{m}_0 + \mathfrak{m}_1$; *in general,* $\sum_{i \in m} \mathfrak{m}_i$ *is denoted by* $\mathfrak{m}_0 + \cdots + \mathfrak{m}_{m-1}$.

Note that $\mathfrak{m} + \mathfrak{n}$ is, by definition, $|\{(0,\alpha) : \alpha \in \mathfrak{m}\} \cup \{(1,\alpha) : \alpha \in \mathfrak{n}\}|$. In the general case, $\{(i,\alpha) : \alpha \in \mathfrak{m}\}$ is the relation whose inverse is the unique function that maps $\mathfrak{m}$ onto $\{i\}$. For $i,j \in I, i \neq j$, we have $(i,\alpha) \neq (j,\beta)$ for any α,β; further, $(i,\alpha) \neq (i,\beta)$ if $\alpha \neq \beta$.

In the general case, intuitively, we imagine I disjoint copies of the cardinals $\mathfrak{m}_i$ laid out, and we count the elements in the union. We first show that not merely the particular way of "disjointing" the $\mathfrak{m}_i$'s given in 20.1 gives the sum, but so does any other way.

Theorem 20.2 *If* $\langle A_i : i \in I \rangle$ *and* $\langle B_i : i \in I \rangle$ *are systems of pairwise disjoint sets, such that* A_i *is equipotent with* B_i *for each* $i \in I$, *then* $\bigcup_{i \in I} A_i$ *is equipotent with* $\bigcup_{i \in I} B_i$.

Proof By the multiplicative principle, let $f \in \prod_{i \in I} \{g : g \text{ is a one-one function mapping } A_i \text{ onto } B_i\}$. Thus for each $i \in I$, f_i is a one-one func-

tion mapping A_i onto B_i. It follows, from Theorem 5.9, that $\bigcup_{i\in I} f_i$ is a one-one function mapping $\bigcup_{i\in I} A_i$ onto $\bigcup_{i\in I} B_i$.

Corollary 20.3 *If* $|A_i| = \mathfrak{m}_i$ *for each* $i \in I$, *and if* $\langle A_i : i \in I\rangle$ *is a system of pairwise disjoint sets, then* $\sum_{i\in I} \mathfrak{m}_i = |\bigcup_{i\in I} A_i|$.

The next theorem is of fundamental importance in cardinal arithmetic.

Theorem 20.4 *For any system* $\langle A_i : i \in I\rangle$ *of sets,*

$$|\bigcup_{i\in I} A_i| \leq \sum_{i\in I} |A_i|.$$

Proof Let $|A_i| = \mathfrak{m}_i$ for each $i \in I$, and, using the axiom of choice, let f_i be a one-one function mapping $\mathfrak{m}_i$ onto A_i for each $i \in I$. Let

$$g = \bigcup_{i\in I} \{((i,\alpha), f_i\alpha) : \alpha < \mathfrak{m}_i\}.$$

Using Theorem 5.9, it is easily checked that g is a function mapping $\bigcup_{i\in I} \{(i,\alpha) : \alpha < \mathfrak{m}_i\}$ onto $\bigcup_{i\in I} A_i$. By Definition 20.1 and Theorem 18.8, the inequality of the theorem is immediate.

If we apply Theorem 20.4 to a system $\langle \mathfrak{m}_i : i \in I\rangle$ of cardinals, we obtain the following corollary.

Corollary 20.5 $\bigcup_{i\in I} \mathfrak{m}_i \leq \sum_{i\in I} \mathfrak{m}_i$.

(Recall from Theorem 18.14 that $\bigcup_{i\in I} \mathfrak{m}_i$ is a cardinal.) We now give some easy properties of addition.

Theorem 20.6 (*i*) $\sum_{i\in I} 0 = 0$.
(*ii*) $\sum_{I\in i} \mathfrak{m}_i = \sum_{I\in i, \mathfrak{m}_i \neq 0} \mathfrak{m}_i$.
(*iii*) $\sum_{i\in 0} \mathfrak{m}_i = 0$.
(*iv*) *If* $I \subseteq J$, *then* $\sum_{i\in I} \mathfrak{m}_i \leq \sum_{j\in J} \mathfrak{m}_j$.
(*v*) *If* $\mathfrak{m}_i \leq \mathfrak{n}_i$ *for each* $i \in I$, *then* $\sum_{i\in I} \mathfrak{m}_i \leq \sum_{i\in I} \mathfrak{n}_i$.
(*vi*) $\sum_{\alpha<\mathfrak{m}} 1 = \mathfrak{m}$.

Proof (*i*) to (*iii*) are immediate from the definition of addition. (*iv*) and (*v*) follow directly from Definition 20.1 and Theorem 18.8. As to (*vi*), $\{(\alpha,(\alpha,0)) : \alpha \in \mathfrak{m}\}$ is a one-one function mapping $\mathfrak{m}$ onto $\bigcup_{\beta<\mathfrak{m}} \{(\beta,\alpha) : \alpha \in 1\}$, so that, by 20.1, $\sum_{\alpha<\mathfrak{m}} 1 = \mathfrak{m}$.

We now prove general commutative and associative laws for addition that will be used in later discussions without citation. The ordinary

laws $\mathfrak{m} + \mathfrak{n} = \mathfrak{n} + \mathfrak{m}$ and $\mathfrak{m} + (\mathfrak{n} + \mathfrak{p}) = (\mathfrak{m} + \mathfrak{n}) + \mathfrak{p}$ easily follow from 20.7 and 20.8, as do their generalizations to any finite sum. Note especially that addition of cardinals is commutative, in contrast to ordinal addition. Thus, in particular, the cardinal sum $+$ of cardinals is, in general, distinct from the ordinal sum $\dot{+}$ (see the comment prior to Theorem 14.7). For natural numbers the two sums coincide, by Theorem 20.9. Similar comments apply to cardinal multiplication and exponentiation, introduced later.

Theorem 20.7 (*General commutative law*) *Let* $\langle \mathfrak{m}_i : i \in I \rangle$ *be a system of cardinals, and* f *a one-one function mapping some set* J *onto* I. *Then*

$$\sum_{i \in I} \mathfrak{m}_i = \sum_{j \in J} \mathfrak{m}_{fj}.$$

Proof As is easily checked, the desired one-one function mapping $\bigcup_{j \in J} \{(j,\beta) : \beta \in \mathfrak{m}_{fj}\}$ onto $\bigcup_{i \in I} \{(i,\alpha) : \alpha \in \mathfrak{m}_i\}$ (see Definition 20.1) is

$$\bigcup_{j \in J} \{((j,\beta),(fj,\beta)) : \beta \in \mathfrak{m}_{fj}\}.$$

Theorem 20.8 (*General associative law*) *Let* $\mathfrak{m} \in {}^{I \times J}Card$. *Then*

$$\sum_{i \in I} \left(\sum_{j \in J} \mathfrak{m}_{ij} \right) = \sum_{i \in I, j \in J} \mathfrak{m}_{ij}.$$

Proof For each $i \in I$ and $j \in J$ let

$$A_{ij} = \{((i,j),\alpha) : \alpha \in \mathfrak{m}_{ij}\}.$$

Then, clearly, for a given $i \in I$, $\langle A_{ij} : j \in J \rangle$ is a system of pairwise disjoint sets such that $|A_{ij}| = \mathfrak{m}_{ij}$ for each $j \in J$; consequently, by Corollary 20.3,

$$(1) \quad \left| \bigcup_{j \in J} A_{ij} \right| = \sum_{j \in J} \mathfrak{m}_{ij} \qquad \text{for each } i \in I.$$

Also, $\langle \bigcup_{j \in J} A_{ij} : i \in I \rangle$ is a system of pairwise disjoint sets; by (1) and Corollary 20.3, we obtain

$$(2) \quad \left| \bigcup_{i \in I} \bigcup_{j \in J} A_{ij} \right| = \sum_{i \in I} \left(\sum_{j \in J} \mathfrak{m}_{ij} \right).$$

But $\bigcup_{i \in I} \bigcup_{j \in J} A_{ij} = \bigcup_{i \in I, j \in J} A_{ij}$, by Theorem 5.7, and $\left| \bigcup_{i \in I, j \in J} A_{ij} \right| = \sum_{i \in I, j \in J} \mathfrak{m}_{ij}$, by Definition 20.1. Hence the theorem follows from (2).

We now connect our discussion of addition with the earlier material on natural numbers and finite sets.

Theorem 20.9 $m \dot{+} n = m + n$.

Proof We proceed by induction on n, proving $\forall n \forall m (m \dot{+} n = m + n)$. For $n = 0$, we have $m \dot{+} 0 = m$, and obviously $m + 0 = m$ [20.6(*ii*) may be applied]. Next we take the case $n = 1$. The function $\langle (0,i) : i \in m \rangle \cup$

$\{(m,(1,0))\}$ maps $m \dot{+} 1$ one-one onto $\{(0,i) : i \in m\} \cup \{(1,0)\}$. Hence

$$\begin{aligned} m \dot{+} 1 &= |m \dot{+} 1| && \text{by 18.10,} \\ &= |\{(0,i) : i \in m\} \cup \{(1,0)\}| \\ &= m + 1 && \text{by 20.1.} \end{aligned}$$

Now we take the general case, assuming inductively that $\forall m(m \dot{+} n = m + n)$. We have

$$\begin{aligned} m \dot{+} \mathcal{S}n &= m \dot{+} (n \dot{+} 1) = (m \dot{+} n) \dot{+} 1 && \text{by 14.6}(i), \\ &= (m + n) \dot{+} 1 && \text{by the inductive assumption,} \\ &= (m + n) + 1 && \text{by the above,} \\ &= m + (n + 1) && \text{by 20.8,} \\ &= m + (n \dot{+} 1) && \text{by the above,} \\ &= m + \mathcal{S}n. \end{aligned}$$

This completes the proof.

Note that Theorem 20.9 implies that, for any m, $\mathcal{S}m = m \dot{+} 1 = m + 1 = m^+$ [recall Definition 18.15(i)].

Theorem 20.10 *If A and B are finite, then so is $A \cup B$. More generally, a finite union of finite sets is finite.*

Proof $|A \cup B| \leq |A| + |B|$, by 20.4, and $|A| + |B| < \omega$, by 20.9. It is easily shown, by induction on $|I|$, that $\bigcup_{i \in I} A_i$ is finite whenever I is finite and A_i is finite for each $i \in I$.

The final theorem of this section is a form of the *pigeon-hole principle*, or the *shoe-box principle*.

Theorem 20.11 *If P is a partition of A, $|P| = |A|$, and A is finite, then $|M| = 1$ for each $M \in P$.*

Proof By the multiplicative axiom, choose $f \in \mathsf{P}_{M \in P} M$. For each $x \in A$ let $gx = f(x/R)$, where R is the equivalence relation with field A associated with P (see Theorem 7.6). Now f is a one-one function mapping P into A, and $|P| = |A|$. Hence, by Theorem 19.7, f maps P onto A. Since $\langle x/R : x \in A\rangle$ maps A onto P, it follows that g maps A onto A. Hence, again by Theorem 19.7, g is one-one. This implies immediately, from the definition of g, that $|M| = 1$ for each $M \in P$.

There are several facts about cardinal addition that we have not mentioned. For example, $\mathfrak{m} + \mathfrak{m} = \mathfrak{m}$ for any infinite cardinal m. These facts are more easily proved after the introduction of cardinal multiplication, so we defer them to the next section. We also save exercises on addition until then.

Remark 20.12 For this and the next two sections see especially Bachmann 1967 and Sierpinski 1965.

21 CARDINAL MULTIPLICATION

Definition 21.1 *For any system* $\langle \mathfrak{m}_i : i \in I \rangle$ *of cardinals we let*

$$\prod_{i\in I} \mathfrak{m}_i = \left| \mathsf{P}_{i\in I} \mathfrak{m}_i \right|.$$

If $I = 2$, *we write* $\mathfrak{m}_0 \cdot \mathfrak{m}_1$ *for* $\prod_{i\in I} \mathfrak{m}_i$.

As for addition $\mathfrak{m}+\mathfrak{n}$ is $\left| \mathsf{P}_{i\in 2} \mathfrak{p}_i \right|$, where $\mathfrak{p}_0 = \mathfrak{m}$ and $\mathfrak{p}_1 = \mathfrak{n}$.

Theorem 21.2 *If* $|A_i| = |B_i|$ *for every* $i \in I$, *then*

$$\left| \mathsf{P}_{i\in I} A_i \right| = \left| \mathsf{P}_{i\in I} B_i \right|.$$

Proof Using the axiom of choice, let f be a function with domain I such that, for each $i \in I$, f_i is itself a one-one function, mapping A_i onto B_i. The desired one-one mapping of $\mathsf{P}_{i\in I} A_i$ onto $\mathsf{P}_{i\in I} B_i$ is then

$$h = \langle \langle f_i x_i : i \in I \rangle : x \in \mathsf{P}_{i\in I} A_i \rangle,$$

as is easily checked. Note that, for $x \in \mathsf{P}_{i\in I} A_i$, hx is a member of $\mathsf{P}_{i\in I} B_i$; in fact, for each $i \in I$, we then have $(hx)_i = f_i x_i \in B_i$.

Corollary 21.3 *If* $\langle A_i : i \in I \rangle$ *is any system of sets, then*

$$\left| \mathsf{P}_{i\in I} A_i \right| = \prod_{i\in I} |A_i|.$$

Proof
$$\left| \mathsf{P}_{i\in I} A_i \right| = \left| \mathsf{P}_{i\in I} |A_i| \right| \quad \text{by 21.2,}$$
$$= \prod_{i\in I} |A_i| \quad \text{by 21.1.}$$

Theorem 21.4 $|A \times B| = |A| \cdot |B|$.

Proof $\{((a,b),\langle a,b\rangle) : a \in A, b \in B\}$ is a one-one function mapping $A \times B$ onto $\mathsf{P}_{i<2} C_i$, where $C_0 = A$ and $C_1 = B$. Hence 21.4 follows from 21.3.

Theorem 21.5 (*i*) *If* $\mathfrak{m}_i = 0$ *for some* $i \in I$, *then* $\prod_{i\in I} \mathfrak{m}_i = 0$.

(*ii*) $\prod_{i\in 0} \mathfrak{m}_i = 1$.

(*iii*) $\prod_{i\in I} \mathfrak{m}_i = \prod_{i\in I, \mathfrak{m}_i \neq 1} \mathfrak{m}_i$.

(*iv*) $\prod_{i\in I} 1 = 1$.

(*v*) $\sum_{i\in I} \mathfrak{m} = |I| \cdot \mathfrak{m}$.

(*vi*) *If* $\mathfrak{m}_i \leq \mathfrak{n}_i$ *for each* $i \in I$, *then* $\prod_{i\in I} \mathfrak{m}_i \leq \prod_{i\in I} \mathfrak{n}_i$.

Proof (*i*) Since then $\mathsf{P}_{i\in I}\,\mathfrak{m}_i = 0$.

(*ii*) Since $\mathsf{P}_{i\in 0}\,\mathfrak{m}_i = \{0\}$.

(*iii*) The function $\langle f\restriction\{i : i \in I, \mathfrak{m}_i \neq 1\} : f \in \mathsf{P}_{i\in I}\,\mathfrak{m}_i\rangle$ is a one-one function mapping $\mathsf{P}_{i\in I}\,\mathfrak{m}_i$ onto $\mathsf{P}_{i\in I,\mathfrak{m}_i\neq 1}\,\mathfrak{m}_i$, and (*iii*) follows.

(*iv*) $\mathsf{P}_{i\in I}\,1$ has just one element, namely $\langle 0 : i \in I\rangle$.

(*v*) Since $I\times\mathfrak{m} = \bigcup_{i\in I}\{(i,\alpha) : \alpha \in \mathfrak{m}\}$, it follows that

$$\begin{aligned} |I|\cdot\mathfrak{m} &= |I\times\mathfrak{m}| && \text{by 21.4,}\\ &= \textstyle\sum_{i\in I}\mathfrak{m} && \text{by 20.1.}\end{aligned}$$

Finally, (*vi*) is trivial since $\mathsf{P}_{i\in I}\,\mathfrak{m}_i \subseteq \mathsf{P}_{i\in I}\,\mathfrak{n}_i$ under the assumption of (*vi*).

Next we give the general commutative, associative, and distributive laws. Again, the usual laws $\mathfrak{m}\cdot\mathfrak{n} = \mathfrak{n}\cdot\mathfrak{m}$, $\mathfrak{m}\cdot(\mathfrak{n}\cdot\mathfrak{p}) = (\mathfrak{m}\cdot\mathfrak{n})\cdot\mathfrak{p}$, and $\mathfrak{m}\cdot(\mathfrak{n}+\mathfrak{p}) = \mathfrak{m}\cdot\mathfrak{n} + \mathfrak{m}\cdot\mathfrak{p}$ are easily derived from them.

