PREFACE

What this book is about. The *theory of sets* is a vibrant, exciting mathematical theory, with its own basic notions, fundamental results and deep open problems, and with significant applications to other mathematical theories. At the same time, *axiomatic set theory* is often viewed as a *foundation of mathematics*: it is alleged that all mathematical objects are sets, and their properties can be derived from the relatively few and elegant axioms about sets. Nothing so simple-minded can be quite true, but there is little doubt that in standard, current mathematical practice, “making a notion precise” is essentially synonymous with “defining it in set theory”. Set theory is the official language of mathematics, just as mathematics is the official language of science.

Like most authors of elementary, introductory books about sets, I have tried to do justice to both aspects of the subject.

From straight set theory, these Notes cover the basic facts about “abstract sets”, including the Axiom of Choice, transfinite recursion, and cardinal and ordinal numbers. Somewhat less common is the inclusion of a chapter on “pointsets” which focuses on results of interest to analysts and introduces the reader to the Continuum Problem, central to set theory from the very beginning. There is also some novelty in the approach to cardinal numbers, which are brought in very early (following Cantor, but somewhat deviously), so that the basic formulas of cardinal arithmetic can be taught as quickly as possible. Appendix **A** gives a more detailed “construction” of the real numbers than is common nowadays, which in addition claims some novelty of approach and detail. Appendix **B** is a somewhat eccentric, mathematical introduction to the study of *natural models* of various set theoretic principles, including Aczel’s Antifoundation. It assumes no knowledge of logic, but should drive the serious reader to study it.

About set theory as a foundation of mathematics, there are two aspects of these Notes which are somewhat uncommon. First, I have taken seriously this business about “everything being a set” (which of course it is not) and have tried to make sense of it in terms of the notion of *faithful representation* of mathematical objects by *structured sets*. An old idea, but perhaps this is the first textbook which takes it seriously, tries to explain it, and applies it consistently. Those who favor category theory will recognize some of its basic notions in places, shamelessly folded into a traditional set theoretical approach to the foundations where categories are never mentioned. Second, *computation theory* is viewed as part of the mathematics "to be founded" and the relevant set theoretic results have been included, along with several examples. The ambition was to explain what every young mathematician or theoretical computer scientist needs to know about sets.

The book includes several historical remarks and quotations which in some places give it an undeserved scholarly gloss. All the quotations (and most of the comments) are from papers reprinted in the following two, marvellous and easily accessible source books, which should be perused by all students of set theory:

Georg Cantor, *Contributions to the founding of the theory of transfinite numbers*, translated and with an Introduction by Philip E. B. Jourdain, Dover Publications, New York.

Jean van Heijenoort, *From Frege to Gödel*, Harvard University Press, Cambridge, 1967.

How to use it. About half of this book can be covered in a Quarter (ten weeks), somewhat more in a longer Semester. Chapters **1** – **6** cover the beginnings of the subject and they are written in a leisurely manner, so that the serious student can read through them alone, with little help. The trick to using the Notes successfully in a class is to cover these beginnings very quickly: skip the introductory Chapter **1**, which mostly sets notation; spend about a week on Chapter **2**, which explains Cantor's basic ideas; and then proceed with all deliberate speed through Chapters **3** – **6**, so that the theory of well ordered sets in Chapter **7** can be reached no later than the sixth week, preferably the fifth. Beginning with Chapter **7**, the results are harder and the presentation is more compact. How much of the "real" set theory in Chapters **7** – **12** can be covered depends, of course, on the students, the length of the course, and what is passed over. If the class is populated by future computer scientists, for example, then Chapter **6** on Fixed Points should be covered in full, with its problems, but Chapter **10** on Baire Space might be omitted, sad as that sounds. For budding young analysts, at the other extreme, Chapter **6** can be cut off after **6.27** (and this too is sad), but at least part of Chapter **10** should be attempted. Additional material which can be left out, if time is short, includes the detailed development of addition and multiplication on the natural numbers in Chapter **5**, and some of the less central applications of the Axiom of Choice in Chapter **9**. The Appendices are quite unlikely to be taught in a course (I devote just one lecture to explain the idea of the construction of the reals in Appendix **A**), though I would like to think that they might be suitable for undergraduate Honors Seminars, or individual reading courses.