Theorem 21.6 (*General commutative law*) *If f is a one-one function mapping J onto I, and if $\langle m_i : i \in I\rangle$ is a system of cardinals, then*

$$\textstyle\prod_{i\in I}\mathfrak{m}_i = \prod_{j\in J}\mathfrak{m}_{fj}.$$

Proof The desired function mapping $\mathsf{P}_{i\in I}\,\mathfrak{m}_i$ onto $\mathsf{P}_{j\in J}\,\mathfrak{m}_{fj}$ is

$$\langle\langle x_{fj} : j \in J\rangle : x \in \mathsf{P}_{i\in I}\,\mathfrak{m}_i\rangle.$$

Theorem 21.7 (*General associative law*) *If $\langle\mathfrak{m}_{ij} : (i,j) \in I\times J\rangle$ is a system of cardinals, then*

$$\textstyle\prod_{i\in I}\left(\prod_{j\in J}\mathfrak{m}_{ij}\right) = \prod_{i\in I,j\in J}\mathfrak{m}_{ij}.$$

Proof By Corollary 21.3, we have

$$\begin{aligned}\textstyle\prod_{i\in I}\left(\prod_{j\in J}\mathfrak{m}_{ij}\right) &= \left|\,\mathsf{P}_{i\in I}\left(\mathsf{P}_{j\in J}\,\mathfrak{m}_{ij}\right)\right|;\\ \textstyle\prod_{i\in I,j\in J}\mathfrak{m}_{ij} &= \left|\,\mathsf{P}_{i\in I,j\in J}\,\mathfrak{m}_{ij}\right|.\end{aligned}$$

Hence it suffices to exhibit a one-one function mapping $\mathsf{P}_{i\in I,j\in J}\,\mathfrak{m}_{ij}$ onto $\mathsf{P}_{i\in I}\left(\mathsf{P}_{j\in J}\,\mathfrak{m}_{ij}\right)$. Such a function is

$$\langle\langle\langle x_{ij} : j \in J\rangle : i \in I\rangle : x \in \mathsf{P}_{i\in I,j\in J}\,\mathfrak{m}_{ij}\rangle.$$

Theorem 21.8 (*General distributive law*) *Let $\langle m_{ij} : i \in I, j \in J\rangle$ be a system of cardinals. Then*

$$\textstyle\prod_{i\in I}\sum_{j\in J}\mathfrak{m}_{ij} = \sum_{f\in{}^IJ}\prod_{i\in I}\mathfrak{m}_{i,fi}.$$

Proof Using Corollary 21.3, as in the preceding proof, and using Definition 20.1, we see that we need to exhibit a one-one mapping between the set $\mathsf{P}_{i\in I} \bigcup_{j\in J} \{(j,\alpha) : \alpha \in \mathfrak{m}_{ij}\}$ and the set $\bigcup_{f\in {}^IJ} \{(f,\alpha) : \alpha \in \prod_{i\in I} \mathfrak{m}_{i,fi}\}$. For any $x \in \mathsf{P}_{i\in I} \bigcup_{j\in J} \{(j,\alpha) : \alpha \in \mathfrak{m}_{ij}\}$ and any $i \in I$, there is a unique $j \in J$ such that $x_i \in \{(j,\alpha) : \alpha \in \mathfrak{m}_{ij}\}$. Hence there is a function F that makes correspond to any such x and i the element $F_{xi} \in J$ such that $x_i \in \{(F_{xi},\alpha) : \alpha \in \mathfrak{m}_{iF_{xi}}\}$. Furthermore, using the axiom of choice, there is a function G such that, for each $f \in {}^IJ$, Gf is a one-one function mapping $\mathsf{P}_{i\in I} \mathfrak{m}_{i,fi}$ onto $\prod_{i\in I} \mathfrak{m}_{i,fi}$. Now for any $x \in \mathsf{P}_{i\in I} \bigcup_{j\in J} \{(j,\alpha) : \alpha \in \mathfrak{m}_{ij}\}$ we let

$$Hx = (\langle F_{xi} : i \in I\rangle, (G\langle F_{xi} : i \in I\rangle)\langle 2^{\text{nd}}x_i : i \in I\rangle)$$

Recall that $2^{\text{nd}}(a,b) = b$ for any b (5.14.) Thus for x and i as above, $2^{\text{nd}}x_i$ is the unique $\alpha \in m_{iF_{xi}}$ such that $x_i = (F_{xi},\alpha)$; hence $\langle 2^{\text{nd}}x_i : i \in I\rangle \in \mathsf{P}_{i\in I} \mathfrak{m}_{iF_{xi}}$. Thus H maps $\mathsf{P}_{i\in I} \bigcup_{j\in J} \{(j,\alpha) : \alpha \in m_{ij}\}$ into $\bigcup_{f\in {}^IJ} \{(f,\alpha) : \alpha \in \prod_{i\in I} m_{i,fi}\}$. To show H maps onto this last set, let $f \in {}^IJ$ and $\alpha \in \prod_{i\in I} m_{i,fi}$. If we let

$$y = \langle (fi, ((Gf)^{-1}\alpha)_i) : i \in I\rangle,$$

it is easily checked that $y \in Dmn\, H$ and $Hy = (f,\alpha)$, as desired. To show that H is one-one, assume that $x,y \in Dmn\, H$ and $x \neq y$. Then there is an $l \in I$ such that $x_l \neq y_l$. Say $x_l = (j,\alpha)$ with $j \in J$ and $\alpha \in \mathfrak{m}_{lj}$, and that $y_l = (k,\beta)$ with $k \in J$ and $\beta \in \mathfrak{m}_{lk}$. Thus $F_{xl} = j$ and $F_{yl} = k$. If $j \neq k$, then $\langle F_{xi} : i \in I\rangle \neq \langle F_{yi} : i \in I\rangle$, and it is clear that $Hx \neq Hy$. Since $x_l \neq y_l$, the other possibility is that $\alpha \neq \beta$. Now $2^{\text{nd}}x_l = \alpha$ and $2^{\text{nd}}y_l = \beta$, so that we again infer easily that $Hx \neq Hy$. This completes the proof.

Although the general distributive law 21.8 is frequently needed, in most of its uses the following special case is sufficient.

Theorem 21.9 *If $\langle \mathfrak{n}_i : i \in I\rangle$ is a system of cardinals, then*

$$\mathfrak{m}\cdot\textstyle\sum_{i\in I}\mathfrak{n}_i = \sum_{i\in I} (\mathfrak{m}\cdot\mathfrak{n}_i).$$

Proof To apply 21.8, we first ignore the trivial case $I = 0$. Let $l \in I$. Let $p_{0l} = m$, and $p_{0i} = 0$ if $i \in I \sim \{l\}$; and let $\mathfrak{p}_{1i} = \mathfrak{n}_i$ for every $i \in I$. It is then clear, using 20.6(ii), that

$$(1) \quad \textstyle\prod_{j\in 2} \sum_{i\in I} \mathfrak{p}_{ji} = m\cdot\sum_{i\in I} \mathfrak{n}_i.$$

By Theorem 21.8, we have

$$(2) \quad \textstyle\prod_{j\in 2} \sum_{i\in I} \mathfrak{p}_{ji} = \sum_{f\in {}^2I} \prod_{j\in 2} \mathfrak{p}_{j,fj}.$$

Now, by Theorem 21.5(i), $\prod_{j\in 2} \mathfrak{p}_{j,fj} = 0$ unless $f0 = l$. Hence, using

20.6(*ii*) again,

$$\begin{aligned}\textstyle\sum_{f\in {}^{2}I}\prod_{j\in 2}\mathfrak{p}_{j,fj} &= \textstyle\sum_{f\in {}^{2}I,\, f0=i}\prod_{j\in 2}\mathfrak{p}_{j,fj}\\ &= \textstyle\sum_{i\in I}(\mathfrak{m}\cdot\mathfrak{n}_i);\end{aligned}$$

combined with (1) and (2) this gives the desired result.

We now turn to relatively deeper matters. The following theorem gives one of the most fundamental properties of infinite cardinals, and it plays an important role even in routine cardinal-number calculations in mathematics.

Theorem 21.10 *For any infinite cardinal* $\mathfrak{m}$, $\mathfrak{m}\cdot\mathfrak{m} = \mathfrak{m}$.

Proof We suppose that the theorem is false, and we let $\mathfrak{m}$ be the least infinite cardinal such that $\mathfrak{m}\cdot\mathfrak{m} \neq \mathfrak{m}$. Let $\preceq$ be the set of all pairs $((\alpha,\beta),(\gamma,\delta))$ such that $\alpha,\beta,\gamma,\delta \in \mathfrak{m}$ and one of the following four conditions holds:

(1) $\alpha = \gamma$ and $\beta = \delta$;
(2) $\alpha \cup \beta < \gamma \cup \delta$;
(3) $\alpha \cup \beta = \gamma \cup \delta$ and $\alpha < \gamma$;
(4) $\alpha \cup \beta = \gamma \cup \delta$, $\alpha = \gamma$, and $\beta < \delta$.

We claim that $\preceq$ is a well-ordering with field $\mathfrak{m}\times\mathfrak{m}$. It is obvious that $\preceq$ is reflexive on $\mathfrak{m}\times\mathfrak{m}$ and antisymmetric. If we assume that $(\alpha,\beta) \preceq (\gamma,\delta) \preceq (\varepsilon,\zeta)$, we easily conclude that $(\alpha,\beta) \preceq (\varepsilon,\zeta)$ in each of these cases, which exhaust all possibilities: $\alpha \cup \beta < \gamma \cup \delta$ or $\gamma \cup \delta < \varepsilon \cup \zeta$; $\alpha \cup \beta = \gamma \cup \delta = \varepsilon \cup \zeta$, and $\alpha < \gamma$ or $\gamma < \varepsilon$; $\alpha \cup \beta = \gamma \cup \delta = \varepsilon \cup \zeta$, $\alpha = \gamma = \varepsilon$, and $\beta < \delta$ or $\delta < \zeta$; $(\alpha,\beta) = (\gamma,\delta) = (\varepsilon,\zeta)$. Thus $\preceq$ is transitive. Clearly $(\alpha,\beta) \preceq (\gamma,\delta)$ or $(\gamma,\delta) \preceq (\alpha,\beta)$, for any α, β, γ, δ, and clearly $\preceq$ has field $\mathfrak{m}\times\mathfrak{m}$. It remains to show that a nonempty subset Γ of $\mathfrak{m}\times\mathfrak{m}$ has a $\preceq$-least element. Let γ be the least element of $\{\alpha \cup \beta : (\alpha,\beta) \in \Gamma\}$, δ the least element of $\{\alpha : (\alpha,\beta) \in \Gamma \text{ for some } \beta \text{ with } \alpha \cup \beta = \gamma\}$, and ε the least element of $\{\beta : (\delta,\beta) \in \Gamma\}$. Clearly (δ,ε) is then the $\preceq$-least element of Γ. Thus, indeed, $\preceq$ is a well-ordering with field $\mathfrak{m}\times\mathfrak{m}$. By Theorem 13.10, let α be an ordinal and f an isomorphism of $\{(\beta,\gamma) : \beta \leq \gamma < \alpha\}$ onto $\preceq$. If $\alpha \leq \mathfrak{m}$, then

$$\begin{aligned}\mathfrak{m}\cdot\mathfrak{m} &= |\mathfrak{m}\times\mathfrak{m}| && \text{by 21.4,}\\ &= |\alpha| \leq |\mathfrak{m}| = \mathfrak{m} = \mathfrak{m}\cdot 1 \leq \mathfrak{m}\cdot\mathfrak{m} && \text{using 21.5,}\end{aligned}$$

giving $\mathfrak{m}\cdot\mathfrak{m} = \mathfrak{m}$, contradicting the initial assumption on $\mathfrak{m}$. Hence $\alpha > \mathfrak{m}$. Thus $\mathfrak{m} \in Dmn f$; say $f\mathfrak{m} = (\beta,\gamma)$. Let $\delta = (\beta \cup \gamma) \dotplus 1$. Note that $\delta < \mathfrak{m}$, since $\mathfrak{m}$ is a limit ordinal, by Theorem 19.3(*ii*). Now $\beta \cup \gamma < \delta$, so that $(\beta,\gamma) \prec (\delta,\delta)$. In fact, for any $\varepsilon < \mathfrak{m}$ we have $f\varepsilon \prec f\mathfrak{m} \prec (\delta,\delta)$, and $f\varepsilon \neq (\delta,\delta)$; hence if $f\varepsilon = (\zeta,\eta)$, we get $\zeta,\eta \leq \zeta \cup \eta \leq \beta \cup \gamma < \delta$.

Hence $f \restriction \mathfrak{m}$ is a one-one function mapping $\mathfrak{m}$ into $\delta \times \delta$; therefore

$$\mathfrak{m} \leq |\delta \times \delta| = |\delta| \cdot |\delta| \qquad \text{by 21.4.}$$

But $\delta < \mathfrak{m}$, so that $|\delta| < \mathfrak{m}$, and so, by the choice of $\mathfrak{m}$, either $\delta < \omega$ and $\delta = |\delta|$, $|\delta| \cdot |\delta| < \omega \leq \mathfrak{m}$, which contradicts the above, or $\omega \leq \delta$ and $|\delta| \cdot |\delta| = |\delta| < \mathfrak{m}$ since $\delta < \mathfrak{m}$, which is again a contradiction. These contradictions show that the initial assumption was untenable, so that the proof is completed.

Note that the axiom of choice is not involved in the proof of 21.10. Introducing $\preceq$, as in the proof, it is easily seen that $\preceq$ is a well-ordering of $\mathfrak{m} \times \mathfrak{m}$ that is isomorphic to $\{(\alpha,\beta) : \alpha \leq \beta < \mathfrak{m}\}$—compare with Exercise 21.31 also.

Many useful corollaries follow at once from Theorem 21.10. They are used in the subsequent work without special citation. They show that addition and multiplication of two cardinals are essentially trivial operations when infinite cardinals are involved.

Corollary 21.11 *If* $\mathfrak{m} \geq \omega$ *or* $\mathfrak{n} \geq \omega$, *then* $\mathfrak{m} + \mathfrak{n} = \mathfrak{m} \cup \mathfrak{n}$; *if in addition* $\mathfrak{m} \neq 0 \neq \mathfrak{n}$, *then* $\mathfrak{m} \cdot \mathfrak{n} = \mathfrak{m} \cup \mathfrak{n}$. *In particular, for* $\mathfrak{m} \geq \omega$ *we have* $\mathfrak{m} + \mathfrak{m} = \mathfrak{m} = \mathfrak{m} \cdot \mathfrak{m}$. *For any* α, β *we have* $\aleph_\alpha + \aleph_\beta = \aleph_{\alpha \cup \beta} = \aleph_\alpha \cdot \aleph_\beta$.

Proof First assume that $\mathfrak{m} \geq \omega$ or $\mathfrak{n} \geq \omega$; say, by symmetry, $\mathfrak{m} \geq \omega$ and $\mathfrak{m} \geq \mathfrak{n}$. Then

$$\begin{aligned} \mathfrak{m} + \mathfrak{n} &\leq \mathfrak{m} + \mathfrak{m} = \mathfrak{m} \cdot 2 && \text{by 21.5}(v), \\ &\leq \mathfrak{m} \cdot \mathfrak{m} = \mathfrak{m} && \text{by 21.10}, \\ &\leq \mathfrak{m} + \mathfrak{n}. \end{aligned}$$

Thus $\mathfrak{m} + \mathfrak{n} = \mathfrak{m} = \mathfrak{m} \cup \mathfrak{n}$ (the maximum of $\mathfrak{m}$ and $\mathfrak{n}$). Now assume that $\mathfrak{m} \neq 0 \neq \mathfrak{n}$ (and still that $\omega \leq \mathfrak{m} \geq \mathfrak{n}$). Then

$$\begin{aligned} \mathfrak{m} \cdot \mathfrak{n} &\leq \mathfrak{m} \cdot \mathfrak{m} = \mathfrak{m} && \text{by 21.10}, \\ &= \mathfrak{m} \cdot 1 \leq \mathfrak{m} \cdot \mathfrak{n}, \end{aligned}$$

so that $\mathfrak{m} \cdot \mathfrak{n} = \mathfrak{m} = \mathfrak{m} \cup \mathfrak{n}$ also. The final equations of the corollary follow from the first part of the corollary, by observing that $\aleph_\alpha \cup \aleph_\beta = \aleph_{\alpha \cup \beta}$.

Corollary 21.12 *Let* $\mathfrak{m}$ *be an infinite cardinal and* $\langle A_i : i \in I \rangle$ *a system of sets with* $|A_i| \leq \mathfrak{m}$ *for each* $i \in I$ *and* $|I| \leq \mathfrak{m}$. *Then* $\left| \bigcup_{i \in I} A_i \right| \leq \mathfrak{m}$ also.

Proof We have

$$\begin{aligned} \left| \bigcup_{i \in I} A_i \right| &\leq \sum_{i \in I} |A_i| && \text{by 20.4}, \\ &\leq \sum_{i \in I} \mathfrak{m} = |I| \cdot \mathfrak{m} && \text{by 21.5}(v), \\ &\leq \mathfrak{m} \cdot \mathfrak{m} = \mathfrak{m} && \text{by 21.10}. \end{aligned}$$

To formulate the next corollary of Theorem 21.10 conveniently, we need the following.

Definition 21.13 *A cardinal* $\mathfrak{m}$ *is said to be* ***regular*** *if, for all* $\Gamma \subseteq \mathfrak{m}$, *if* $|\Gamma| < \mathfrak{m}$ *then* $\bigcup\Gamma < \mathfrak{m}$. *If* $\mathfrak{m}$ *is not regular, it is said to be* ***singular.***

Corollary 21.14 ω *is regular, and for any infinite cardinal* $\mathfrak{m}$, $\mathfrak{m}^+$ *is regular.*

Proof Assume that $\Gamma \subseteq \omega$ and $|\Gamma| < \omega$. Then, by Theorem 19.8, Γ has a greatest element m. It is clear that $\bigcup\Gamma = m$, so that $\bigcup\Gamma \in \omega$. Therefore, ω is regular. Now suppose that $\mathfrak{m}$ is an infinite cardinal, and suppose that $\Gamma \subseteq \mathfrak{m}^+$ with $|\Gamma| < \mathfrak{m}^+$; that is, $|\Gamma| \leq \mathfrak{m}$. For each $\alpha \in \Gamma$ we have $\alpha < \mathfrak{m}^+$ and hence $|\alpha| < \mathfrak{m}^+$; that is, $|\alpha| \leq \mathfrak{m}$. Therefore, by Corollary 21.12, $\left|\bigcup_{\alpha\in\Gamma} \alpha\right| \leq \mathfrak{m}$. Since $\bigcup_{\alpha\in\Gamma} \alpha = \bigcup\Gamma$ is an ordinal, it follows that $\bigcup\Gamma < \mathfrak{m}^+$, as desired.

An example of a singular cardinal is $\aleph_\omega$, since $\aleph_\omega = \bigcup_{n\in\omega} \aleph_n$ and $\omega < \aleph_\omega$. In fact, by Corollary 21.14, $\aleph_\omega$ is the first infinite singular cardinal.

As a special case of 21.12 we have the following important corollary.

Corollary 21.15 *A countable union of countable sets is countable.*

Next we want to give an equivalent definition of a cardinal's being regular. The equivalence depends upon a lemma.

Lemma 21.16 *If* $\langle \mathfrak{m}_\alpha : \alpha \in \mathfrak{n} \rangle$ *is a system of infinite cardinals such that for all* α, β, $\alpha < \beta < \mathfrak{n}$ *implies that* $\mathfrak{m}_\alpha \leq \mathfrak{m}_\beta$, *then*

$$\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha = \sum_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha.$$

If actually $\mathfrak{m}_\alpha < \mathfrak{m}_\beta$ *whenever* $\alpha < \beta < \mathfrak{n}$, *then* $\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha = \bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha$.

Proof We have

$$\begin{aligned}
\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha &\leq \mathfrak{n}\cdot\sum_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha && \text{by 20.5,}\\
&\leq \mathfrak{n}\cdot\sum_{\alpha<\mathfrak{n}} \left(\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha\right) && \text{by 20.6}(v),\\
&= \mathfrak{n}\cdot\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha && \text{by 21.5}(v),\\
&= \mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha.
\end{aligned}$$

Here the last equality follows immediately if $\mathfrak{n}$ is infinite, or if $\mathfrak{n} = 0$; if $0 < \mathfrak{n} < \omega$, then $\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha = \bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha$ since $\bigcup_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha$ is infinite, so that the equality again follows. Now

$$\begin{aligned}
\mathfrak{n} &= \sum_{\alpha<\mathfrak{n}} 1 && \text{by 21.5}(v),\\
&\leq \sum_{\alpha<\mathfrak{n}} \mathfrak{m}_\alpha && \text{by 20.6}(v);
\end{aligned}$$

together with the above calculation this easily yields $\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}\mathfrak{m}_\alpha = \sum_{\alpha<\mathfrak{n}}\mathfrak{m}_\alpha$. If actually $\mathfrak{m}_\alpha < \mathfrak{m}_\beta$ whenever $\alpha < \beta < \mathfrak{n}$, then $\mathfrak{m}$ is a strictly increasing function; by 12.2, $\alpha \leq \mathfrak{m}_\alpha$ for each $\alpha < \mathfrak{n}$, and so $\mathfrak{n} \subseteq \bigcup_{\alpha<\mathfrak{n}}\mathfrak{m}_\alpha$ and $\mathfrak{n} \leq \bigcup_{\alpha<\mathfrak{n}}\mathfrak{m}_\alpha$. This gives $\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}\mathfrak{m}_\alpha = \bigcup_{\alpha<\mathfrak{n}}\mathfrak{m}_\alpha$.