Since elementary courses in set theory are not offered regularly and they are seldom long enough to cover all the basics, I have tried to make these Notes accessible to the serious student who is studying the subject on their own. There are numerous, simple Exercises strewn throughout the text, which test understanding of new notions immediately after they are introduced. In class I present about half of them, as examples, and I assign some of the rest for easy homework. The Problems at the end of each chapter vary widely in difficulty, some of them covering additional material. The hardest problems are marked with an asterisk (*).

Acknowledgments. I am grateful to the Mathematics Department of the University of Athens for the opportunity to teach there in Fall 1990, when I wrote the first draft of these Notes, and especially to Prof. A. Tsarpalias who usually teaches that Set Theory course and used a second draft in Fall 1991; and to Dimitra Kitsiou and Stratos Paschos for struggling with PCs and laser printers at the Athens Polytechnic in 1990 to produce the first "hard copy" version. I am grateful to my friends and colleagues at UCLA and Caltech (hotbeds of activity in set theory) from whom I have absorbed what I know of the subject, over many years of interaction. I am especially grateful to my wife Joan Moschovakis and my student Darren Kessner for reading large parts of the preliminary edition, doing the problems and discovering a host of errors; and to Larry Moss who taught out of the preliminary edition in the Spring Term of 1993, found the remaining host of errors and wrote out solutions to many of the problems.

The book was written more-or-less simultaneously in Greek and English, by the magic of bilingual LATEXand in true reflection of my life. I have dedicated it to Prof. Nikos Kritikos (a student of Caratheodory), in fond memory of many unforgettable hours he spent with me back in 1973, patiently teaching me how to speak and write mathematics in my native tongue, but also much about the love of science and the nature of scholarship. In this connection, I am also greatly indebted to Takis Koufopoulos, who read critically the preliminary Greek version, corrected a host of errors and made numerous suggestions which (I believe) improved substantially the language of the final Greek draft.

Palaion Phaliron, Greece November 1993

About the 2nd edition. Perhaps the most important changes I have made are in small things, which (I hope) will make it easier to teach and learn from this book: simplifying proofs, streamlining notation and terminology, adding a few diagrams, rephrasing results (especially those justifying *definition by recursion*) to ease their applications, and, most significantly, correcting errors, typographical and other. For spotting these errors and making numerous, useful suggestions over the years, I am grateful to Serge Bozon, Joel Hamkins, Peter Hinman, Aki Kanamori, Joan Moschovakis, Larry Moss, Thanassis Tsarpalias and many, many students.

The more substantial changes include:

— A proof of *Suslin's Theorem* in Chapter **10**, which has also been significantly massaged.

— A better exposition of ordinal theory in Chapter **12** and the addition of some material, including the basic facts about ordinal arithmetic.

— The last chapter, a compilation of solutions to the Exercises in the main part of the book – in response to popular demand. This eliminates the most obvious, easy homework assignments, and so I have added some easy problems.

I am grateful to Thanos Tsouanas, who copy-edited the manuscript and caught the worst of my mistakes.

Palaion Phaliron, Greece July 2005

CHAPTER 2

EQUINUMEROSITY

After these preliminaries, we can formulate the fundamental definitions of Cantor about the size or cardinality of sets.

2.1. Definition. Two sets A, B are **equinumerous** or **equal in cardinality** if there exists a (one-to-one) correspondence between their elements, in symbols

$$A =_c B \iff_{\text{df}} (\exists f)[f : A \rightarrowtail\!\!\!\!\twoheadrightarrow B].$$

This definition of equinumerosity stems from our intuitions about finite sets, e.g., we can be sure that a shoe store offers for sale the same number of left and right shoes without knowing exactly what that number is: the correspondence of each left shoe with the right shoe in the same pair establishes the equinumerosity of these two sets. The radical element in Cantor's definition is the proposal to accept the existence of such a correspondence as the characteristic property of equinumerosity for all sets, despite the fact that its application to infinite sets leads to conclusions which had been viewed as counterintuitive. A finite set, for example, cannot be equinumerous with one of its proper subsets, while the set of natural numbers $\mathbb{N}$ is equinumerous with $\mathbb{N} \setminus \{0\}$ via the correspondence $(x \mapsto x + 1)$,

$$\{0, 1, 2, \dots\} =_c \{1, 2, 3, \dots\}.$$

In the real numbers, also,

$$(0, 1) =_c (0, 2)$$

via the correspondence $(x \mapsto 2x)$, where as usual, for any two reals $\alpha < \beta$

$$(\alpha, \beta) = \{x \in \mathbb{R} \mid \alpha < x < \beta\}.$$

We will use the analogous notation for the closed and half-closed intervals $[\alpha, \beta]$, $[\alpha, \beta)$, etc.

2.2. Proposition. *For all sets A, B, C,*

$$A =_c A,$$
$$\textit{if } A =_c B, \textit{ then } B =_c A,$$
$$\textit{if } (A =_c B \,\&\, B =_c C), \textit{ then } A =_c C.$$

Proof. To show the third implication as an example, suppose that the bijections $f : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ and $g : B \rightarrowtail\!\!\!\!\twoheadrightarrow C$ witness the equinumerosities of the hypothesis; their composition $gf : A \rightarrowtail\!\!\!\!\twoheadrightarrow C$ then witnesses that $A =_c C$. $\dashv$

$$\begin{array}{cccccccccc} \pi : & 3 & 1 & 0 & \cancel{1} & 5 & \cancel{3} & 9 & 7 & \cdots \\ & \uparrow & \uparrow & \uparrow & & \uparrow & & \uparrow & \uparrow & \\ & f(0) & f(1) & f(2) & & f(3) & & f(4) & f(5) & \cdots \end{array}$$

FIGURE **2.1**. Deleting repetitions.

2.3. Definition. The set A is **less than or equal to** B **in size** if it is equinumerous with some subset of B, in symbols:

$$A \leq_c B \iff (\exists C)[C \subseteq B \,\&\, A =_c C].$$

2.4. Proposition. $A \leq_c B \iff (\exists f)[f : A \rightarrowtail B]$.

PROOF. If $A =_c C \subseteq B$ and $f : A \rightarrowtail\!\!\!\!\rightarrow C$ witnesses this equinumerosity, then f is an injection from A into B. Conversely, if there exists an injection $f : A \rightarrowtail B$, then the same f is a bijection of A with its image $f[A]$, so that $A =_c f[A] \subseteq B$ and so $A \leq_c B$ by the definition. ⊣

2.5. Exercise. *For all sets* A, B, C,

$$A \leq_c A,$$
$$\textit{if } (A \leq_c B \,\&\, B \leq_c C), \textit{ then } A \leq_c C.$$

2.6. Definition. A set A is **finite** if there exists some natural number n such that

$$A =_c \{i \in \mathbb{N} \mid i < n\} = \{0, 1, \dots, n-1\},$$

otherwise A is **infinite**. (Thus the empty set is finite, since $\emptyset = \{i \in \mathbb{N} \mid i < 0\}$.)

A set A is **countable** if it is finite or equinumerous with the set of natural numbers $\mathbb{N}$, otherwise it is **uncountable**. Countable sets are also called **denumerable**, and correspondingly, uncountable sets are **non-denumerable**.

2.7. Proposition. *The following are equivalent for every set* A:

(1) A *is countable.*

(2) $A \leq_c \mathbb{N}$.

(3) *Either* $A = \emptyset$, *or* A *has an* **enumeration**, *a surjection* $\pi : \mathbb{N} \twoheadrightarrow A$, *so that*

$$A = \pi[\mathbb{N}] = \{\pi(0), \pi(1), \pi(2), \dots\}.$$

PROOF. We give what is known as a "round robin proof".

(1) $\Longrightarrow$ (2). If A is countable, then either $A =_c \{i \in \mathbb{N} \mid i < n\}$ for some n or $A =_c \mathbb{N}$, so that, in either case, $A =_c C$ for some $C \subseteq \mathbb{N}$ and hence $A \leq_c \mathbb{N}$.

(2) $\Longrightarrow$ (3). Suppose $A \neq \emptyset$, choose some $x_0 \in A$, and assume by (2) that $f : A \rightarrowtail \mathbb{N}$. For each $i \in \mathbb{N}$, let

$$\pi(i) = \begin{cases} x_0, & \text{if } i \notin f[A], \\ f^{-1}(i), & \text{otherwise, i.e., if } i \in f[A]. \end{cases}$$

The definition works (because f is an injection, and so $f^{-1}(i)$ is uniquely determined in the second case), and it defines a surjection $\pi : \mathbb{N} \twoheadrightarrow A$, because $x_0 \in A$ and for every $x \in A$, $x = \pi(f(x))$.