Corollary 21.17 *If $\langle A_\alpha : \alpha \in \mathfrak{n}\rangle$ is a system of infinite sets such that, for all α, β, $\alpha < \beta < \mathfrak{n}$ implies that $A_\alpha \subseteq A_\beta$, then*

$$\mathfrak{n}\cdot\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| = \mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| = \sum_{\alpha<\mathfrak{n}}|A_\alpha|.$$

If in addition $|A_\alpha| < |A_\beta|$ whenever $\alpha < \beta < \mathfrak{n}$, then $\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| = \bigcup_{\alpha<\mathfrak{n}}|A_\alpha| = \mathfrak{n}\cdot\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right|$.

Proof The equality $\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| = \sum_{\alpha<\mathfrak{n}}|A_\alpha|$ is immediate from 21.16. Since for any $\alpha \in \mathfrak{n}$ we have $A_\alpha \subseteq \bigcup_{\alpha<\mathfrak{n}} A_\alpha$, we get $|A_\alpha| \leq \left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right|$, and so

$$(1)\quad \bigcup_{\alpha<\mathfrak{n}}|A_\alpha| \leq \left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right|.$$

Also, $\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| \leq \sum_{\alpha<\mathfrak{n}}|A_\alpha|$, by Theorem 20.4, so that

$$\begin{aligned}\mathfrak{n}\cdot\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| &\leq \mathfrak{n}\cdot\sum_{\alpha<\mathfrak{n}}|A_\alpha| \\ &= \mathfrak{n}\cdot\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| \\ &= \mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| \\ &\leq \mathfrak{n}\cdot\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| \qquad \text{by (1)},\end{aligned}$$

so that $\mathfrak{n}\cdot\left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| = \mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha|$, as desired. Under the additional assumptions, $\mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| = \bigcup_{\alpha<\mathfrak{n}}|A_\alpha|$, by 21.16, and

$$\begin{aligned}\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| &\leq \left|\bigcup_{\alpha<\mathfrak{n}} A_\alpha\right| \qquad \text{by (1)} \\ &\leq \sum_{\alpha<\mathfrak{n}}|A_\alpha| \\ &= \mathfrak{n}\cdot\bigcup_{\alpha<\mathfrak{n}}|A_\alpha| \\ &= \bigcup_{\alpha<\mathfrak{n}}|A_\alpha|,\end{aligned}$$

and the proof is complete.

Now the equivalent definition of regular cardinal is given by the following theorem.

Theorem 21.18 *For $\mathfrak{m}$ infinite, $\mathfrak{m}$ is regular iff for every system $\langle \mathfrak{n}_i : i \in I\rangle$ of cardinals with $\mathfrak{n}_i < \mathfrak{m}$ for each $i \in I$ and $|I| < \mathfrak{m}$ we have $\sum_{i\in I}\mathfrak{n}_i < \mathfrak{m}$.*

Proof $\Rightarrow$ Suppose the conclusion fails; let I be a set of minimum cardinality for which there is a system $\langle \mathfrak{n}_i : i \in I\rangle$ of cardinals such that $\mathfrak{n}_i < \mathfrak{m}$ for each $i \in I$, $|I| < \mathfrak{m}$, and $\sum_{i\in I}\mathfrak{n}_i \geq \mathfrak{m}$. Clearly I is infinite. Let $|I| = \mathfrak{p}$,

and let γ be a one-one mapping of $\mathfrak{p}$ onto I. For each $\alpha < \mathfrak{p}$, let $\mathfrak{q}_\alpha = \sum_{\beta \leq \alpha} \mathfrak{n}_{\gamma\beta}$; since $\alpha \dotplus 1 < \mathfrak{p} = |I|$, it follows, by the choice of I, that $\mathfrak{q}_\alpha < \mathfrak{m}$ for every $\alpha < \mathfrak{p}$. Furthermore, $\mathfrak{q}_\alpha \leq \mathfrak{q}_\beta$ whenever $\alpha < \beta < \mathfrak{p}$. Hence, by 21.16, $\sum_{\alpha<\mathfrak{p}} \mathfrak{q}_\alpha = \mathfrak{p} \cdot \bigcup_{\alpha<\mathfrak{p}} \mathfrak{q}_\alpha$. But $\mathfrak{q}_\alpha < \mathfrak{m}$ for each $\alpha < \mathfrak{p}$, and $\mathfrak{p} = |I| < m$, so that, by the regularity of $\mathfrak{m}$, $\bigcup_{\alpha<\mathfrak{p}} \mathfrak{q}_\alpha < \mathfrak{m}$. It easily follows that $\mathfrak{p} \cdot \bigcup_{\alpha<\mathfrak{p}} \mathfrak{q}_\alpha < \mathfrak{m}$, so that $\sum_{\alpha<\mathfrak{p}} \mathfrak{q}_\alpha < \mathfrak{m}$. But clearly $\mathfrak{n}_{\gamma\alpha} \leq \mathfrak{q}_\alpha$ for every $\alpha < \mathfrak{p}$, so that

$$\begin{aligned} \sum_{i\in I} \mathfrak{n}_i &= \sum_{\alpha<\mathfrak{p}} \mathfrak{n}_{\gamma\alpha} \\ &\leq \sum_{\alpha<\mathfrak{p}} \mathfrak{q}_\alpha < \mathfrak{m}, \end{aligned}$$

a contradiction of the assumption on I.

$\Leftarrow$ Assume that $\Gamma \subseteq \mathfrak{m}$ and $|\Gamma| < \mathfrak{m}$. Then

$$\left| \bigcup \Gamma \right| \leq \sum_{\alpha\in\Gamma} |\alpha| < \mathfrak{m},$$

so that we must have $\bigcup \Gamma < \mathfrak{m}$, since $\bigcup \Gamma$ is an ordinal.

We will discuss regular and singular cardinals in more detail in Sec. 23.

The next theorem gives an important tool in more advanced cardinal arithmetic.

Theorem 21.19 (*Zermelo's inequality*) *If* $\mathfrak{m}_i < \mathfrak{n}_i$ *for all* $i \in I$, *then*

$$\sum_{i\in I} \mathfrak{m}_i < \prod_{i\in I} \mathfrak{n}_i.$$

Proof First we want to prove the nonstrict inequality $\leq$. By the multiplicative principle, choose $f \in \mathsf{P}_{i\in I} (\mathfrak{n}_i \sim \mathfrak{m}_i)$. We define a function g with domain $\bigcup_{i\in I} \{(i,\alpha) : \alpha \in \mathfrak{m}_i\}$ and range $\subseteq \mathsf{P}_{i\in I} \mathfrak{n}_i$ by setting, for $i \in I$, $\alpha \in \mathfrak{m}_i$, and $j \in I$,

$$[g(i,\alpha)]_j = \begin{cases} f_j & \text{if } j \neq i, \\ \alpha & \text{if } j = i. \end{cases}$$

We claim that g is one-one. Suppose that $g(i,\alpha) = g(j,\beta)$, with $i,j \in I$, $\alpha \in \mathfrak{m}_i$, and $\beta \in m_j$. If $i \neq j$, then $[g(i,\alpha)]_j = f_j \notin \mathfrak{m}_j$, and $[g(j,\beta)]_j = \beta \in \mathfrak{m}_j$, a contradiction. Hence $i = j$. Therefore, $\alpha = [g(i,\alpha)]_i = [g(i,\beta)]_i = \beta$. This proves that g is one-one. Hence the inequality $\leq$ follows.

Keeping the preceding notation, suppose also, to get a contradiction, that $\sum_{j\in I} \mathfrak{m}_i = \prod_{i\in I} \mathfrak{n}_i$. Then there is a one-one function h mapping $\bigcup_{i\in I} \{(i,\alpha) : \alpha \in \mathfrak{m}_i\}$ onto $\mathsf{P}_{i\in I} \mathfrak{n}_i$. Let

$$k = \langle \langle (hx)_i : x \in \{(i,\alpha) : \alpha \in \mathfrak{m}_i\} \rangle : i \in I \rangle.$$

Thus k is a function with domain I such that k_i maps $\{(i,\alpha) : \alpha \in \mathfrak{m}_i\}$

into $\mathfrak{n}_i$ for each $i \in I$. Since $|\{(i,\alpha) : \alpha \in \mathfrak{m}_i\}| = \mathfrak{m}_i < \mathfrak{n}_i$, also $|k_i^*\{(i,\alpha) : \alpha \in \mathfrak{m}_i\}| < \mathfrak{n}_i$, and so $k_i^*\{(i,\alpha) : \alpha \in \mathfrak{m}_i\} \neq \mathfrak{n}_i$. Therefore, again using the multiplicative principle, there is a function l with domain I such that $l_i \in \mathfrak{n}_i \sim k_i^*\{(i,\alpha) : \alpha \in \mathfrak{m}_i\}$ for every $i \in I$. Since, thus, $l \in \mathsf{P}_{i\in I}\, \mathfrak{n}_i$, and h maps onto $\mathsf{P}_{i\in I}\, \mathfrak{n}_i$, there exist $i \in I$ and $\alpha \in \mathfrak{m}_i$ such that $h(i,\alpha) = l$. But then

$$l_i = [h(i,\alpha)]_i = k_i(i,\alpha) \in k_i^*\{(i,\beta) : \beta \in \mathfrak{m}_i\},$$

contradicting the choice of l. This completes the proof.

We conclude the section with an easy pair of theorems on finite sets.

Theorem 21.20 $m \cdot n = m \cdot n$.

Proof We proceed by induction on n. $m \cdot 0 = 0$, and $m \cdot 0 = 0$, so that the case $n = 0$ works. Now assume that $m \cdot n = m \cdot n$. Then

$$\begin{aligned} m \cdot \mathcal{S}n &= m \cdot n \dot{+} m \\ &= m \cdot n + m && \text{by the induction assumption, and 20.9,} \\ &= m \cdot (n + 1) && \text{by 21.8,} \\ &= m \cdot \mathcal{S}n && \text{by 20.9.} \end{aligned}$$

Corollary 21.21 *A finite product of finite cardinals is finite. If A_i is finite for each $i \in I$, and I is finite, then $\mathsf{P}_{i\in I}\, A_i$ is finite.*

EXERCISES

21.22 If $\mathfrak{m}$ is an infinite cardinal, $|A| = \mathfrak{m}$, and $0 < \mathfrak{n} \leq \mathfrak{m}$, show that there is a partition $\mathcal{P}$ of A into $\mathfrak{n}$ pairwise disjoint sets, each of power $\mathfrak{m}$.

21.23 Prove that, for any α, $\sum_{\beta \leq \alpha} \aleph_\beta = \aleph_\alpha$.

21.24 Give an example of a system $\langle A_i : i \in I \rangle$ of distinct infinite sets for which equality fails in Theorem 20.4. Can this be done for I finite?

21.25 Show that $|\alpha \dot{+} \beta| = |\alpha| + |\beta|$.

21.26 Show that $|\alpha \cdot \beta| = |\alpha| \cdot |\beta|$.

21.27 Show that, if $\alpha > 1$ and $\beta \geq \omega$, then $|\alpha^\beta| = |\alpha| \cup |\beta|$.

21.28 If $\mathfrak{m}$ is an infinite cardinal, prove that $\mathfrak{m}$ is an ϵ-number (see Sec. 15).

21.29 Show that, if $\mathfrak{m}_i \leq \mathfrak{n}_i > 1$ for every $i \in I$, then $\sum_{i\in I} \mathfrak{m}_i \leq \prod_{i\in I} \mathfrak{n}_i$.

21.30 Give a direct proof of 21.9, not depending on 21.8.

21.31 Show that for any infinite cardinal $\mathfrak{m}$ there is a function σ mapping $\mathfrak{m} \times \mathfrak{m}$ into $\mathfrak{m}$ such that for all $\alpha \in \mathfrak{m}$, (1) $\sigma(\alpha,0) = \alpha$; (2) $\forall \beta < \mathfrak{m}\, \forall \gamma < \mathfrak{m}[\sigma(\beta,\gamma) = \alpha \Rightarrow \beta \leq \alpha]$; (3) $\{(\beta,\gamma) : \sigma(\beta,\gamma) = \alpha\}$ is finite. *Hint:* Analyze and modify the proof of 21.10.

22 CARDINAL EXPONENTIATION

We finish our exposition of the fundamentals of set theory with a discussion of the third common operation on cardinals, exponentiation.

Definition 22.1 $\mathfrak{m}^{\mathfrak{n}} = |{}^{\mathfrak{n}}\mathfrak{m}|$.

Thus the *cardinal* $\mathfrak{m}^{\mathfrak{n}}$ is, by definition, the power of ${}^{\mathfrak{n}}\mathfrak{m}$, the *set of all functions mapping* $\mathfrak{n}$ *into* $\mathfrak{m}$. We first give a pair of theorems that expand this definition.

Theorem 22.2 $|{}^{J}I| = |I|^{|J|}$.

Proof For brevity let $|I| = \mathfrak{m}$, $|J| = \mathfrak{n}$. Thus $|I|^{|J|} = |{}^{\mathfrak{n}}\mathfrak{m}|$. Thus it suffices to show that ${}^{J}I$ and ${}^{\mathfrak{n}}\mathfrak{m}$ are equipotent. Let f be a one-one function mapping $\mathfrak{m}$ onto I, and let g be a one-one function mapping $\mathfrak{n}$ onto J. As is easily checked, the following function is one-one and maps ${}^{\mathfrak{n}}\mathfrak{m}$ onto ${}^{J}I$:

$$\langle f \circ x \circ g^{-1} : x \in {}^{\mathfrak{n}}\mathfrak{m} \rangle.$$

Corollary 22.3 *If* $|I| = |K|$ *and* $|J| = |L|$, *then* $|{}^{J}I| = |{}^{L}K|$.

Quite trivial properties of exponentiation are given in the following.

Theorem 22.4 (*i*) $\mathfrak{m}^{0} = 1$.
(*ii*) $\mathfrak{m}^{1} = \mathfrak{m}$.
(*iii*) *If* $\mathfrak{m} \neq 0$, *then* $0^{\mathfrak{m}} = 0$.
(*iv*) $1^{\mathfrak{m}} = 1$.
(*v*) $\prod_{i \in I} \mathfrak{m} = \mathfrak{m}^{|I|}$.

Proof (*i*) is clear, since ${}^{0}A = \{0\}$ for any set A. The function $\langle \{(0,\alpha)\} : \alpha \in \mathfrak{m} \rangle$ is one-one and maps $\mathfrak{m}$ onto ${}^{1}\mathfrak{m}$, and this establishes (*ii*). (*iii*) is clear, since ${}^{A}0 = 0$ if A is any nonempty set. (*iv*) follows from the fact that $\langle 0 : \alpha \in \mathfrak{m} \rangle$ is the only member of ${}^{\mathfrak{m}}1$, for any $\mathfrak{m}$. To establish (*v*), we must exhibit a one-one function mapping $\mathsf{P}_{i \in I} \mathfrak{m}$ onto ${}^{I}\mathfrak{m}$, by Theorem 22.2; but actually $\mathsf{P}_{i \in I} \mathfrak{m} = {}^{I}\mathfrak{m}$, so that $I\restriction(\mathsf{P}_{i \in I} \mathfrak{m})$ is the desired function.

Theorem 22.5 *If* $\langle \mathfrak{n}_i : i \in I \rangle$ *is a system of cardinals, then*
(*i*) $\mathfrak{m}^{\sum_{i \in I} \mathfrak{n}_i} = \prod_{i \in I} \mathfrak{m}^{\mathfrak{n}_i}$.
(*ii*) $\left(\prod_{i \in I} \mathfrak{n}_i\right)^{\mathfrak{m}} = \prod_{i \in I} \mathfrak{n}_i{}^{\mathfrak{m}}$.

Proof (*i*) By Theorem 22.2, Corollary 21.3, and Definition 20.1, it suffices to find a one-one function mapping ${}^{A}\mathfrak{m}$ onto $\mathsf{P}_{i \in I} {}^{\mathfrak{n}_i}\mathfrak{m}$, where $A = \bigcup_{i \in I} \{(i,\alpha) : \alpha \in \mathfrak{n}_i\}$. As is easily checked, the following function works:

$$\langle \langle f(i,\alpha) : \alpha \in \mathfrak{n}_i \rangle : i \in I \rangle : f \in {}^{A}\mathfrak{m} \rangle.$$

(*ii*) It suffices to find a one-one function mapping ${}^{\mathfrak{m}}(\mathsf{P}_{i\in I}\,\mathfrak{n}_i)$ onto $\mathsf{P}_{i\in I}\,{}^{\mathfrak{m}}\mathfrak{n}_i$. The following function is as desired:

$$\langle\langle(f_\alpha i : \alpha \in \mathfrak{m}\rangle : i \in I\rangle : f \in {}^{\mathfrak{m}}(\mathsf{P}_{i\in I}\,\mathfrak{n}_i)\rangle.$$

Theorem 22.6 $(\mathfrak{m}^{\mathfrak{n}})^{\mathfrak{p}} = \mathfrak{m}^{\mathfrak{n}\cdot\mathfrak{p}}$.

Proof By Definition 22.1, Theorem 22.2, and Theorem 21.4, it is enough to find a one-one function mapping ${}^{\mathfrak{p}}({}^{\mathfrak{n}}\mathfrak{m})$ onto ${}^{(\mathfrak{p}\times\mathfrak{n})}\mathfrak{m}$; and such a function is

$$\langle\langle f_\alpha(\beta) : (\alpha,\beta) \in \mathfrak{p}\times\mathfrak{n}\rangle : f \in {}^{\mathfrak{p}}({}^{\mathfrak{n}}\mathfrak{m})\rangle.$$

Theorem 22.7 *If* $\mathfrak{m} \leq \mathfrak{p} \neq 0$ *and* $\mathfrak{n} \leq \mathfrak{q}$, *then* $\mathfrak{m}^{\mathfrak{n}} \leq \mathfrak{p}^{\mathfrak{q}}$.

Proof For any $f \in {}^{\mathfrak{n}}\mathfrak{m}$, let $f^+ = f \cup \langle 0 : \alpha \in \mathfrak{q} \sim \mathfrak{n}\rangle$. Thus $f^+ \in {}^{\mathfrak{q}}\mathfrak{p}$, and if $f,g \in {}^{\mathfrak{n}}\mathfrak{m}$ with $f \neq g$, then $f^+ \neq g^+$. Hence $\langle f^+ : f \in {}^{\mathfrak{n}}\mathfrak{m}\rangle$ is a one-one function mapping ${}^{\mathfrak{n}}\mathfrak{m}$ into ${}^{\mathfrak{q}}\mathfrak{p}$. Therefore, by Theorem 18.8, $|{}^{\mathfrak{n}}\mathfrak{m}| \leq |{}^{\mathfrak{q}}\mathfrak{p}|$, and we infer, by Definition 22.1, that $\mathfrak{m}^{\mathfrak{n}} \leq \mathfrak{p}^{\mathfrak{q}}$.

The hypothesis that $\mathfrak{p} \neq 0$ is essential in 22.7, since $0^0 = 1$, and $0^1 = 0$. However, it is only the case in which $\mathfrak{m} = \mathfrak{n} = \mathfrak{p} = 0$ and $\mathfrak{q} \neq 0$ that the implication $(\mathfrak{m} \leq \mathfrak{p} \wedge \mathfrak{n} \leq \mathfrak{q} \Rightarrow \mathfrak{m}^{\mathfrak{n}} \leq \mathfrak{p}^{\mathfrak{q}})$ breaks down. But this implication is of little interest when $\mathfrak{p} = 0$, anyway.

Theorem 22.8 $\aleph_\alpha{}^m = \aleph_\alpha$ *for* $m \neq 0$.

Proof By induction on m, using 21.10.

Theorem 22.9 *For any set* A, $|SA| = 2^{|A|}$.

Proof It is enough to exhibit a one-one correspondence between SA and ${}^A2$. In fact, here we need the notion of a *characteristic function*, which proves useful in many situations in mathematics. With each $B \subseteq A$ we associate its *A-characteristic function* χ_B defined by

$$\chi_B a = \begin{cases} 1 & a \in B, \\ 0 & a \notin B, \end{cases}$$

for every $a \in A$. Then χ is a one-one correspondence between SA and ${}^A2$, as is easily checked.

Combining 22.9 with an earlier observation about SA (Theorem 18.12), we obtain the following important theorem.

Theorem 22.10 $\mathfrak{m} < 2^{\mathfrak{m}}$.

If we recall our earlier definition of $\mathfrak{m}^+$ as the least cardinal greater than $\mathfrak{m}$, 22.10 suggests the obvious question whether $\mathfrak{m}^+ = 2^{\mathfrak{m}}$ or not. For $\mathfrak{m}$

finite this is easily answered: $0^+ = 2^0 = 1$, $1^+ = 2^1 = 2$, but for $\omega > \mathfrak{m} > 1$, $\mathfrak{m}^+ < 2^{\mathfrak{m}}$ (see Exercise 22.28). For $\mathfrak{m}$ infinite the situation is more complicated. The case $\mathfrak{m} = \aleph_0$ is of special interest since it can be shown that the set of real numbers has power $2^{\aleph_0}$. The statement $2^{\aleph_0} = \aleph_1$ is known as the *continuum hypothesis;* the statement $2^{\mathfrak{m}} = \mathfrak{m}^+$ for every infinite cardinal $\mathfrak{m}$ is known as the *generalized continuum hypothesis.* It is known that these two hypotheses are independent of our axioms—neither can be derived from our axioms, nor can their negations be derived from our axioms, assuming always that our axioms are consistent. On the one hand it is consistent to assume that $2^{\aleph_0} = \aleph_1$, and on the other hand it is consistent to assume that $2^{\aleph_0} = \aleph_2$ (or $\aleph_3$, $\aleph_4$, . . . , but not $\aleph_\omega$; see Theorem 22.12). Here again, as in the case of the integers (cf. Remark 13.15), we run into a seeming paradox. For it is known that the set of real numbers can be characterized up to isomorphism, as an ordered field satisfying Dedekind's postulate, but it appears that we definitely get two isomorphic fields of real numbers by settling on one of the two assumptions $2^{\aleph_0} = \aleph_1$ or, say, $2^{\aleph_0} = \aleph_2$. Again, the paradox is only apparent, for the isomorphism of any two fields of real numbers can be proved without taking either assumption as an axiom, although one may say that, within any given fixed conception of set theory, only one of the two possibilities can be true. At any rate, it is not clear which hypothesis, the continuum hypothesis or its negation, should be taken as an axiom. We will see in Sec. 23, however, that under the assumption of the generalized continuum hypothesis cardinal arithmetic is greatly simplified.

In connection with this discussion the following two theorems are significant (see Sec. 23 also).

Theorem 22.11 $\aleph_\omega < \aleph_\omega^{\aleph_0}$.

Proof By Lemma 21.16 and the definition of $\aleph_\omega$, we have $\aleph_\omega = \sum_{m\in\omega} \aleph_m$. Now $\aleph_m < \aleph_\omega$ for every m, so that

$$\begin{aligned} \aleph_\omega &< \prod_{m\in\omega} \aleph_\omega && \text{by Zermelo's inequality (21.19),} \\ &= \aleph_\omega^{\aleph_0} && \text{by 22.4}(v), \end{aligned}$$

as desired.

Corollary 22.12 $2^{\aleph_0} \neq \aleph_\omega$.

Proof If $2^{\aleph_0} = \aleph_\omega$, then $2^{\aleph_0} < (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$, a contradiction.

The following equality is another of the very useful facts in cardinal arithmetic.

Theorem 22.13 *If* $1 < \mathfrak{m} \leq \mathfrak{n} \geq \aleph_0$, *then* $\mathfrak{m}^{\mathfrak{n}} = 2^{\mathfrak{n}}$.

Proof We have

$$\begin{aligned} \mathfrak{m}^{\mathfrak{n}} &\leq (2^{\mathfrak{m}})^{\mathfrak{n}} && \text{by 22.7,} \\ &= 2^{\mathfrak{m}\cdot\mathfrak{n}} && \text{by 22.6,} \\ &= 2^{\mathfrak{n}} && \text{by 21.11,} \\ &\leq \mathfrak{m}^{\mathfrak{n}} && \text{by 22.7,} \end{aligned}$$

and this completes the proof.

We now give two rather special equalities that are sometimes useful in calculations.

Theorem 22.14 (*Hausdorff*) $\aleph_{\alpha+1}^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1}$.

Proof We consider two cases; first suppose that $\alpha \dotplus 1 \leq \beta$. Then $\aleph_{\alpha+1} \leq \aleph_\beta < 2^{\aleph_\beta}$, and so

$$\begin{aligned} \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1} &= 2^{\aleph_\beta} \cdot \aleph_{\alpha+1} && \text{by 22.13,} \\ &= 2^{\aleph_\beta} && \text{by 21.11,} \\ &= \aleph_{\alpha+1}^{\aleph_\beta} && \text{by 22.13.} \end{aligned}$$

The other case is $\alpha \dotplus 1 > \beta$. Now

$$(1) \quad {}^{\aleph_\beta}\aleph_{\alpha+1} = \bigcup_{\gamma<\aleph_{\alpha+1}} {}^{(\aleph_\beta)}\gamma.$$

Indeed, the inclusion $\supseteq$ is obvious. On the other hand, if $f \in {}^{\aleph_\beta}\aleph_{\alpha+1}$, then $|Rng\ f| \leq |Dmn\ f| = \aleph_\beta < \aleph_{\alpha+1}$, so that, by Corollary 21.14, $\bigcup Rng f < \aleph_{\alpha+1}$. With $\gamma = \bigcup Rng f \dotplus 1$ we have, by 19.3(*ii*), $\gamma < \aleph_{\alpha+1}$. Thus $f \in {}^{\aleph_\beta}\gamma$, as desired. Hence (1) holds. It follows that

$$\begin{aligned} \aleph_{\alpha+1}^{\aleph_\beta} &= \left| \bigcup_{\gamma<\aleph_{\alpha+1}} ({}^{\aleph_\beta}\gamma) \right| && \text{by (1)} \\ &\leq \sum_{\gamma<\aleph_{\alpha+1}} |\gamma|^{\aleph_\beta} \\ &\leq \sum_{\gamma<\aleph_{\alpha+1}} \aleph_\alpha^{\aleph_\beta} \\ &= \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1} \\ &\leq \aleph_{\alpha+1}^{\aleph_\beta} \cdot \aleph_{\alpha+1} \\ &= \aleph_{\alpha+1}^{\aleph_\beta}, \end{aligned}$$

and this completes the proof.

Corollary 22.15 (*Bernstein*) $\aleph_m^{\aleph_\alpha} = 2^{\aleph_\alpha} \cdot \aleph_m$.

Proof We proceed by induction on m. We first have $\aleph_0^{\aleph_\alpha} = 2^{\aleph_\alpha} =$

$2^{\aleph_\alpha \cdot \aleph_0}$, using 22.13 and 21.11. Assuming our result for m, we obtain

$$\begin{aligned} \aleph_{m+1}^{\aleph_\alpha} &= \aleph_m^{\aleph_\alpha} \cdot \aleph_{m+1} && \text{by 22.14,} \\ &= 2^{\aleph_\alpha \cdot \aleph_m} \cdot \aleph_{m+1} && \text{by the induction assumption,} \\ &= 2^{\aleph_\alpha} \cdot \aleph_{m+1} && \text{by 21.11.} \end{aligned}$$

This completes the proof.

The following theorem is easily established by induction on n.

Theorem 22.16 $m^{\cdot n} = m^n$.

We have now developed all of the cardinal arithmetic needed for most routine cardinality calculations in mathematics. We conclude this section with two examples of such calculations. More examples are found in the exercises.

Theorem 22.17 *For any infinite set A, A is equipotent with the set of all finite subsets of A.*

Proof We first prove

(1) $|\{F : F \subseteq A, F \text{ finite}\}| \leq \left| \bigcup_{m \in \omega} {}^m A \right|$.

Indeed, for each m and each $f \in {}^m A$ let $Ff = Rng\, f$. Then F is clearly a function with domain $\bigcup_{m \in \omega} {}^m A$ and range $\{F : F \subseteq A, F \text{ finite}\}$, so that (1) follows from Theorem 18.8. Next

(2) $|A| \leq |\{F : F \subseteq A, F \text{ finite}\}|$.

For $\langle \{a\} : a \in A \rangle$ is a one-one function mapping A into $\{F : F \subseteq A, F \text{ finite}\}$. Now, using (1) and (2),

$$\begin{aligned} |A| \leq |\{F : F \subseteq A, F \text{ finite}\}| &\leq \left| \bigcup_{m \in \omega} {}^m A \right| \leq \sum_{m \in \omega} |A|^m \\ &\leq \sum_{m \in \omega} |A| = |A| \cdot \aleph_0 = |A|, \end{aligned}$$

and the proof is complete.

For the last theorem, recall the definition of a *ring with identity* (Definition 17.5).

Theorem 22.18 *For any infinite set A, there are at most $2^{|A|}$ rings with identity of the form $\mathfrak{A} = \langle A, +, \cdot, -, 0, 1 \rangle$.*

Proof A ring with identity $\mathfrak{A}$ of the kind described is a six-termed sequence, i.e., a function with domain 6. If we assign to any such ring $\mathfrak{A}$ the pair

$$((((+, \cdot), -), 0), 1),$$

we define a one-one function from the set of all such $\mathfrak{A}$ into the set

$$({}^{A\times A}A)\times({}^{A\times A}A)\times({}^{A}A)\times A\times A$$

(cf. Definitions 17.5 and 3.11). Hence it suffices to show that the latter set has power at most $2^{|A|}$. And, indeed,

$$\begin{aligned}|({}^{A\times A}A)\times({}^{A\times A}A)\times({}^{A}A)\times A\times A| &= (|A|^{|A|\cdot|A|})\cdot(|A|^{|A|\cdot|A|})\cdot|A|^{|A|}\cdot|A|\cdot|A| \\ &= |A|^{|A|}\cdot|A|^{|A|}\cdot|A|^{|A|}\cdot|A| \\ &= |A|^{|A|} = 2^{|A|},\end{aligned}$$

as desired.

Remark 22.19 For further discussion of the continuum hypothesis see Cohen 1963 to 1964, Gödel 1940 and 1947, and Solovay 1964.

EXERCISES

Prove the following.

22.20 There are $2^{|A|}$ partitions of an infinite set A.

22.21 $\aleph_{\alpha+m}^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta}\cdot\aleph_{\alpha+m}$.

22.22 If $\mathfrak{m}$ is infinite, then $\prod_{\alpha<\mathfrak{m}}\aleph_\alpha = \aleph_{\mathfrak{m}}{}^{\mathfrak{m}}$. *Hint:* Using Zermelo's inequality, the exercise quickly reduces to showing that $(\prod_{\alpha<\mathfrak{m}}\aleph_\alpha)^{\mathfrak{m}} = \prod_{\alpha<\mathfrak{m}}\aleph_\alpha$. This is easily seen by observing that $\prod_{\alpha<\mathfrak{m}}\prod_{\beta<\mathfrak{m}}\aleph_\beta = \prod_{\gamma<\mathfrak{m}}\prod_{\sigma(\beta,\alpha)=\gamma}\aleph_\beta$, where σ is as in Exercise 21.31.

22.23 For every α there is an $\mathfrak{m} > \alpha$ with $\mathfrak{m}^{\aleph_0} = \mathfrak{m}$.

22.24 For every α there is an $\mathfrak{m} > \alpha$ with $\mathfrak{m}^{\aleph_0} > \mathfrak{m}$.

22.25 An infinite set A has $|A|^{\mathfrak{m}}$ subsets of power $\mathfrak{m}$, for any cardinal $\mathfrak{m} \leq |A|$.

22.26 For any infinite cardinal $\mathfrak{m}$, let $A = \{f : f \in {}^{\mathfrak{m}}\mathfrak{m}$, and $\{\alpha : f\alpha \neq 0\}$ is finite$\}$. Then $|A| = \mathfrak{m}$.

22.27 Let $\mathfrak{A} = \langle A,+,\cdot,-,0,1\rangle$ be a field, and let B be an infinite set. Then the number of vector spaces over $\mathfrak{A}$ of the form $\langle B,+',\cdot',-',0'\rangle$ is at most $2^{|A|\cup|B|}$.

22.28 $m^+ < 2^m$ for $m > 1$.

23 REGULAR AND SINGULAR CARDINALS

In this section we give a few further theorems about regular and singular cardinals (see Definition 21.13).

Definition 23.1 *α is cofinal with β if there is a strictly increasing function f with domain β such that*

$$\bigcup_{\gamma<\beta}(f\gamma \dot{+} 1) = \alpha.$$

Thus if α is a successor ordinal $\delta \dotplus 1$, then α is cofinal with 1 (via the function $f = \{(0,\delta)\}$). If α is cofinal with β and α is a limit ordinal, then β is a limit ordinal, and in this case, if f is an increasing function with domain β such that $\bigcup_{\gamma<\beta}(f\gamma \dotplus 1) = \alpha$, then $\bigcup_{\gamma<\beta}(f\gamma \dotplus 1) = \bigcup_{\gamma<\beta} f\gamma$. Obviously an ordinal α is always cofinal with itself. The notion of cofinality is connected with the previously defined notion of regular cardinal by the following.

Theorem 23.2 *For any infinite ordinal α, α is a regular cardinal iff α is not cofinal with any ordinal $< \alpha$.*

Proof $\Rightarrow$ By Definition 21.13.

$\Leftarrow$ By the above remarks, α is a limit ordinal. Now it suffices to prove the following statement:

(1) For any $\Gamma \subseteq \alpha$, if $|\Gamma| < \alpha$, then $\bigcup\Gamma < \alpha$.

Indeed, if (1) holds, then, taking $\Gamma = \alpha$, we infer that $|\alpha| = \alpha$, because $\bigcup\alpha = \alpha$; thus α is a cardinal, and then (1) yields that α is regular.

To prove (1), let h be a one-one correspondence between $|\Gamma|$ and Γ, and for any $\beta < |\Gamma|$ let, by recursion,

$$k\beta = \bigcap\{\gamma : k^*\beta \subseteq \gamma \text{ and } h\beta \leq \gamma\}.$$

Then one sees, by transfinite induction on β, that $k\restriction\beta$ is strictly increasing and $k\beta \in \alpha$ for every $\beta < |\Gamma|$. Thus k itself is strictly increasing, and, since $h\beta \leq k\beta$ for each $\beta < |\Gamma|$, $\bigcup\Gamma \subseteq \bigcup_{\beta<|\Gamma|} k\beta$. Since this latter union is less than α by virtue of the assumption of the theorem, we get $\bigcup\Gamma < \alpha$, as desired.

By the proof of 23.2, we could have defined an *ordinal* α to be regular if condition (1) of the proof held; then we would have as a theorem that every infinite regular ordinal is a cardinal.

We can now extend Theorem 13.9, concerning fixed points of normal functions.

Theorem 23.3 *Let μ be an $\mathfrak{m}$-termed normal function with range $\subseteq \mathfrak{m}$, $\mathfrak{m}$ an infinite regular cardinal $> \aleph_0$. Then for every $\alpha < \mathfrak{m}$ there is a β with $\alpha < \beta < \mathfrak{m}$ and $\mu\beta = \beta$.*

Proof We proceed just as in the proof of Theorem 13.9. We define a function ν by iteration:

$$\nu 0 = \mu(\alpha \dotplus 1);$$
$$\nu(m+1) = \mu\nu m.$$

Also, let $\nu\omega = \bigcup_{m\in\omega} \nu_m$. Note that $\alpha < \alpha \dotplus 1 \leq \mu(\alpha \dotplus 1) = \nu 0$, by Theorem 12.2.

By induction on m, it is easily seen that $\nu m < \mathfrak{m}$ for every $m < \omega$, and so, by the regularity of $\mathfrak{m}$, $\nu\omega = \bigcup_{m<\omega} \nu m < \mathfrak{m}$. Furthermore, ν is a half-normal function, and so

$$\nu\omega \leq \mu\nu\omega = \bigcup_{m\in\omega} \mu\nu m = \bigcup_{m\in\omega} \nu(m + 1) = \nu\omega,$$

so that $\mu\nu\omega = \nu\omega$, as desired.

As with 13.9, the proof of Theorem 23.3 yields the least fixed point $> \alpha$. Also, from Theorem 23.3 itself, it is easy to see that the set of all fixed points has power $\mathfrak{m}$.

Another important property of regular cardinals is given in the following, concerning regressive functions.

Definition 23.4 *Let μ be an A-termed sequence of members of A, where A is an ordinal. We say that μ is* **regressive** *if $\mu 0 = 0$ and $\mu\alpha < \alpha$ for all $\alpha \in A \sim \{0\}$.*

Theorem 23.5 *Let μ be a regressive function mapping A into A, where $A = Ord$ or A is an uncountable regular cardinal. Then there is a $\beta \in A$ such that:*
(i) If $A = Ord$, then $\{\alpha : \mu\alpha = \beta\}$ is a proper class.
(ii) If A is an uncountable regular cardinal, then $|\{\alpha : \mu\alpha = \beta\}| = A$.