(3) $\Longrightarrow$ (1). If A is finite then (1) is automatically true, so assume that A is infinite but it has an enumeration $\pi : \mathbb{N} \twoheadrightarrow A$. We must find another enumeration $f : \mathbb{N} \twoheadrightarrow A$ which is *without repetitions*, so that it is in fact a bijection of $\mathbb{N}$ with A, and hence $A =_c \mathbb{N}$. The proof is suggested by Figure **2.1**: we simply delete the repetitions from the given enumeration π of A. To get a precise definition of f by recursion, notice that because A is not finite, for every finite sequence $a_0, \dots, a_n$ of members of A there exists some m such that $\pi(m) \notin \{a_0, \dots, a_n\}$. Set

$$\begin{aligned} f(0) &= \pi(0), \\ m_n &= \text{the least } m \text{ such that } \pi(m) \notin \{f(0), \dots, f(n)\}, \\ f(n+1) &= \pi(m_n). \end{aligned}$$

It is obvious that f is an injection, so it is enough to verify that every $x \in A$ is a value of f, i.e., that *for every* $n \in \mathbb{N}$, $\pi(n) \in f[\mathbb{N}]$. This is immediate for 0, since $\pi(0) = f(0)$. If $x = \pi(n+1)$ for some n and $x \in \{f(0), \dots, f(n)\}$, then $x = f(i)$ for some $i \leq n$; and if $x \notin \{f(0), \dots, f(n)\}$, then $m_n = n+1$ and $f(n+1) = \pi(m_n) = x$ by the definition. ⊣

2.8. Exercise. *If A is countable and there exists an injection $f : B \rightarrowtail A$, then B is also countable*; *in particular, every subset of a countable set is countable.*

2.9. Exercise. *If A is countable and there exists a surjection $f : A \twoheadrightarrow B$, then B is also countable.*

The next, simple theorem is one of the most basic results of set theory.

2.10. Theorem (Cantor). *For each sequence $A_0, A_1, \dots$ of countable sets, the union*

$$A = \bigcup_{n=0}^{\infty} A_n = A_0 \cup A_1 \cup \dots$$

is also a countable set.

In particular, the union $A \cup B$ of two countable sets is countable.

PROOF. The second claim follows by applying the first to the sequence

$$A, B, B, \cdots$$

For the first, it is enough (why?) to consider the special case where none of the A_n is empty, in which case we can find for each A_n an enumeration $\pi^n : \mathbb{N} \twoheadrightarrow A_n$. If we let

$$a_i^n = \pi^n(i)$$

to simplify the notation, then for each n

$$A_n = \{a_0^n, a_1^n, \dots\},$$

and we can construct from these enumerations a table of elements which lists all the members of the union A. This is pictured in Figure **2.2**, and the arrows

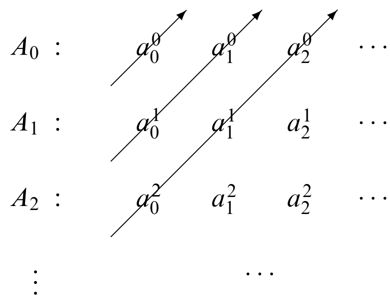

FIGURE **2.2**. Cantor's first diagonal method.

in that picture show how to enumerate the union:

$$A = \{a_0^0, a_0^1, a_1^0, a_0^2, a_1^1, \dots\}. \qquad \dashv$$

2.11. Corollary. *The set of rational* (*positive and negative*) *integers*

$$\mathbb{Z} = \{\dots - 2, -1, 0, 1, 2, \dots\}$$

is countable.

PROOF. $\mathbb{Z} = \mathbb{N} \cup \{-1, -2, \dots, \}$ and the set of negative integers is countable via the correspondence $(x \mapsto -(x + 1))$. $\dashv$

2.12. Corollary. *The set $\mathbb{Q}$ of rational numbers is countable.*

PROOF. The set $\mathbb{Q}^+$ of non-negative rationals is countable because

$$\mathbb{Q}^+ = \bigcup_{n=1}^{\infty} \{\frac{m}{n} \mid m \in \mathbb{N}\}$$

and each $\{\frac{m}{n} \mid m \in \mathbb{N}\}$ is countable via the enumeration $(m \mapsto \frac{m}{n})$. The set $\mathbb{Q}^-$ of negative rationals is countable by the same method, and then the union $\mathbb{Q}^+ \cup \mathbb{Q}^-$ is countable. $\dashv$

This corollary was Cantor's first significant result in the program of classification of infinite sets by their size, and it was considered somewhat "paradoxical" because $\mathbb{Q}$ appears to be so much larger than $\mathbb{N}$. Immediately afterwards, Cantor showed the existence of uncountable sets.