Proof We define a function f with domain A by transfinite recursion. Choose $a \notin Ord$ [for example, $a = (0,0)$; see 9.5]. For each $\alpha \in A$, let

$$f0 = \begin{cases} \bigcap\{\beta : \beta \in A \wedge \forall\gamma(\beta \leq \gamma \in A \Rightarrow 0 < \mu\gamma)\} & \text{if such a } \beta \text{ exists,} \\ a & \text{otherwise;} \end{cases}$$

$$f(\alpha \dotplus 1) = \begin{cases} \bigcap\{\beta : \beta \in A \wedge f\alpha < \beta \\ \wedge\ \forall\gamma(\beta \leq \gamma \in A \Rightarrow \alpha < \mu\gamma)\} & \text{if such a } \beta \text{ exists,} \\ a & \text{otherwise;} \end{cases}$$

$$f\alpha = \bigcup_{\beta<\alpha} f\beta \quad \text{if } \alpha = \bigcup\alpha \neq 0.$$

Assume, now, that the conclusion of the theorem fails to hold. Then $a \neq f\alpha \in A$ for all $\alpha \in A$, as we now show by transfinite induction on α. First consider the case $\alpha = 0$. If $f\alpha = a$, then $\forall\beta \in A\exists\gamma(\beta \leq \gamma \in A \wedge \mu\gamma = 0)$, from which it is easy to infer that $\{\gamma : \mu\gamma = 0\}$ is a proper class if $A = Ord$, or has power A if $A \neq Ord$; both conclusions contradict the assumption. Now assume that $a \neq f\alpha \in A$, and $\alpha \dotplus 1 \in A$; we want to show that $a \neq f(\alpha \dotplus 1) \in A$. If this is not the case, then $\forall\beta \in A(f\alpha < \beta \Rightarrow \exists\gamma(\beta \leq \gamma \in A \wedge \mu\gamma \leq \alpha))$. Let $\Gamma_\delta = \{\gamma : f\alpha < \gamma \in A \wedge \mu\gamma = \delta\}$, for each $\delta \leq \alpha$. Then it is easily seen that $\bigcup_{\delta\leq\alpha} \Gamma_\delta$ is a proper class if $A = Ord$, or has power A if $A \neq Ord$. It follows that there is a $\delta \leq \alpha$ such that Γ_δ is a proper class if $A = Ord$, or has power A if $A \neq Ord$, and this contradicts the assumption. Thus $a \neq f(\alpha \dotplus 1) \in A$. The limit step

is trivial, so that $\forall \alpha \in A(a \neq f\alpha \in A)$. It easily follows that

(1) For any $\alpha,\beta \in A$, if $f(\alpha \dot{+} 1) \leq \beta$, then $\alpha < \mu\beta$.
(2) For any $\alpha,\beta \in A$, if $\alpha < \beta$, then $\alpha < \mu f\beta$.

Statement (2) follows from (1) since for $\alpha < \beta$ we have $\alpha \dot{+} 1 \leq \beta$, $f(\alpha \dot{+} 1) \leq f\beta$, and hence $\alpha < \mu f\beta$.

Now f is a normal function, so that, by 13.9 or 23.3, let $\alpha \in A$ be a fixed point of $f : f\alpha = \alpha$. For $\beta < \alpha$ we then have, by (2), $\beta < \mu f\alpha$; hence $\alpha \leq \mu f\alpha = \mu\alpha < \alpha$, a contradiction. This completes the proof.

Definition 23.6 *For any ordinal α, $cf\alpha$, the* ***cofinality character*** *of α, is the least ordinal β such that α is cofinal with β.*

Theorem 23.7 *If α is a limit ordinal, then $cf\alpha$ is a regular cardinal.*

Proof Clearly $cf\alpha \geq \omega$, so that it is enough to use 23.2; to do so, we need to show that $cf\alpha$ is not cofinal with any ordinal $< cf\alpha$. Suppose, on the contrary, that $cf\alpha$ is cofinal with $\beta < cf\alpha$. Then there exist strictly increasing functions f, g with domains $cf\alpha$, β respectively such that $\bigcup_{\gamma<cf\alpha} f_\gamma = \alpha$ and $\bigcup_{\gamma<\beta} g_\gamma = cf\alpha$. But then $f \circ g$ is strictly increasing with domain β and, as is easily checked, $\bigcup_{\gamma<\beta} fg_\gamma = \alpha$. This contradicts the choice of $cf\alpha$.

Note that $cf\aleph_\omega = \aleph_0$; $cf\aleph_{\aleph_1} = \aleph_1$; $cf\aleph_{\aleph_\omega} = \aleph_0$. We can now generalize Theorem 22.11.

Theorem 23.8 *For any infinite cardinal* $\mathfrak{m}$, $\mathfrak{m}^{cf\mathfrak{m}} > \mathfrak{m}$.

Proof Let f be a strictly increasing sequence of members of $\mathfrak{m}$, with $Dmn\, f = cf\mathfrak{m}$ and $\bigcup_{\alpha<cf\mathfrak{m}} f_\alpha = \mathfrak{m}$. Thus

$$\mathfrak{m} = \bigcup_{\alpha<cf\mathfrak{m}} f_\alpha \leq \sum_{\alpha<cf\mathfrak{m}} |f_\alpha| < \prod_{\alpha<cf\mathfrak{m}} \mathfrak{m} = \mathfrak{m}^{cf\mathfrak{m}},$$

as desired.

Theorem 23.8 enables us to compute exponentials of infinite cardinals easily, under the assumption of the generalized continuum hypothesis.

Theorem 23.9 *Assume the generalized continuum hypothesis. and suppose that* $\mathfrak{m}$ *and* $\mathfrak{n}$ *are infinite cardinals. Then*
(*i*) *If* $\mathfrak{n} < cf\mathfrak{m}$, *then* $\mathfrak{m}^{\mathfrak{n}} = \mathfrak{m}$.
(*ii*) *If* $cf\mathfrak{m} \leq \mathfrak{n} \leq \mathfrak{m}$, *then* $\mathfrak{m}^{\mathfrak{n}} = \mathfrak{m}^+$.
(*iii*) *If* $\mathfrak{m} < \mathfrak{n}$, *then* $\mathfrak{m}^{\mathfrak{n}} = \mathfrak{n}^+$.

Proof (*i*) Assume that $\mathfrak{n} < cf\mathfrak{m}$. If $f \in {}^{\mathfrak{n}}\mathfrak{m}$, then $f \in {}^{\mathfrak{n}}\alpha$ for some $\alpha < \mathfrak{m}$.

Hence

$$\begin{aligned} \mathfrak{m} \leq \mathfrak{m}^{\mathfrak{n}} = |{}^{\mathfrak{n}}\mathfrak{m}| &\leq \left| \bigcup_{\alpha<\mathfrak{m}} {}^{\mathfrak{n}}\alpha \right| \\ &\leq \sum_{\alpha<\mathfrak{m}} |\alpha|^{\mathfrak{n}} \\ &\leq \sum_{\alpha<\mathfrak{m}} (|\alpha| \cup \mathfrak{n})^{(|\alpha|\cup\mathfrak{n})} \\ &= \sum_{\alpha<\mathfrak{m}} (|\alpha| \cup \mathfrak{n})^{+} \qquad \text{by 22.13,} \\ &\leq \sum_{\alpha<\mathfrak{m}} \mathfrak{m} = \mathfrak{m}\cdot\mathfrak{m} = \mathfrak{m}, \end{aligned}$$

as desired.

(*ii*) For $cf\mathfrak{m} \leq \mathfrak{n} \leq \mathfrak{m}$ we have

$$\mathfrak{m} < \mathfrak{m}^{cf\mathfrak{m}} \leq \mathfrak{m}^{\mathfrak{n}} \leq \mathfrak{m}^{\mathfrak{m}} = \mathfrak{m}^{+}$$

so that $\mathfrak{m}^{cf\mathfrak{m}} = \mathfrak{m}^{+}$.

(*iii*) follows directly from 22.13.

Definition 23.10 (*i*) $\mathfrak{m}$ *is* ***weakly inaccessible*** *if* $\mathfrak{m} = \aleph_\alpha$ *for some limit ordinal* α, *and* $\mathfrak{m}$ *is regular.*

(*ii*) $\mathfrak{m}$ *is* ***strongly inaccessible*** *if* $\mathfrak{m} > \aleph_0$, $\mathfrak{m}$ *is regular, and* $2^{\mathfrak{n}} < \mathfrak{m}$ *whenever* $\mathfrak{n} < \mathfrak{m}$.

Clearly every strongly inaccessible cardinal is also weakly inaccessible, and the two concepts coincide if we assume the generalized continuum hypothesis. It is consistent with our axioms to assume that inaccessibles do not exist, but it is known to be impossible to prove the consistency of the existence of inaccessibles. Despite these negative results, the notion of an inaccessible cardinal seems natural. They have turned out to play an important role in many mathematical discussions. We shall give only one indication of their importance, in justifying the related concept of a universe, introduced below.

The notion of a strongly inaccessible cardinal is perhaps made a little more intuitive by the following result.

Theorem 23.11 *Let* $\mathfrak{m}$ *be an uncountable cardinal. Then* $\mathfrak{m}$ *is strongly inaccessible iff the following condition holds:*

(*i*) *For every* $\mathfrak{n} < \mathfrak{m}$ *and for every system* $\langle \mathfrak{p}_\alpha : \alpha < \mathfrak{n} \rangle$ *of cardinals with* $\mathfrak{p}_\alpha < \mathfrak{m}$ *for every* $\alpha < \mathfrak{n}$, *we have* $\sum_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha < \mathfrak{m}$ *and* $\prod_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha < \mathfrak{m}$.

Thus $\mathfrak{m}$ is strongly inaccessible iff it cannot be attained by arithmetic operations formulated exclusively in terms of preceding cardinals.

Proof $\Rightarrow$ We have $\sum_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha < \mathfrak{m}$, by 21.18. Second, with $\mathfrak{q} = (\bigcup_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha) \cup \mathfrak{n}$,

$$\begin{aligned} \prod_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha &\leq \prod_{\alpha<\mathfrak{n}} (\bigcup_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha) = (\bigcup_{\alpha<\mathfrak{n}} \mathfrak{p}_\alpha)^{\mathfrak{n}} \\ &\leq \mathfrak{q}^{\mathfrak{q}} = 2^{\mathfrak{q}} < \mathfrak{m}. \end{aligned}$$

$\Leftarrow$ Obviously $\mathfrak{m}$ is regular, by 21.18. For any $\mathfrak{n} < \mathfrak{m}$ we have

$$2^{\mathfrak{n}} = \prod_{\alpha<\mathfrak{n}} 2 < \mathfrak{m},$$

so that $\mathfrak{m}$ is strongly inaccessible.

Closely connected with the notion of an inaccessible cardinal is the following notion of a *universe*.

Definition 23.12 *A is a* ***universe*** *if the following conditions hold.*

(*i*) *A is a set.*
(*ii*) $\omega \in A$.
(*iii*) *For all x, y, if $x \in y \in A$, then $x \in A$.*
(*iv*) *For all x, if $x \in A$, then $Sx \in A$.*
(*v*) *For all x, if $x \subseteq A$ and there is no one-one correspondence between x and A, then $x \in A$.*

This notion is not to be confused with the concept of the universe V, introduced earlier (2.14). The few properties required of a universe are very strong. We give two theorems that indicate their strength.

Theorem 23.13 *Let A be a universe. Then*

(*i*) *If $x \in A$, then $|x| < |A|$.*
(*ii*) $\omega < |A|$.
(*iii*) *If $x \subseteq y \in A$, then $x \in A$.*
(*iv*) *If $x,y \in A$, then $\{x,y\} \in A$.*
(*v*) *If $x,y \in A$, then $(x,y) \in A$.*
(*vi*) *If $x,y \in A$, then $x \times y \in A$.*
(*vii*) *If $x \in A$ and f is a function mapping x into A, then $\mathrm{Rng}\, f \in A$ and $f \in A$.*
(*viii*) *$|A|$ is strongly inaccessible.*
(*ix*) *If $x \in A$, then $\bigcup x \in A$.*
(*x*) *If $\langle x_i : i \in I \rangle$ is a system of members of A, and $I \in A$, then $\bigcup_{i\in I} x_i \in A$ and $\bigcap_{i\in I} x_i \in A$.*

Proof (*i*) If $x \in A$, then $Sx \in A$, by 23.12(*iv*), $Sx \subseteq A$, by 23.12(*iii*), and so $|x| < |Sx| \leq |A|$.

(*ii*) By (*i*), since $\omega \in A$ by 23.12(*ii*).

(*iii*) If $x \subseteq y \in A$, then $x \in Sy \in A$, by 23.12(*iv*), so that $x \in A$, by 23.12(*iii*).

(*iv*) If $x,y \in A$, then $\{x,y\} \subseteq A$, and $|\{x,y\}| \neq |A|$, since $\omega < |A|$, by (*ii*). Thus, by 23.12(*v*), $\{x,y\} \in A$.

(*v*) Directly from (*iv*).

(*vi*) Assume that $x,y \in A$. For any $u \in x$ and $v \in y$, we have $u,v \in A$, by 23.12(*iii*), and hence $(u,v) \in A$, by (*v*). Thus $x \times y \subseteq A$. Now $|A|$ is

an infinite cardinal, by (*ii*). By (*i*), $|x| < |A|$ and $|y| < |A|$. We infer from all of this that $|x \times y| = |x| \cdot |y| < |A|$. Now it follows from 23.12(*v*) that $x \times y \in A$.

(*vii*) Assume that $x \in A$ and f is a function mapping x into A. Then $Rng\, f \subseteq A$ and $|Rng\, f| \leq |x| < |A|$, by (*i*), so that $Rng\, f \in A$, by 23.12(*v*). Hence $f \subseteq x \times Rng\, f \in A$, by (*vi*), so that $f \in A$, by (*iii*).

(*viii*) By (*ii*), $|A|$ is uncountable. $|A|$ is regular. To prove this, suppose, on the contrary, that $cf|A| < |A|$ (recall Definition 23.6 and Theorem 23.2). Then there is a subset of X of A such that $|X| = cf|A|$. By 23.12(*v*), $X \in A$, and then, by (*vii*), ${}^{X}A \subseteq A$. But then

$$\begin{aligned} |A| &< |A|^{cf|A|} \quad \text{by 23.8,} \\ &= |{}^{X}A| \leq |A|, \end{aligned}$$

a contradiction. Therefore $|A|$ is regular. If $\mathfrak{m} < |A|$, then $|X| = \mathfrak{m}$ for some $X \subseteq A$. Applying, in succession, 23.12(*v*), (*iv*), and (*i*), we get $X \in A$; $SX \in A$; $|Sx| < |A|$; i.e., $2^{\mathfrak{m}} < |A|$. Thus (*viii*) holds.

(*ix*) Assume that $x \in A$. For any y, $y \in \bigcup x$ implies that $y \in z \in x \in A$ for some z; hence $y \in z \in A$, by 23.12(*iii*), and finally $y \in A$, by 23.12(*iii*). Thus $\bigcup x \subseteq A$. Furthermore

$$|\bigcup x| \leq \textstyle\sum_{y \in x} |y| < |A|,$$

since $|A|$ is regular, where we use (*i*). Hence, by 23.12(*v*), $\bigcup x \in A$.

(*x*) By 23.12(*v*), $\{x_i : i \in I\} \in A$, so that, by (*ix*), $\bigcup_{i \in I} x_i \in A$. By (*vii*), $\bigcap_{i \in I} x_i \subseteq A$, and $|\bigcap_{i \in I} x_i| < |A|$, by (*i*) and (*viii*). Hence $\bigcap_{i \in I} x_i \in A$, by 23.12(*v*).

Theorem 23.14 *A is a universe iff $A = \boldsymbol{M}_\theta$ for some strongly inaccessible cardinal θ.*

For the definition of the function $\boldsymbol{M}$, see 15.19.

Proof $\Rightarrow$ Let $\theta = |A|$. By 23.13(*viii*), θ is strongly inaccessible, so that we need establish only that $A = \boldsymbol{M}_\theta$. Suppose that $\boldsymbol{M}_\theta \not\subseteq A$. Let x be a set of least rank such that $x \notin A$. Thus $\rho x < \theta$. Now, by transfinite induction on α, one easily establishes

(1) For every $\alpha < \theta$, $|\boldsymbol{M}_\alpha| < \theta$

(see Theorem 15.20). Since $x \in \boldsymbol{M}_{\rho x + 1}$, it follows, using 15.20(*iii*), that $|x| < \theta$. By the choice of x, $y \in A$ for each $y \in x$; i.e., $x \subseteq A$. But 23.12(*v*) now yields $x \in A$, a contradiction. Thus we must have $\boldsymbol{M}_\theta \subseteq A$ after all. Next, we claim

(2) For all x, if $x \in A$, then $\rho x \in A$.

For if (2) is not true, choose $x \in A$ with ρx minimal such that $\rho x \notin A$. Thus $|\{\rho y : y \in x\}| < \theta$, by 23.13($i$), and $\{\rho y : y \in x\} \subseteq A$. Hence, by 23.13($x$), $\bigcup_{y \in x} \rho y \in A$. Let $\alpha = \bigcup_{y \in x} \rho y$. Then $\rho x = \alpha$ or $\rho x = \alpha \dotplus 1$. By 23.13($iv$) and ($ix$), we know that $\alpha \dotplus 1 = \alpha \cup \{\alpha\} \in A$. Thus $\rho x \in A$, a contradiction. This establishes (2).

Now (2) implies that $|\rho x| < \theta$ for each $x \in A$, which in turn implies that $\rho x < |\rho x|^+ < \theta$, since θ is strongly inaccessible. Thus $A \subseteq \boldsymbol{M}_\theta$, and the equality $A = \boldsymbol{M}_\theta$ has been established.

$\Leftarrow$ The conditions of 23.12(i) to (iv) are all easily checked for $A = \boldsymbol{M}_\theta$, θ a strongly inaccessible cardinal. To check 23.12(v), suppose that $x \subseteq \boldsymbol{M}_\theta$, $|x| < |\boldsymbol{M}_\theta|$. As above, in proving (1), it is seen that $|\boldsymbol{M}_\theta| = \theta$. Thus $|x| < \theta$. θ being regular, it follows that $\bigcup_{y \in x} \rho y < \theta$, and hence $\rho x \leq (\bigcup_{y \in x} \rho y) \dotplus 1 < \theta$, as desired. This completes the proof.

Theorems 23.13 and 23.14 indicate that essentially all the usual set-theoretical operations can be performed completely "inside" a given universe; thus, for example, the direct product $\mathsf{P}_{i \in I} x_i$ does not lead outside a universe so long as the x_i's as well as the index set I are in the given universe. For this reason many mathematicians have, implicitly or explicitly, suggested working within universes rather than within the full scope of set theory. A big advantage of this procedure is that in rare instances when it is desirable to go "outside" a universe, it is perfectly legal to do so. For example, one might wish to consider an algebraic structure of the form $\langle A, \times \rangle$, where A is the class of all rings with identity and $\times$ is the operation of direct product. In our full set theory this is illegal: $\langle A, \times \rangle$ cannot be considered as a two-termed sequence with 0-term A and 1-term $\times$, since A and $\times$ are both proper classes. But by restricting A to be the set of all rings with identity in a given universe and $\times$ the operation of direct product restricted to A the construction becomes quite normal. $\langle A, \times \rangle$ then lies outside the given universe, but it can be discussed within the full set theory, or in a "higher universe."

However, as follows from 23.14 and our previous remarks, our axioms do not guarantee the existence of universes. The desired additional axioms can be put in one of two forms, in accordance with the next theorem, whose proof is obvious on the basis of Theorems 23.14 and 15.20(iv).

Theorem 23.15 *The following two conditions are equivalent.*
(i) *For every cardinal* $\mathfrak{m}$, *there is an inaccessible cardinal* $\theta > \mathfrak{m}$.
(ii) *For every set* A, *there is a universe* B *such that* $A \in B$.

Condition 23.15(ii) will be called the *universe axiom*. It has the advan-

tage over 23.15(i) of being formulated in very elementary terms. As indicated in Sec. 1, normally one does not take this as an axiom of set theory, but in special situations it may prove valuable. Concerning its consistency and independence, the remarks above following 23.10 are all applicable.

For the notion of universe see Tarski 1938. For further discussion of inaccessibles see Keisler, Tarski 1963.

EXERCISES

23.16 Strengthen Theorem 23.3 by replacing the assumption that $\mathfrak{m}$ is regular by the condition $\aleph_0 < cf\mathfrak{m}$.

23.17 If $\mathfrak{n} \geq cf\mathfrak{m}$, then $\mathfrak{p}^{\mathfrak{n}} \neq \mathfrak{m}$.

23.18 $2^{\aleph_\alpha} = \aleph_\beta$ iff β is the least ordinal γ such that $\aleph_\gamma^{\aleph_\alpha} < \aleph_{\gamma+1}^{\aleph_\alpha}$.

23.19 The generalized continuum hypothesis is equivalent to the condition that, for all $\mathfrak{m} \geq \aleph_0$, $\mathfrak{m}^{cf\mathfrak{m}} = \mathfrak{m}^+$.

23.20 If $\mathfrak{m}$ is strongly inaccessible, then $\mathfrak{m}^{\mathfrak{n}} = \mathfrak{m}$ for every $\mathfrak{n} < \mathfrak{m}$.

23.21 For any set A the following conditions are equivalent:

(a) A is a universe;

(b) The following conditions hold:

(1) $\omega \in A$.

(2) For all x, f, if $x \in A$ and $f \in {}^xA$, then $\bigcup Rng\, f \in A$.

(3) For all x, y, if $x \in y \in A$, then $x \in A$.

(4) For all x, if $x \in A$, then $Sx \in A$.

24 APPLICATIONS

In this section we give four important theorems of abstract set theory to indicate something of its flavor when developed beyond the basic facts presented so far. Each of the four theorems has been generalized considerably in the literature; each represents a first step in a well-developed area of interest in abstract set theory.

We begin with the notion of almost disjoint sets.

Definition 24.1 *Two sets A and B are* ***almost disjoint*** *if $A \cap B$ is finite.*

Theorem 24.2 (*Sierpinski*) *There is an uncountable family A of infinite pairwise almost disjoint sets of natural numbers.*

Proof Let $C = \bigcup_{m<\omega} ({}^m2)$. Thus C is the set of all finite sequences of 0's and 1's. By induction on m, it is easily seen that $m < 2^m$ for every m, and it follows that $m < |C|$ for every m; that is, C is infinite. Further

$$|C| \leq \sum_{m\in\omega} 2^m \leq \sum_{m\in\omega} \omega = \omega\cdot\omega = \omega;$$

hence $|C| = \omega$. Thus it is sufficient to find an uncountable family of pairwise almost disjoint subsets of C.

For each $x \in {}^{\omega}2$ let $fx = \{c : c \in C, c \subseteq x\}$. Clearly $\bigcup\{fx : x \in {}^{\omega}2\} = C$. Suppose $x,y \in {}^{\omega}2$ and $x \neq y$. Then there is an $m \in \omega$ such that $x_m \neq y_m$. If $c \in fx \cap fy$, then $Dmn\ c \leq m$, for if $m \in Dmn\ c$, we should have, from $c \subseteq x \cap y$, that $x_m = c_m = y_m$. Thus

$$fx \cap fy \subseteq \bigcup_{n \leq m} {}^{n}2,$$

and this union is finite, according to Theorem 20.10. Thus $Rng\ f$ is the desired uncountable family of pairwise almost disjoint subsets of C.

Theorem 24.2 can be found in Sierpinski 1928, and various generalizations of it in Tarski 1928 and Tarski 1929. The results have found varied applications in mathematics.

The second theorem is a classical and basic result in the theory of infinite graphs.

Definition 24.3 *For any set A, let* $\binom{A}{2} = \{X : X \subseteq A, |X| = 2\}$. $\binom{A}{2}$ *is called the* ***complete graph*** *on A; subsets of* $\binom{A}{2}$ *are referred to as (undirected)* ***graphs over*** *A.*

Theorem 24.4 *(Ramsey) If A is infinite and* $\binom{A}{2} = M \cup N$, *then one of the following two conditions must hold.*

(i) There is an infinite subset B of A such that $\binom{B}{2} \subseteq M$.

(ii) There is an infinite subset B of A such that $\binom{B}{2} \subseteq N$.

Proof Assume that *(i)* is false; we shall show that *(ii)* must hold. For any nonempty subset B of A and any $x \in B$ we set $S(B,x) = \{y : y \in B, x \neq y, \text{ and } \{x,y\} \in N\}$. We claim

(1) If B is an infinite subset of A, then there is an $x \in B$ such that $S(B,x)$ is infinite.

To prove (1), let $A = \left\{F : F \subseteq B, \binom{F}{2} \subseteq M\right\}$. Since we are assuming that *(i)* fails, we know that every member of A is finite. It is clear that $\bigcup B \in A$ for any subset B of A simply ordered by inclusion, so that, by Zorn's lemma, A has a maximal element F under inclusion. For any $y \in B \sim F$ we have $F \cup \{y\} \notin A$, by the maximality of F, and it follows

$\left(\text{because } \binom{F}{2} \subseteq M \text{ but } \binom{F \cup \{y\}}{2} \not\subseteq M\right)$ that there is an $x \in F$ such that $\{x,y\} \notin M$; that is, $\{x,y\} \in N$. Restated, this means that $B \sim F \subseteq \bigcup_{x \in F} S(B,x)$. Because $B \sim F$ is infinite, $\bigcup_{x \in F} S(B,x)$ is infinite, and because F is finite, $S(B,x)$ must be infinite for some $x \in F$. This establishes (1).

Now let f be a choice function for nonempty subsets of A. We define functions x, T by recursion. Let x_0 be an element of A such that $S(A,x_0)$ is infinite—such an element exists, by (1)—and let $T_0 = S(A,x_0)$. If x_m and T_m have been defined, we let

$$x_{m+1} = \begin{cases} f\{z : z \in T_m \text{ and } S(T_m,z) \text{ is infinite}\} & \text{if such a } z \text{ exists;} \\ x_0 & \text{otherwise;} \end{cases}$$

and we let $T_{m+1} = S(T_m,x_{m+1})$. One can easily establish, by induction, that for every m, the "otherwise" case above does not happen, T_m is infinite, and $x_{m+1} \in T_m$. Note that for $m < n$ we have $T_n \subseteq T_m$ (induction on n), $x_m \neq x_n$, and $\{x_m,x_n\} \in N$. Thus with $B = Rng\ x$ we have $\binom{B}{2} \subseteq N$, as desired.

For extensions of Ramsey's theorem see Erdös, Rado 1956, and Erdös, Hajnal, Rado 1965. Ramsey's theorem has had many important applications, particularly in metamathematics.

The third topic is the theory of trees, or ramification systems, closely related to graph theory.

Definition 24.5 *A* ***tree*** *is a pair* $\mathfrak{A} = \langle A, \preceq \rangle$ *such that* $\preceq$ *is a partial ordering with field* A, *and for all* $x \in A$, $\{y : y \preceq x\}$ *is well-ordered by* $\preceq$. *A* ***branch*** *of* A *is an* $\subseteq$*-maximal subset of* A *simply ordered by* $\preceq$. *For each* $x \in A$, *the unique ordinal* α *such that for some function* f, $\beta \leq \gamma < \alpha$ *iff* $f\beta \preceq f\gamma \prec x$, *is called the* ***level*** *of the element* x *of* A (*see Theorem* 13.10).

A tree is finitary provided that, for each m, *there is an element of level* m, *but there are only finitely many such elements, although the tree has no elements of infinite level.*

Theorem 24.6 (*König*) *Any finitary tree has an infinite branch.*

Proof Let $A = \langle A, \preceq \rangle$ be a finitary tree. Let f be a choice function for nonempty subsets of A. We define a function x by recursion. Suppose that x_m has been defined for all $m < n$. Then we set

$$x_n = \begin{cases} f\{y : y \text{ is of level } n,\ x_m \prec y \text{ for all } m < n, \text{ and} \\ \qquad\qquad \{z : y \prec z\} \text{ is infinite}\} & \text{if such a } y \text{ exists;} \\ 0 & \text{otherwise.} \end{cases}$$

As usual, we want to establish that the "otherwise" case never occurs; that is, we prove the following statement by induction on n:

(1) x_n is of level n, $x_m \prec x_n$ for all $m < n$, and $\{z : x_n \prec z\}$ is infinite.

Assuming (1) true for all $m < n$, we then have $x_0 \prec \cdots \prec x_{n-1}$. Let $\{y_0, \ldots, y_{p-1}\}$ be all the elements of A of level n all $\succ x_{n-1}$ (in case $n = 0$, only all the elements of level 0). Then

$$\{z : x_{n-1} \prec z\} = \{y_0, \ldots, y_{p-1}\} \cup \bigcup_{i<p} \{z : y_i \prec z\}$$

(in case $n = 0$, the left side of this equation is replaced by A). Since $\{z : x_{n-1} \prec z\}$ (or A, if $n = 0$) is infinite, there is an $i < p$ such that $\{z : y_i \prec z\}$ is infinite. This shows that in our definition of x_n we obtain an element satisfying the conditions of (1).

From (1) it follows, of course, that $\{x_i : i \in \omega\}$ is an infinite branch of A, as desired.

The fourth, and last, topic is a celebrated theorem of Cantor concerning dense orderings. The method of proof is more important than the theorem itself; it has since been applied in many diverse situations.

Definition 24.7 *A dense ordering, without first or last elements, is a simple ordering $\preceq$ such that the following three conditions hold.*
(i) For every $x \in Fld\ (\preceq)$, there is a y such that $x \prec y$.
(ii) For every $y \in Fld\ (\preceq)$, there is an x such that $x \prec y$.
(iii) For all x, y, if $x \prec y$, then there is a z such that $x \prec z \prec y$.

Theorem 24.8 *(Cantor) Any two denumerable dense orderings without first or last elements are isomorphic.*

Proof Let $\preceq$ and $\preceq'$ be denumerable dense orderings without first or last elements, with fields $\{a_m : m \in \omega\}$ and $\{b_m : m \in \omega\}$ respectively, where a and b are one-one functions. We define sequences $\langle c_m : m \in \omega\rangle$ and $\langle d_m : m \in \omega\rangle$ by recursion. Let $c_0 = a_0$ and $d_0 = b_0$. Assume that c_{2m} and d_{2m} have been defined; we now define c_{2m+1}, d_{2m+1}, c_{2m+2}, and d_{2m+2}. Let $c_{2m+1} = a_n$, where n is the first element of $\{p : a_p \notin \{c_0, \ldots, c_{2m}\}\}$. We now distinguish three cases.
Case 1 $a_n \prec c_i$ for each $i \leq 2m$. Let $d_{2m+1} = b_q$, q the least element of $\{p : b_p \prec' d_i$ for each $i \leq 2m\}$.
Case 2 $c_i \prec a_n$ for each $i \leq 2m$. Let $d_{2m+1} = b_q$, q the least element of $\{p : d_i \prec' b_p$ for each $i \leq 2m\}$.
Case 3 $c_i \prec a_n \prec c_j$ for certain $i,j \leq 2m$. Choose i, j so that c_i is a $\preceq$-maximal element of $\{c_k : c_k \prec a_n\}$ and c_j is a $\preceq$-minimal element of

$\{c_k : a_n \prec c_k\}$. If $d_i \not\prec' d_j$, we let $d_{2m+1} = 0$. If $d_i \prec' d_j$, let $d_{2m+1} = b_q$, q the least element of $\{p : d_i \prec' b_p \prec' d_j\}$.

d_{2m+2} and c_{2m+2} are defined similarly, interchanging the roles of $\preceq$ and $\preceq'$, the a's and b's, and the c's and d's. The following facts are now easily established by induction on m:

(1) $\{(c_i,d_i) : i < m\}$ is an isomorphism of $\preceq \cap (c^*m \times c^*m)$ onto $\preceq' \cap (d^*m \times d^*m)$;
(2) $a^*m \subseteq c^*(2m)$ and $b^*m \subseteq d^*(2m)$.

From (1) and (2) it follows at once that $\{(c_i,d_i) : i < \omega\}$ is the desired isomorphism between $\preceq$ and $\preceq'$.

EXERCISES

Prove the following statements.

24.9 Let $\mathfrak{m}$ be an infinite cardinal, and let $\mathfrak{n}$ be the least cardinal such that $\mathfrak{m} < 2^{\mathfrak{n}}$. Suppose that A and I are sets such that $|A| \geq \mathfrak{m}$ and $|I| = \mathfrak{n}$. Then there is a set $F \subseteq {}^IA$ such that $|F| = 2^{\mathfrak{n}}$ and $|\{i : i \in I, fi = gi\}| < \mathfrak{n}$ for any two distinct $f,g \in F$.

24.10 If A is a denumerable set simply ordered by a relation $\preceq$, then A has an infinite subset B such that B is well-ordered either by $\preceq$ or $\succeq$.

There is a set A of power $\aleph_1$ simply ordered by a relation $\preceq$ such that no subset of A of power $\aleph_1$ is well-ordered by $\preceq$ or $\succeq$ (compare Theorem 17.3).

24.11 There is a set A of power $\aleph_1$ such that $\binom{A}{2}$ can be written as $R \cup S$, with neither R nor S including a complete graph on a set of power $\aleph_1$.

APPENDIX
Axiomatic Logic

This appendix is devoted to a rigorous development of a portion of mathematical logic sufficient to found axiomatic set theory on. Because of the many symbols used in axiomatic set theory, it is necessary to give a fairly general description of the underlying logic involved. And because the present precise development of logic is intended to precede the development of set theory in the text, we wish to assume as little as possible; in particular, the mathematical or set-theoretical apparatus we use in this appendix is to be understood strictly in the naïve intuitive sense. For a more thorough treatment of logic using minimum mathematical apparatus we suggest Church 1956.

The fundamental idea in making logic precise is to fix on a formal language, in principle different from English, for which we describe various notions (like *term*, *sentence*, *theorem*, and so on) with complete precision, and no ambiguity as in ordinary languages. Even so, we want the possibility of expanding our language from time to time. Thus in set theory we start with a very limited language, having only the symbol ϵ in addition to logical symbols, but we soon introduce many other symbols, like

$\subseteq$, $\cap$, $\S$, and so on, by definitions. Hence we need to describe, in a precise way, not a single language, but a whole class of languages, which have, however, many essential properties in common.

A given formal language has the following *primitive symbols.*

1 *Individual variables.* $A, B, C, \ldots, Z, A', B', C', \ldots, Z', A'', B'', C'', \ldots$.

2 *Logical symbols.* $\vee, \wedge, \Rightarrow, \neg, \Leftrightarrow, \forall, \exists, =, (,)$.

3 *Operation symbols.* These vary from one language to another in number (perhaps even none are in a particular language), shape, and rank; with each operation symbol there is associated an integer 0, 1, 2, 3, 4, or 5 called its *rank* (it is not necessary here to assume general properties of integers as known, because we speak of the first six integers only). An operation symbol of rank 0 is called an *individual constant.*

4 *Relation symbols.* These again vary from one language to another, and each has a rank that is a *positive* integer 1, 2, 3, 4, or 5.

No other symbols than these are allowed, and the specification of the symbols (along with the ranks of the operation and relation symbols involved) determines the formal language completely.

Given a formal language, we define the notion of a *term*, which corresponds roughly to the notion of a name in ordinary languages.

5 Every individual variable is a term.

6 If O is an operation symbol of rank m, and $\sigma_0, \ldots, \sigma_{m-1}$ are terms, then $O(\sigma_0, \ldots, \sigma_{m-1})$ is a term. In particular, if O is an individual constant, then O itself is a term.

7 Terms can only be formed by finitely many applications of rules 5 and 6.

Next we define the notion of a *formula*, which corresponds roughly to the notion of a sentence in ordinary languages; we give a special technical meaning to the word *sentence* later on.

8 If σ and τ are terms, then $(\sigma = \tau)$ is a formula.

9 If R is a relation symbol of rank m and $\sigma_0, \ldots, \sigma_{m-1}$ are terms, then $R(\sigma_0, \ldots, \sigma_{m-1})$ is a formula.

10 If φ and ψ are formulas, then so are $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, $(\varphi \Rightarrow \psi)$, $(\neg \varphi)$, and $(\varphi \Leftrightarrow \psi)$.

11 If φ is a formula and α is a variable, then $(\forall \alpha \varphi)$ and $(\exists \alpha \varphi)$ are formulas.

12 Formulas can only be constructed by finitely many applications of rules 8 to 11.

To explain the next terminology, *sentential combination*, we must expand the preceding definition. Let $\varphi_0, \ldots, \varphi_{m-1}$ be formulas (again it is enough for our purposes to take $0 < m < 6$).

13 φ_i is a sentential combination of $\varphi_0, \ldots, \varphi_{m-1}$, for $i = 0, \ldots, m-1$.

14 If χ and θ are sentential combinations of $\varphi_0, \ldots, \varphi_{m-1}$, then so are $(\chi \vee \theta)$, $(\chi \wedge \theta)$, $(\chi \Rightarrow \theta)$, $(\neg\chi)$, and $(\chi \Leftrightarrow \theta)$.

15 ψ is a sentential combination of $\varphi_0, \ldots, \varphi_{m-1}$ only by finitely many applications of rules 13 and 14.

Now suppose that values T and F are assigned to $\varphi_0, \ldots, \varphi_{m-1}$ in some way; say i_0 assigned to $\varphi_0, \ldots, i_{m-1}$ assigned to φ_{m-1}, each i_j either T or F. We then assign values to any sentential combination of $\varphi_0, \ldots, \varphi_{m-1}$ by the following rules.

16 φ_j has the value i_j for any $j < m$.

17 If χ has the value k, $k = T$ or $k = F$, then $(\neg\chi)$ has the value F or T respectively.

18 If χ has the value k and θ has the value l, then $(\chi \vee \theta)$, $(\chi \wedge \theta)$, $(\chi \Rightarrow \theta)$, and $(\chi \Leftrightarrow \theta)$ have values given in the following table, depending on k and l:

k	l	$(\chi \vee \theta)$	$(\chi \wedge \theta)$	$(\chi \Rightarrow \theta)$	$(\chi \Leftrightarrow \theta)$
T	T	T	T	T	T
T	F	T	F	F	F
F	T	T	F	T	F
F	F	F	F	T	T

19 Values are assigned to sentential combinations of $\varphi_0, \ldots, \varphi_{m-1}$ only by finitely many applications of rules 16 to 18.