2.13. Theorem (Cantor). *The set of infinite, binary sequences*

$$\Delta = \{(a_0, a_1, \dots,) \mid (\forall i)[a_i = 0 \vee a_i = 1]\}$$

is uncountable.

PROOF. Suppose (towards a contradiction) that Δ is countable, so there exists an enumeration

$$\Delta = \{\alpha_0, \alpha_1, \dots\},$$

where for each n,

$$\alpha_n = (a_0^n, a_1^n, \dots)$$

$$
\begin{tikzcd}
\alpha_0 : & a_0^0 \arrow[dr] & a_1^0 & a_2^0 & \cdots \\
\alpha_1 : & a_0^1 & a_1^1 \arrow[dr] & a_2^1 & \cdots \\
\alpha_2 : & a_0^2 & a_1^2 & a_2^2 \arrow[dr] & \cdots \\
\vdots & & \cdots & & {}
\end{tikzcd}
$$

FIGURE **2.3**. Cantor's second diagonal method.

is a sequence of 0's and 1's.[3] We construct a table with these sequences as before, and then we define the sequence β by interchanging 0 and 1 in the "diagonal" sequence $a_0^0, a_1^1, \ldots$:

$$\beta(n) = 1 - a_n^n.$$

It is obvious that for each α_n, $\beta \neq \alpha_n$, since

$$\beta(n) = 1 - \alpha_n(n) \neq \alpha_n(n),$$

so that the sequence $\alpha_0, \alpha_1, \ldots$ does not enumerate the entire Δ, contrary to our hypothesis. $\dashv$

2.14. Corollary (Cantor). *The set $\mathbb{R}$ of real numbers is uncountable.*

PROOF. We define first a sequence of sets $\mathcal{C}_0, \mathcal{C}_1, \ldots$, of real numbers which satisfy the following conditions:

1. $\mathcal{C}_0 = [0, 1]$.
2. Each $\mathcal{C}_n$ is a union of 2^n closed intervals and
$$\mathcal{C}_0 \supseteq \mathcal{C}_1 \supseteq \cdots \mathcal{C}_n \supseteq \mathcal{C}_{n+1} \supseteq \cdots .$$
3. $\mathcal{C}_{n+1}$ is constructed by removing the (open) middle third of each interval in $\mathcal{C}_n$, i.e., by replacing each $[a, b]$ in $\mathcal{C}_n$ by the two closed intervals
$$L[a, b] = [a, a + \frac{1}{3}(b - a)],$$
$$R[a, b] = [a + \frac{2}{3}(b - a), b].$$

With each binary sequence $\delta \in \Delta$ we associate now a sequence of closed intervals,

$$F_0^\delta, F_1^\delta, \ldots ,$$

[3]To prove a proposition θ by the method of *reduction to a contradiction*, we assume its negation $\neg\theta$ and derive from that assumption something which violates known facts, a contradiction, something *absurd:* we conclude that θ cannot be false, so it must be true. Typically we will begin such arguments with the code-phrase *towards a contradiction*, which alerts the reader that the supposition which follows is the negation of what we intend to prove.

FIGURE **2.4**. The first four stages of the Cantor set construction.

by the following recursion:

$$F_0^\delta = \mathcal{C}_0 = [0, 1],$$
$$F_{n+1}^\delta = \begin{cases} LF_n^\delta, & \text{if } \delta(n) = 0, \\ RF_n^\delta, & \text{if } \delta(n) = 1. \end{cases}$$

By induction, for each n, F_n^δ is one of the closed intervals of $\mathcal{C}_n$ of length 3^{-n} and obviously

$$F_0^\delta \supseteq F_1^\delta \supseteq \cdots,$$

so by the fundamental **completeness property** of the real numbers the intersection of this sequence is not empty; in fact, it contains exactly one real number, call it

$$f(\delta) = \text{ the unique element in the intersection } \bigcap_{n=0}^\infty F_n^\delta.$$

The function f maps the uncountable set Δ into the set

$$\mathcal{C} = \bigcap_{n=0}^\infty \mathcal{C}_n,$$

the so-called **Cantor set**, so to complete the proof it is enough to verify that f is one-to-one. But if n is the least number for which $\delta(n) \neq \varepsilon(n)$ and (for example) $\delta(n) = 0$, we have $F_n^\delta = F_n^\varepsilon$ from the choice of n,

$$f(\delta) \in F_{n+1}^\delta = LF_n^\delta,\ f(\varepsilon) \in F_{n+1}^\varepsilon = RF_n^\delta, \text{ and } LF_n^\delta \cap RF_n^\delta = \emptyset,$$

so that indeed f is an injection. $\dashv$

The basic mathematical ingredient of this proof is the appeal to the completeness property of the real numbers, which we will study carefully in Appendix **A**. Some use of a special property of the reals is necessary: the rest of Cantor's construction relies solely on arithmetical properties of numbers which are also true of the rationals, so if we could avoid using completeness we would also prove that $\mathbb{Q}$ is uncountable, contradicting Corollary **2.12**.

The fundamental importance of this theorem was instantly apparent, the more so because Cantor used it immediately in a significant application to the theory of algebraic numbers. Before we prove this corollary we need some definitions and lemmas.

2.15. Definition. For any two sets A, B, the set of **ordered pairs** of members of A and members of B is denoted by

$$A \times B = \{(x, y) \mid x \in A \,\&\, y \in B\}.$$

In the same way, for each $n \geq 2$,

$$A_1 \times \cdots \times A_n = \{(x_1, \dots, x_n) \mid x_1 \in A_1, \dots, x_n \in A_n\},$$
$$A^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in A\}.$$

We call $A_1 \times \cdots \times A_n$ the **Cartesian product** of $A_1, \dots, A_n$.

2.16. Lemma. (1) *If $A_1, \dots, A_n$ are all countable, so is their Cartesian product $A_1 \times \cdots \times A_n$.*

(2) *For every countable set A, each A^n $(n \geq 2)$ and the union*

$$\bigcup_{n=2}^{\infty} A^n = \{(x_1, \dots, x_n) \mid n \geq 2, x_1, \dots, x_n \in A\}$$

are all countable.

PROOF. (1) If some A_i is empty, then the product is empty (by the definition) and hence countable. Otherwise, in the case of two sets A, B, we have some enumeration

$$B = \{b_0, b_1, \dots\}$$

of B, obviously

$$A \times B = \bigcup_{n=0}^{\infty} (A \times \{b_n\}),$$

and each $A \times \{b_n\}$ is equinumerous with A (and hence countable) via the correspondence $(x \mapsto (x, b_n))$. This gives the result for $n = 2$. To prove the proposition for all $n \geq 2$, notice that

$$A_1 \times \cdots \times A_n \times A_{n+1} =_c (A_1 \times \cdots \times A_n) \times A_{n+1}$$

via the bijection

$$f(a_1, \dots, a_n, a_{n+1}) = ((a_1, \dots, a_n), a_{n+1}).$$

Thus, if every product of $n \geq 2$ countable factors is countable, so is every product of $n + 1$ countable factors, and so (1) follows by induction.

(2) Each A^n is countable by (1), and then $\bigcup_{n=2}^{\infty} A^n$ is also countable by another appeal to Theorem **2.10**. $\dashv$

2.17. Definition. A real number α is **algebraic** if it is a root of some polynomial

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n$$

with integer coefficients $a_0, \dots, a_n \in Z$ $(n \geq 1, a_n \neq 0)$, i.e., if

$$P(\alpha) = 0.$$

Typical examples of algebraic numbers are $\sqrt{2}$, $(1 + \sqrt{2})^2$ (why?) but also the real root of the equation $x^5 + x + 1 = 0$ which exists (why?) but cannot be expressed in terms of radicals, by a classical theorem of Abel. The basic fact (from algebra) about algebraic numbers is that *a polynomial of degree $n \geq 1$ has at most n real roots*; this is all we need for the next result.

2.18. Corollary. *The set K of algebraic real numbers is countable* (Cantor), *and hence there exist real numbers which are not algebraic* (Liouville).