A formula ψ is a *tautology* iff there exist formulas $\varphi_0, \ldots, \varphi_{m-1}$ such that ψ is a sentential combination of $\varphi_0, \ldots, \varphi_{m-1}$ and ψ receives the value T for any assignment of values to $\varphi_0, \ldots, \varphi_{m-1}$.

Some typical examples of tautologies are the following formulas (where φ, ψ, and χ are arbitrary formulas):

$\varphi \Leftrightarrow \varphi$
$\varphi \vee \neg\varphi$
$(\varphi \vee \psi) \Leftrightarrow (\psi \vee \varphi)$
$\varphi \wedge (\psi \vee \chi) \Leftrightarrow (\varphi \wedge \psi) \wedge (\varphi \wedge \chi)$
$\varphi \Rightarrow \varphi \vee \psi$
$\varphi \Rightarrow (\neg\varphi \Rightarrow \psi)$
$\neg(\varphi \wedge \neg\varphi)$
$\varphi \wedge \psi \Rightarrow \varphi$
$\neg(\varphi \wedge \psi) \Leftrightarrow \neg\varphi \vee \neg\psi$
$\neg(\varphi \vee \psi) \Leftrightarrow \neg\varphi \wedge \neg\psi$

Here, typically, we omit various parentheses that should have been put in, in the hope that the formulas are more readable and no confusion results.

We still need more concepts before we can formulate the axioms for logic. A specific occurrence of a variable α in a formula φ is said to be a *bound occurrence of α in φ* provided that the occurrence is within a formula ψ, itself a part of φ, ψ of one of the two forms $(\forall\alpha\chi)$ or $(\exists\alpha\chi)$; we say that the occurrence of α is *within the scope of the quantifier* $\forall\alpha$ or $\exists\alpha$ respectively. If an occurrence of α in φ is not a bound occurrence, then it is called a *free occurrence* of α *in* φ. A formula φ is said to be a *sentence* if no variable occurs free in it.

The axioms for logic (in a fixed but arbitrary language, as with the whole discussion) are as follows, where φ and ψ are arbitrary formulas, α is an arbitrary variable, and σ and τ are arbitrary terms (all subject to limitations mentioned in the axioms).

A1 φ, if φ is a tautology.

A2 $(\exists\alpha\varphi) \Leftrightarrow \neg\forall\alpha\,\neg\varphi$.

A3 $\forall\alpha\varphi \Rightarrow \psi$ provided that ψ is obtained from φ by replacing every free occurrence of α in φ by σ, where no free occurrence of α in φ is within the scope of a quantifier on a variable occurring in σ.

A4 $\forall\alpha(\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \forall\alpha\psi)$ provided that α does not occur free in φ.

A5 $\alpha = \alpha$.

A6 $\sigma = \tau \Rightarrow (\varphi \Leftrightarrow \psi)$ provided that ψ is obtained from φ by replacing one or more occurrences of σ in φ all simultaneously by τ, provided that none of the occurrences of σ in φ have occurrences of variables bound in φ, and similarly for the occurrences of τ in ψ.

At this primitive level of mathematics, it is not enough simply to give axioms. For logic itself we must even specify precisely what we mean by proof and theorem.

In a language we usually expect not only logical axioms but also a body of nonlogical axioms that express the mathematical content of the theory we are interested in. Hence we now assume, in addition to the logical axioms, a collection Γ of *sentences* as given. Now, by a Γ-*formal proof* we mean a finite sequence $\psi_0, \ldots, \psi_{m-1}$ of formulas (where, this time, we must allow m to be quite large in general) such that for each $i < m$ one of the following conditions holds.

20 ψ_i is a logical axiom.
21 ψ_i is a member of Γ.
22 There exist $j,k < i$ such that ψ_j is the formula $\psi_k \Rightarrow \psi_i$ (rule of *modus ponens*, or *detachment*).
23 There is a $j < i$ and a variable α such that ψ_i is the formula $\forall\alpha\psi_j$ (rule of *generalization*).

We then say that $\psi_0, \ldots, \psi_{m-1}$ is a Γ-formal proof of ψ_{m-1}, and we abbreviate the existence of such a Γ-formal proof by writing $\Gamma \vdash \psi_{m-1}$; then ψ_{m-1} is called a Γ-*formal theorem*. If Γ is empty, we write $\vdash \psi_{m-1}$ and call ψ_{m-1} a *formal theorem*.

We now give some common formal theorems that are found useful in practice. Throughout the remainder of the discussion, φ, ψ, χ, ξ denote arbitrary formulas, α, β, γ arbitrary variables, and $\sigma, \tau, \rho, \theta$ arbitrary terms. As a special case of axiom A3 we get

24 $\vdash \forall\alpha\varphi \Rightarrow \varphi$.
25 $\vdash \psi \Rightarrow \exists\alpha\varphi$ if ψ is obtained from φ in the manner, and with the restrictions, of A3.

Indeed, a formal proof of $\psi \Rightarrow \exists\alpha\varphi$ is supplied by the sequence

$\forall\alpha \neg\varphi \Rightarrow \neg\psi$,
$(\forall\alpha \neg\varphi \Rightarrow \neg\psi) \Rightarrow [(\exists\alpha\varphi \Leftrightarrow \neg\forall\alpha \neg\varphi) \Rightarrow (\psi \Rightarrow \exists\alpha\varphi)]$,
$(\exists\alpha\varphi \Leftrightarrow \neg\forall\alpha \neg\varphi) \Rightarrow (\psi \Rightarrow \exists\alpha\varphi)$,
$\exists\alpha\varphi \Leftrightarrow \neg\forall\alpha \neg\varphi$,
$\psi \Rightarrow \exists\alpha\varphi$.

As a special case of 25 we get

26 $\vdash \varphi \Rightarrow \exists\alpha\varphi$.
27 $\vdash \forall\alpha\varphi \Rightarrow \exists\alpha\varphi$.
28 $\vdash \forall\alpha(\varphi \Rightarrow \psi) \Rightarrow (\forall\alpha\varphi \Rightarrow \forall\alpha\psi)$.

To prove 28, observe the following facts in succession:

(*a*)	$\vdash \forall\alpha\varphi \Rightarrow \varphi$	by 24,
(*b*)	$\vdash \forall\alpha(\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \psi)$	by 24,
(*c*)	$\vdash \forall\alpha(\varphi \Rightarrow \psi) \Rightarrow (\forall\alpha\varphi \Rightarrow \psi)$	by (*a*), (*b*), A1,
(*d*)	$\vdash \forall\alpha[\forall\alpha(\varphi \Rightarrow \psi) \Rightarrow (\forall\alpha\varphi \Rightarrow \psi)]$	by generalization,
(*e*)	$\vdash \forall\alpha(\varphi \Rightarrow \psi) \Rightarrow \forall\alpha(\forall\alpha\varphi \Rightarrow \psi)$	by (*d*), A4,
(*f*)	$\vdash \forall\alpha(\forall\alpha\varphi \Rightarrow \psi) \Rightarrow (\forall\alpha\varphi \Rightarrow \forall\alpha\psi)$	by A4,
(*g*)	$\vdash \forall\alpha(\varphi \Rightarrow \psi) \Rightarrow (\forall\alpha\varphi \Rightarrow \forall\alpha\psi)$	by (*e*), (*f*), A1.

29 $\vdash \forall\alpha(\varphi \wedge \psi) \Leftrightarrow \forall\alpha\varphi \wedge \forall\alpha\psi$.

Proof

(*a*)	$\vdash \forall\alpha(\varphi \wedge \psi) \Rightarrow \varphi \wedge \psi$	by 24,
(*b*)	$\vdash \forall\alpha(\varphi \wedge \psi) \Rightarrow \varphi$	by (*a*), A1,
(*c*)	$\vdash \forall\alpha[\forall\alpha(\varphi \wedge \psi) \Rightarrow \varphi]$	by (*b*), generalization,
(*d*)	$\vdash \forall\alpha(\varphi \wedge \psi) \Rightarrow \forall\alpha\varphi$	by (*c*), A4,
(*e*)	$\vdash \forall\alpha(\varphi \wedge \psi) \Rightarrow \forall\alpha\psi$	similarly,
(*f*)	$\vdash \forall\alpha(\varphi \wedge \psi) \Rightarrow \forall\alpha\varphi \wedge \forall\alpha\psi$	by (*d*), (*e*), A1,
(*g*)	$\vdash \forall\alpha\varphi \wedge \forall\alpha\psi \Rightarrow \forall\alpha\varphi$	by A1,
(*h*)	$\vdash \forall\alpha\varphi \wedge \forall\alpha\psi \Rightarrow \varphi$	by (*g*), 24, A1,
(*i*)	$\vdash \forall\alpha\varphi \wedge \forall\alpha\psi \Rightarrow \psi$	similarly,
(*j*)	$\vdash \forall\alpha\varphi \wedge \forall\alpha\psi \Rightarrow \varphi \wedge \psi$	by (*h*), (*i*), A1,
(*k*)	$\vdash \forall\alpha\varphi \wedge \forall\alpha\psi \Rightarrow \forall\alpha(\varphi \wedge \psi)$	by generalization, A4,
(*l*)	$\vdash \forall\alpha(\varphi \wedge \psi) \Leftrightarrow A\alpha\varphi \wedge \forall\alpha\psi$	by (*f*), (*k*), A1.

30 $\vdash \exists\alpha(\varphi \vee \psi) \Leftrightarrow \exists\alpha\varphi \vee \exists\alpha\psi$.

Proof

(*a*)	$\vdash \exists\alpha(\varphi \vee \psi) \Leftrightarrow \neg\forall\alpha\, \neg(\varphi \vee \psi)$	by A2,
(*b*)	$\vdash \neg(\varphi \vee \psi) \Leftrightarrow \neg\varphi \wedge \neg\psi$	by A1,
(*c*)	$\vdash \forall\alpha\, \neg(\varphi \vee \psi) \Leftrightarrow \forall\alpha(\neg\varphi \wedge \neg\psi)$	by 28, A1,
(*d*)	$\vdash \forall\alpha(\neg\varphi \wedge \neg\psi) \Leftrightarrow \forall\alpha\, \neg\varphi \wedge \forall\alpha\, \neg\psi$	by 29,
(*e*)	$\vdash \exists\alpha(\varphi \vee \psi) \Leftrightarrow \exists\alpha\varphi \vee \exists\alpha\psi$	by (*a*), (*c*), (*d*), A2, A1.

31 $\vdash \varphi \Leftrightarrow \forall\alpha\varphi$ provided that α does not occur free in φ.

Proof

(*a*)	$\vdash \forall\alpha\varphi \Rightarrow \varphi$	by 24,
(*b*)	$\vdash \forall\alpha(\varphi \Rightarrow \varphi)$	by A1, generalization,
(*c*)	$\vdash \forall\alpha(\varphi \Rightarrow \varphi) \Rightarrow (\varphi \Rightarrow \forall\alpha\varphi)$	by A4,
(*d*)	$\vdash \varphi \Rightarrow \forall\alpha\varphi$	by (*b*), (*c*),
(*e*)	$\vdash \varphi \Leftrightarrow \forall\alpha\varphi$	by (*a*), (*d*), A1.

32 $\vdash \forall\alpha\, \forall\beta\varphi \Leftrightarrow \forall\beta\, \forall\alpha\varphi$.

Proof

(*a*)	$\vdash \forall\alpha\, \forall\beta\varphi \Rightarrow \forall\beta\varphi$	by 24,
(*b*)	$\vdash \forall\alpha\, \forall\beta\varphi \Rightarrow \varphi$	by (*a*), 24, A1,
(*c*)	$\vdash \forall\alpha(\forall\alpha\, \forall\beta\varphi \Rightarrow \varphi)$	by (*b*), generalization,
(*d*)	$\vdash \forall\alpha\, \forall\beta\varphi \Rightarrow \forall\alpha\varphi$	by (*c*), A4,
(*e*)	$\vdash \forall\alpha\, \forall\beta\varphi \Rightarrow \forall\beta\, \forall\alpha\varphi$	by (*d*), with similar reasoning,
(*f*)	$\vdash \forall\beta\, \forall\alpha\varphi \Rightarrow \forall\alpha\, \forall\beta\varphi$	similarly,
(*g*)	$\vdash \forall\alpha\, \forall\beta\varphi \Leftrightarrow \forall\beta\, \forall\alpha\varphi$	by (*e*), (*f*), A1.

33 $\vdash \exists\alpha\, \forall\beta\varphi \Rightarrow \forall\beta\, \exists\alpha\varphi$.

Proof

(*a*)	$\vdash \forall\beta\varphi \Rightarrow \varphi$	by 24,
(*b*)	$\vdash \varphi \Rightarrow \exists\alpha\varphi$	by 26,
(*c*)	$\vdash \forall\beta\varphi \Rightarrow \exists\alpha\varphi$	by (*a*), (*c*), A1,
(*d*)	$\vdash \neg\exists\alpha\varphi \Rightarrow \neg\forall\beta\varphi$	by (*c*), A1,
(*e*)	$\vdash \neg\exists\alpha\varphi \Rightarrow \forall\alpha\, \neg\forall\beta\varphi$	by generalization, (*d*), A4,
(*f*)	$\vdash \exists\alpha\, \forall\beta\varphi \Rightarrow \exists\alpha\varphi$	by (*e*), A1, A2,
(*g*)	$\vdash \exists\alpha\, \forall\beta\varphi \Rightarrow \forall\beta\, \exists\alpha\varphi$	by (*f*), generalization, A4.

We now give a series of formal theorems concerning equality.

34 If α does not occur in σ, then $\vdash \exists\alpha(\alpha = \sigma)$.

Proof

(*a*)	$\vdash \sigma = \sigma \Rightarrow \exists\alpha(\alpha = \sigma)$	by 25,
(*b*)	$\vdash \forall\alpha(\alpha = \alpha)$	by A5, generalization,
(*c*)	$\vdash \sigma = \sigma$	by (*b*), A3,
(*d*)	$\vdash \exists\alpha(\alpha = \sigma)$	by (*a*), (*c*).

35 $\vdash \alpha = \beta \Rightarrow \beta = \alpha$.

Proof

(*a*)	$\vdash \alpha = \beta \Rightarrow (\alpha = \alpha \Rightarrow \beta = \alpha)$	by A6,
(*b*)	$\vdash \alpha = \alpha \Rightarrow (\alpha = \beta \Rightarrow \beta = \alpha)$	by (*a*), A1,
(*c*)	$\vdash \alpha = \beta \Rightarrow \beta = \alpha$	by (*b*), A5.

36 $\vdash \alpha = \beta \wedge \beta = \gamma \Rightarrow \alpha = \gamma$.

Proof

(*a*)	$\vdash \beta = \alpha \Rightarrow (\beta = \gamma \Rightarrow \alpha = \gamma)$	by A6,
(*b*)	$\vdash \alpha = \beta \Rightarrow \beta = \alpha$	by 35,
(*c*)	$\vdash \alpha = \beta \wedge \beta = \gamma \Rightarrow \alpha = \gamma$	by (*a*), (*b*), A1.

37 $\vdash \sigma = \tau \Rightarrow \rho = \theta$ provided that θ is obtained from ρ by replacing one or more occurrences of σ in ρ by τ all simultaneously.

Proof

(*a*)	$\vdash \sigma = \tau \Rightarrow (\rho = \rho \Rightarrow \rho = \theta)$	by A6,
(*b*)	$\vdash \rho = \rho \Rightarrow (\sigma = \tau \Rightarrow \rho = \theta)$	by (*a*), A1,
(*c*)	$\vdash \rho = \rho$	by A5, generalization, A3,
(*d*)	$\vdash \sigma = \tau \Rightarrow \rho = \theta$	by (*b*), (*c*).

The formal theorems 24 to 37 form a useful basis for beginning the development of the mathematics one is really interested in within a

language, and their proofs are typical of purely logical arguments and should be valuable in trying to prove additional logical theorems that may be needed in a mathematical development. We now want to give a few more results of a deeper nature that will be useful in connecting our development of logic in general with set theory, as developed in the main text.

38 If $\Gamma \vdash \varphi \Leftrightarrow \psi$, then $\Gamma \vdash \chi \Leftrightarrow \xi$ provided that ξ is obtained from χ by replacing one or more occurrences of φ in χ by ψ.

Proof We parallel the definition of formula—8 to 12—proceeding thus by induction on the length of χ. If χ satisfies clause 8 or 9, then we obviously have $\chi = \varphi$ and $\xi = \psi$, and nothing needs proving. Proceeding inductively, suppose that χ is the formula $\chi' \vee \chi''$. Two cases present themselves.

Case 1 φ is $\chi' \vee \chi''$ also, and then ξ is ψ, so that, again, nothing needs proving.

Case 2 ξ has the form $\xi' \vee \xi''$, where ξ' (respectively ξ'') is obtained from χ' (respectively χ'') by replacing zero or more occurrences of φ in χ' (respectively χ'') by ψ. By the induction hypothesis, $\Gamma \vdash \chi' \Leftrightarrow \xi'$ and $\Gamma \vdash \chi'' \Leftrightarrow \xi''$. An easy application of A1 yields $\Gamma \vdash \chi \Leftrightarrow \xi$, as desired.

If χ is a formula $\chi' \wedge \chi''$, $\chi' \Rightarrow \chi''$, $\neg\chi'$, or $\chi' \Leftrightarrow \chi''$, the procedure is the same as the case just treated. Now suppose that χ is the formula $\forall\alpha\chi'$. Again we may ignore the trivial case in which $\varphi = \chi$. In the nontrivial case, ξ has the form $\forall\alpha\xi'$, where ξ' is obtained from χ' by replacing one or more occurrences of φ by ψ. By the induction assumption, $\Gamma \vdash \chi' \Leftrightarrow \xi'$. Generalization, 28, and A1 are then easily used to obtain $\Gamma \vdash \chi \Leftrightarrow \xi$, as desired. The case in which χ is $\exists\alpha\chi'$ is treated similarly. We get $\Gamma \vdash \chi' \Leftrightarrow \xi'$, as above; $\Gamma \vdash \neg\chi' \Leftrightarrow \neg\xi'$, by A1; $\Gamma \vdash \forall\alpha\ \neg\chi' \Leftrightarrow \forall\alpha\ \neg\xi'$, by using 28; and then A1 and A2 yield $\Gamma \vdash \exists\alpha\chi' \Leftrightarrow \exists\alpha\xi'$, as desired.

We are mainly interested in applying 38 in conjunction with the following more elementary result.

39 If β does not occur in φ, and if ψ is obtained from φ by replacing each free occurrence of α in φ by β, then $\vdash \forall\alpha\varphi \Leftrightarrow \forall\beta\psi$.

Proof

(*a*)	$\vdash \forall\alpha\varphi \Rightarrow \psi$	by A3,
(*b*)	$\vdash \forall\alpha\varphi \Rightarrow \forall\beta\psi$	by (*a*), generalization, A4,
(*c*)	$\vdash \forall\beta\psi \Rightarrow \forall\alpha\varphi$	similarly,
(*d*)	$\vdash \forall\alpha\varphi \Leftrightarrow \forall\beta\psi$	by (*b*), (*c*), A1.

Using 38 and 39, we can change bound variables in any formula; for example, if we want to banish all bound occurrences of α in a formula φ, we choose a new variable β, look at an innermost quantifier on α in φ, and replace α there and in the scope of that quantifier by β. 39 together with 38 guarantees that the formula so obtained is provably equivalent to φ. The process is repeated until all bound occurrences of α are wiped out.

The next "deeper" result allows us to eliminate unwanted mathematical axioms (at an expense, of course).

40 (*Deduction metatheorem*) If $\Gamma \cup \{\varphi\}$ is a set of sentences and $\Gamma \cup \{\varphi\} \vdash \psi$, then $\Gamma \vdash \varphi \Rightarrow \psi$.

Proof Let $\chi_0, \ldots, \chi_{m-1}$ be a $(\Gamma \cup \{\varphi\})$-formal proof of ψ. We replace each part χ_i of the sequence to get a Γ-formal proof of $\varphi \Rightarrow \psi$. In accordance with definitions 20 to 23 of formal proofs, one of the following cases holds for an arbitrary part χ_i of the above sequence.
Case 1 χ_i is a logical axiom, or χ_i is a member of Γ. We replace χ_i by the three formulas

$$\chi_i,$$
$$\chi_i \Rightarrow (\varphi \Rightarrow \chi_i),$$
$$\varphi \Rightarrow \chi_i.$$

Case 2 χ_i is φ. We replace χ_i by $\varphi \Rightarrow \varphi$.
Case 3 There exist $j,k < i$ such that χ_j is the formula $\chi_k \Rightarrow \chi_i$. We replace χ_i by the three formulas

$$[\varphi \Rightarrow (\chi_k \Rightarrow \chi_i)] \Rightarrow [(\varphi \Rightarrow \chi_k) \Rightarrow (\varphi \Rightarrow \chi_i)],$$
$$(\varphi \Rightarrow \chi_k) \Rightarrow (\varphi \Rightarrow \chi_i),$$
$$\varphi \Rightarrow \chi_i.$$

Case 4 There exist $j < i$ and a variable α such that χ_i is $\forall \alpha \chi_j$. We replace χ_i by the three formulas

$$\forall \alpha(\varphi \Rightarrow \chi_j),$$
$$\forall \alpha(\varphi \Rightarrow \chi_j) \Rightarrow (\varphi \Rightarrow \forall \alpha \chi_j),$$
$$\varphi \Rightarrow \forall \alpha \chi_j.$$

It is then easy to check that we get a Γ-formal proof of $\varphi \Rightarrow \psi$, as desired.

As discussed in the Introduction, a frequent method of arguing using a mathematical axiom of the form $\exists \alpha \varphi$ is to introduce a "constant" which plays a role of the object assumed to exist in the axiom, deduce a conclusion not involving the object, and then get rid of the constant. We now want to explain and justify this within our axiomatic logic. The

deduction theorem clearly plays a big role here. The following easy result is also important.

41 If σ is an individual constant, α does not occur bound in φ, ψ is obtained from φ by replacing every free occurrence of α in φ by σ, $\Gamma \vdash \psi \Rightarrow \chi$, and σ does not occur in any sentence of Γ, in φ, or in χ, then $\Gamma \vdash \exists \alpha \varphi \Rightarrow \chi$; and also $\Gamma \vdash \psi \Rightarrow \chi$ refers to provability in a language of which σ is a symbol, and $\Gamma \vdash \exists \alpha \varphi \Rightarrow \chi$ to provability in a language of which σ is not a symbol.

Proof Let $\theta_0, \ldots, \theta_{m-1}$ be a Γ-formal proof of $\psi \rightarrow \chi$. Let β be a new variable, one that does not occur in φ, ψ, χ; $\beta \neq \alpha$; β does not occur in any of $\theta_0, \ldots, \theta_{m-1}$. Replace σ by β in each of $\theta_0, \ldots, \theta_{m-1}$, obtaining $\theta'_0, \ldots, \theta'_{m-1}$. By checking A1 to A6 and 20 to 23, we see that $\theta'_0, \ldots, \theta'_{m-1}$ is a Γ-formal proof in a language without σ. Because θ_{m-1} is $\psi \Rightarrow \chi$, θ'_{m-1} is $\psi' \Rightarrow \chi$, where ψ' is obtained from ψ by replacing σ everywhere by β. Thus $\Gamma \vdash \psi' \Rightarrow \chi$ in the σ-less language; moreover, β does not occur in χ, because σ does not occur in χ. Hence $\Gamma \vdash \neg\chi \Rightarrow \neg\psi'$; $\Gamma \vdash \forall\beta(\neg\chi \Rightarrow \neg\psi')$; $\Gamma \vdash \neg\chi \Rightarrow \forall\beta\, \neg\psi'$; $\Gamma \vdash \exists\beta\psi' \Rightarrow \chi$. Now α does not occur in ψ', so that, by change of bound variable, we get $\Gamma \vdash \exists \alpha \varphi \Rightarrow \chi$.

Clearly 41 justifies the intuitive procedure previously described. If, in addition to the assumptions of 41 we have $\Gamma \vdash \exists \alpha \varphi$, then we can conclude that $\Gamma \vdash \chi$.

Only one topic remains to be discussed to found our set-theoretical development properly on axiomatic logic. This is the role of *definitions* in an axiomatic development. Working within a language, we frequently wish to expand it by introducing new symbols, but defining them in terms of the old ones. The procedure differs slightly, depending on whether the new symbol is a relation symbol or an operation symbol. Suppose Γ is a set of sentences in a language L. We *expand L by adjoining a new relation symbol R* (say of rank m) *by a definition* φ provided that the following condition holds.

42 In the language with R adjoined, we take $\Gamma \cup \{\psi\}$ as axioms, where ψ is the sentence $\forall\alpha_0 \forall\alpha_1 \cdots \forall\alpha_{m-1} (\varphi \Leftrightarrow R(\alpha_0, \ldots, \alpha_{m-1}))$, where φ has exactly m variables $\alpha_0, \ldots, \alpha_{m-1}$ occurring free in it ($\alpha_0, \ldots, \alpha_{m-1}$ in their natural alphabetic order).

In case of an operation symbol O (say of rank m), we require the following conditions to hold.

43 φ has exactly $m + 1$ variables $\alpha_0, \ldots, \alpha_m$ occurring free in it ($\alpha_0, \ldots, \alpha_m$ in their natural alphabetic order).

44 $\Gamma \vdash \forall \alpha_0 \cdots \forall \alpha_{m-1} \exists \beta \forall \alpha_m (\varphi \Leftrightarrow \alpha_m = \beta)$, for $\beta \neq \alpha_0, \ldots, \alpha_m$.

45 In the language with O adjoined, we take $\Gamma \cup \{\psi\}$ as axioms, where ψ is the sentence

$$\forall \alpha_0 \cdots \forall \alpha_m (\varphi \Leftrightarrow O(\alpha_0, \ldots, \alpha_{m-1}) = \alpha_m).$$

The following two results are the basic facts about definitions.

46 If L is expanded to L' by adding a definition, then for every formula ψ in L' there is a formula χ of L such that $\Gamma \cup \{\varphi\} \vdash \psi \Leftrightarrow \chi$ (using the above meaning of φ).

Condition 46 is easily proved by induction on the length of ψ.

47 Under the assumptions of 46, if ψ is a formula of L and $\Gamma \cup \{\varphi\} \vdash \psi$, then $\Gamma \vdash \psi$.

Condition 47 is proved essentially by replacing the defined symbol by its definition throughout a $(\Gamma \cup \{\varphi\})$-formal proof of ψ; the details are left to the reader.

This completes our exposition of axiomatic logic.

In the remainder of this appendix we want to indicate how the main part of the text can be brought into the framework of axiomatic logic as we have developed it here. Theoretically, our development of set theory begins within a language with only one nonlogical symbol, the binary relation symbol ϵ [and we have preferred to write $\alpha \epsilon \beta$ instead of $\epsilon(\alpha,\beta)$, as in the general case of this appendix]. A sequence of definitional expansions of this simple language then occurs throughout our book. Although we have not tried to keep the set-theoretical development within the rigid axiomatic logic described here, we hope the reader has little trouble in seeing that this is possible in principle. To aid the reader in this regard, we give a series of disconnected remarks concerning typical difficult points that arise in fitting the text into the axiomatic logic framework.

Remark 1 Definition 1.3 has three parts. The first two may be considered definitions of the one-place relation symbols "is a set" and "is a proper class." The third cannot be put into the framework of this appendix. Instead, one should imagine lowercase letters eliminated throughout the text: $\forall a \varphi(a)$ replaced by $\forall A(A \text{ is a set} \Rightarrow \varphi(A))$, $\exists a \varphi(a)$ by $\exists A(A \text{ is a set} \wedge \varphi(A))$. We could have developed logic with several "sorts" of variables, and then such a replacement would not have been necessary; but such a development is complicated.

Remark 2 The notion of *set-theoretical expression* defined in Definition 1.6 is the same as that of a *formula* in the language with sole nonlogical constant ϵ. Axioms 1.7 have the form

$$\exists\alpha\ \forall\beta(\beta\ \epsilon\ \alpha \Leftrightarrow \beta \text{ is a set } \wedge\ \varphi),$$

where φ is a formula, α and β are distinct variables, and α does not occur in φ. The symbolism $\{X : \varphi(X)\}$ again does not fit into the framework of this appendix; one should replace phrases like "let $A = \{X : \varphi(X)\}$" in the text by "let A be a class such that $\forall X(X\ \epsilon\ A \Leftrightarrow X$ is a set $\wedge\ \varphi(X))$." Again we could have developed logic so as to encompass this notion formally (technically, by using a description operator), but we preferred simplicity.

Remark 3 $\S$, introduced in Definition 1.20, is a one-place operation symbol. It should be distinguished from the *function* $F = \langle \S x : x \text{ a set}\rangle$, which is a certain set of ordered pairs.

Remark 4 Definition 1.24 introduces a binary operation symbol $\{\ ,\ \}$.

Remark 5 Definition 3.9 introduces a binary relation symbol R *is on* A. Of course we do not adhere closely to a formal language in the text, and we allow a phrase such as "let R be on A."

Remark 6 In Definition 4.9, $: \rightarrow$ is to be considered a ternary relation symbol.

Remark 7 Definition 4.17(*i*) coincides, of course, with the definition of term given in this appendix. It is necessary to eliminate all lowercase letters in favor of capital letters in the use of Definition 4.17 (as in Remark 1), and one must eliminate uses of Definition 4.17(*ii*) and (*iii*) entirely, as in Remark 2.

Remark 8 Definitions 5.4 and 5.16 again cannot be formalized in the logic described in this appendix. $\bigcup_{i\in I} A_i$ should always be eliminated in favor of $\bigcup Rng\ A$, for example.

Remark 9 Definition 6.1 defines $\mathsf{P}A$ only for functions A. P is to be treated as an operation symbol; hence, to make the definition completely rigorous, $\mathsf{P}A$ should be defined for any class A. How it is defined for classes A that are not functions is of no importance. A similar comment applies to some later definitions.

Remark 10 As in Remark 1, the use of special variables α, β, γ, . . . , as introduced in Definition 9.12, would have to be proscribed.

Axioms of Set Theory

1.2 Extensionality $\forall A\ \forall B[\forall C(C \in A \Leftrightarrow C \in B) \Rightarrow A = B]$.

1.7 Class-building $\exists A\ \forall X(X \in A \Leftrightarrow X \text{ is a set } \wedge \varphi(X))$, where A does not occur in $\varphi(X)$.

1.12 Power set $\forall a\ \exists b\ \forall C(C \subseteq a \Rightarrow C \in b)$.

1.13 Pairing $\forall a\ \forall b\ \exists c(a \in c \wedge b \in c)$.

1.14 Union $\forall a\ \exists b\ \forall C(C \in a \Rightarrow C \subseteq b)$.

1.18 Regularity $\forall A[A \neq 0 \Rightarrow \exists X(X \in A \wedge X \cap A = 0)]$.

1.23 Infinity $\exists a[0 \in a \wedge \forall X(X \in a \Rightarrow \mathcal{S}X \in a)]$.

1.35 Substitution If F is a function and $Dmn\ F$ is a set, then $Rng\ F$ is a set.

1.36 Relational axiom of choice If R is a relation, then there is a function F such that $F \subseteq R$ and $Dmn\ F = Dmn\ R$.

Bibliography

***Bachmann*, 1967.** *Transfinite Zahlen*, Berlin, 228 pp.

***Bernays, Fraenkel*, 1958.** *Axiomatic Set Theory*, Amsterdam, 226 pp.

***Birkhoff, MacLane*, 1965.** *A Survey of Modern Algebra*, New York, 437 pp. (third edition).

***Bourbaki*, 1939–*present*.** *Eléments de Mathématique*, many volumes, Paris.

***Chin, Tarski*, 1951.** *Distributive and modular laws in the arithmetic of relation algebras*, Univ. Calif. Publ. Math., v. 1, 341–384.

***Church*, 1941.** *The calculi of lamda-conversion*, Ann. of Math. Studies, no. 6, Princeton, 77 pp.

***Church*, 1956.** *Introduction to Mathematical Logic*, Princeton, 376 pp.

***Cohen*, 1963–1964.** *The independence of the continuum hypothesis*, Proc. Nat. Acad. Sci., v. 50, 1143–1148; v. 51, 105–110.

***Cohen*, 1966.** *Set Theory and the Continuum Hypothesis*, New York, 154 pp.

***Dwinger*, 1961.** *Introduction to Boolean Algebras*, Würzburg, 60 pp.

***Erdös, Hajnal, Rado*, 1965.** *Partition relations for cardinal numbers*, Acta Math. Hung., 16, 93–196.

Erdös, Rado, **1956.** *A partition calculus in set theory,* Bull. Amer. Math. Soc., 62, 427–489.

Feferman, **1960.** *Arithmetization of metamathematics in a general setting,* Fund. Math., v. 49, 35–92.

Fraenkel, Bar-Hillel, **1958.** *Foundations of Set Theory,* Amsterdam, 415 pp.

Gentzen, **1936.** *Die Widerspruchsfreiheit der reinen Zahlentheorie,* Math. Ann., 112, 493–565.

Gödel, **1931.** *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I,* Monatsh. Math., 38, 173–198.

Gödel, **1940.** *The Consistency of the Axiom of Choice and of the Generalized Continuum-hypothesis with the Axioms of Set Theory,* Princeton, 69 pp.

Gödel, **1947.** *What is Cantor's continuum problem?* Amer. Math. Monthly, 54, 515–525.

Goodstein, **1963.** *Boolean Algebra,* New York, 140 pp.

Halmos, **1950.** *Measure Theory,* xi + 304 pp.

Halmos, **1960.** *Naive Set Theory,* Princeton, 104 pp.

Halmos, **1963.** *Boolean Algebras,* Princeton, 147 pp.

Hausdorff, **1914.** *Grundzüge der Mengenlehre,* viii + 476 pp.

Henkin, **1960.** *On mathematical induction,* Amer. Math. Monthly, 67, 323–338.

Heyting, **1966.** *Intuitionism,* Amsterdam, 137 pp.

Janiczak, **1952.** *Undecidability of some simple formalized theories,* Fund. Math., 40, 131–139.

Jónsson, **1953.** *On the representation of lattices,* Math. Scand., v. 1, 193–206.

Jónsson, Tarski, **1952.** *Boolean algebras with operators II,* Amer. J. Math., v. 74, 127–162.

Keisler, Tarski, **1964.** *From accessible to inaccessible cardinals,* Fund. Math., 53, 225–308.

Kelley, **1955.** *General Topology,* New York, 298 pp.

Klaua, **1964.** *Allgemeine Mengenlehre,* Berlin, 581 pp.

Kuratowski, **1966.** *Topology,* New York, 560 pp.

Kuroš, Livšic, Šulgeifer, **1960.** *Foundations of the theory of categories,* Uspehi Mat. Nauk, 15, 3–52 (Russian); Russian Math. Surv., 15, 1–45 (English transl.).

Lyndon, **1961.** *Relation algebras and projective geometries,* Mich. Math. J., v. 8, 21–28.

Mates, **1965.** *Elementary Logic,* New York, 227 pp.

Mendelson, **1964.** *Introduction to Mathematical Logic,* Princeton, 300 pp.

Monk, **1964.** *On representable relation algebras,* Mich. Math. J., v. 11, 207–210.

Montague, Scott, Tarski, **1956**. *Abstracts on the notion of rank*, Bull. Amer. Math. Soc., 62.

Rosser, Turquette, **1952**. *Many-valued Logics*, Amsterdam, 124 pp.

Rubin, **1967**. *Set Theory for the Mathematician*, San Francisco, 387 pp.

Rubin, Rubin, **1963**. *Equivalents of the Axiom of Choice*, Amsterdam, 134 pp.

Russell, Whitehead, **1925–1927**. *Principia Mathematica*, 3 vols., Cambridge, 674 pp., 742 pp., 491 pp. (second edition).

Schröder, **1895**. *Vorlesungen über die Algebra der Logik (exakte Logik)*, vol. 3, part 1, Leipzig, 819 pp.

Sierpinski, **1928**. *Sur une décomposition d'ensembles*, Monatsh. Math., 35, 239–248.

Sierpinski, **1965**. *Cardinal and Ordinal Numbers*, Warsaw, 491 pp.

Sikorski, **1964**. *Boolean Algebras*, Berlin, 237 pp. (second edition).

Solovay, **1964**. $2^{\aleph_0}$ *can be anything it ought to be*, in *The Theory of Models*, p. 435.

Suppes, **1960**. *Axiomatic Set Theory*, Princeton, 265 pp.

Tarski, **1924**. *Sur les ensembles finis*, Fund. Math., 6, 45–95.

Tarski, **1928**. *Sur la décomposition des ensembles en sous-ensembles presque disjoints*, Fund. Math., 12, 186–205.

Tarski, **1929**. *Sur la décomposition des ensembles en sous-ensembles presque disjoints (Supplément à la note sous le même titre)*, Fund. Math., 14, 205–215.

Tarski, **1938**. *Über unerreichbare Kardinalzahlen*, Fund. Math., 30, 68–89.

Tarski, **1941**. *On the calculus of relations*, J. Symb. Logic, 6, 73–89.

Tarski, **1949**. *Cardinal Algebras*, Oxford, 327 pp.

Tarski, **1956**. *Ordinal Algebras*, Amsterdam, 133 pp.

Tarski, **1965**. *Introduction to Logic and to the Methodology of Deductive Sciences*, New York, 252 pp.

Van der Waerden, **1940**. *Modern Algebra*, New York, 264 pp.

Vopěnka, **1964**. *Independence of the continuum hypothesis*, Comment. Math. Univ. Carol.; A. M. S. Translations 1967.

Index of Notations[1]

[1] Entries are listed in the order in which they are defined in the text.

Subject Index