PROOF. The set Π of all polynomials with integer coefficients is countable, because each such polynomial is determined by the sequence of its coefficients, so that Π can be injected into the countable set $\bigcup_{n=2}^{\infty} Z^n$. For each polynomial $P(x)$, the set of its roots

$$\Lambda(P(x)) = \{\alpha \mid P(\alpha) = 0\}$$

is finite and hence countable. It follows that the set of algebraic numbers K is the union of a sequence of countable sets and hence it is countable. $\dashv$

This first application of the (then) new theory of sets was instrumental in ensuring its quick and favorable acceptance by the mathematicians of the period, particularly since the earlier proof of Liouville (that there exist non-algebraic numbers) was quite intricate. Cantor showed something stronger, that "almost all" real numbers are not algebraic, and he did it with a much simpler proof which used just the fact that a polynomial of degree n cannot have more than n real roots, the completeness of $\mathbb{R}$, and, of course, the new method of *counting the members of infinite sets*.

So far we have shown the existence of only two "orders of infinity", that of $\mathbb{N}$—the countable, infinite sets—and that of $\mathbb{R}$. There are many others.

2.19. Definition. The **powerset** $\mathcal{P}(A)$ of a set A is the set of all its subsets,

$$\mathcal{P}(A) = \{X \mid X \text{ is a set and } X \subseteq A\}.$$

2.20. Exercise. *For all sets A, B,*

$$A =_c B \Longrightarrow \mathcal{P}(A) =_c \mathcal{P}(B).$$

2.21. Theorem (Cantor). *For every set A,*

$$A <_c \mathcal{P}(A),$$

i.e., $A \leq_c \mathcal{P}(A)$ but $A \neq_c \mathcal{P}(A)$; in fact there is no surjection $\pi : A \twoheadrightarrow \mathcal{P}(A)$.

PROOF. That $A \leq_c \mathcal{P}(A)$ follows from the fact that the function

$$(x \mapsto \{x\})$$

which associates with each member x of A its **singleton** $\{x\}$ is an injection. (Careful here: the singleton $\{x\}$ is a set with just the one member x and it is not the same object as x, which is probably not a set to begin with!)

To complete the proof, we assume (towards a contradiction) that there exists a surjection

$$\pi : A \twoheadrightarrow \mathcal{P}(A),$$

and we define the set

$$B = \{x \in A \mid x \notin \pi(x)\},$$

so that for every $x \in A$,

$$x \in B \iff x \notin \pi(x). \tag{2-1}$$

Now B is a subset of A and π is a surjection, so there must exist some $b \in A$ such that $B = \pi(b)$; and setting $x = b$ and $\pi(b) = B$ in (2-1), we get

$$b \in B \iff b \notin B$$

which is absurd. $\dashv$

So there are many orders of infinity, and specifically (at least) those of the sets

$$\mathbb{N} <_c \mathcal{P}(\mathbb{N}) <_c \mathcal{P}(\mathcal{P}(\mathbb{N})) <_c \cdots .$$

If we name these sets by the recursion

$$\begin{aligned} T_0 &= \mathbb{N}, \\ T_{n+1} &= \mathcal{P}(T_n), \end{aligned} \tag{2-2}$$

then their union $T_\infty = \bigcup_{n=0}^{\infty} T_n$ has a larger cardinality than each T_n, Problem **x2.8**. The classification and study of these orders of infinity is one of the central problems of set theory.

Somewhat more general than powersets are *function spaces*.

2.22. Definition. For any two sets A, B,

$$\begin{aligned} (A \to B) =_{\text{df}} & \{f \mid f : A \to B\} \\ = & \quad \text{the set of all functions from } A \text{ to } B. \end{aligned}$$

2.23. Exercise. *If* $A_1 =_c A_2$ *and* $B_1 =_c B_2$, *then* $(A_1 \to B_1) =_c (A_2 \to B_2)$.

Function spaces are "generalizations" of powersets because each subset $X \subseteq A$ can be represented by its *characteristic function* $c_X : A \to \{0, 1\}$,

$$c_X(t) = \begin{cases} 1, & \text{if } t \in A \cap X, \\ 0, & \text{if } t \in A \setminus X, \end{cases} \qquad (t \in A). \tag{2-3}$$

We can recover X from c_X,

$$X = \{t \in A \mid c_X(t) = 1\},$$

and so the mapping $(X \mapsto c_X)$ is a correspondence of $\mathcal{P}(A)$ with $(A \to \{0, 1\})$. Thus

$$(A \to \{0, 1\}) =_c \mathcal{P}(A) >_c A, \tag{2-4}$$

and the function space operation also leads to large, uncountable sets. The next obvious problem is to compare for size these uncountable sets, starting with the two simplest ones, $\mathcal{P}(\mathbb{N})$ and the set $\mathbb{R}$ of real numbers.

2.24. Lemma. $\mathcal{P}(\mathbb{N}) \leq_c \mathbb{R}$.

PROOF. It is enough to prove that $\mathcal{P}(\mathbb{N}) \leq_c \Delta$, since we have already shown that $\Delta \leq_c \mathbb{R}$. This follows immediately from (2-4), as $\Delta = (\mathbb{N} \to \{0, 1\})$. $\dashv$

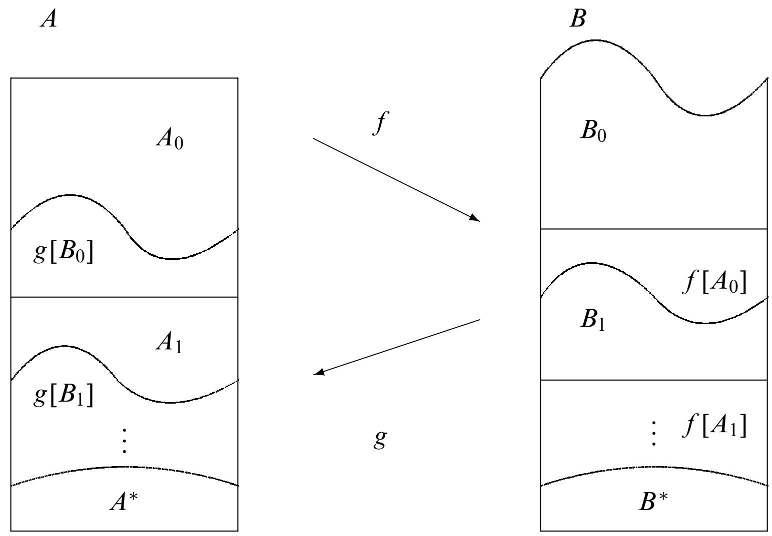

FIGURE **2.5**. Proof of the Schröder-Bernstein Theorem.

2.25. Lemma. $\mathbb{R} \leq_c \mathcal{P}(\mathbb{N})$.

PROOF. It is enough to show that $\mathbb{R} \leq_c \mathcal{P}(\mathbb{Q})$, since the set of rationals $\mathbb{Q}$ is equinumerous with $\mathbb{N}$ and hence $\mathcal{P}(\mathbb{N}) =_c \mathcal{P}(\mathbb{Q})$. This follows from the fact that the function

$$x \mapsto \pi(x) = \{q \in \mathbb{Q} \mid q < x\} \subseteq \mathbb{Q}$$

is an injection, because if $x < y$ are distinct real numbers, then there exists some rational q between them, $x < q < y$ and $q \in \pi(y) \setminus \pi(x)$. $\dashv$

With these two simple Lemmas, the equinumerosity $\mathbb{R} =_c \mathcal{P}(\mathbb{N})$ will follow immediately from the following basic theorem.

2.26. Theorem (Schröder-Bernstein). *For any two sets A, B,*

$$\textit{if } A \leq_c B \textit{ and } B \leq_c A, \textit{ then } A =_c B.$$

PROOF.[4] We assume that there exist injections

$$f : A \rightarrowtail B, \; g : B \rightarrowtail A,$$

and we define the sets A_n, B_n by the following recursive definitions:

$$\begin{aligned} A_0 &= A, & B_0 &= B, \\ A_{n+1} &= gf[A_n], & B_{n+1} &= fg[B_n], \end{aligned}$$

[4]A different proof of this theorem is outlined in Problems **x4.26**, **x4.27**.

where $fg[X] = \{f(g(x)) \mid x \in X\}$ and correspondingly for the function gf. By induction on n (easily)

$$\begin{aligned} A_n &\supseteq g[B_n] \supseteq A_{n+1}, \\ B_n &\supseteq f[A_n] \supseteq B_{n+1}, \end{aligned}$$

so that we have the "chains of inclusions"

$$\begin{aligned} A_0 &\supseteq g[B_0] \supseteq A_1 \supseteq g[B_1] \supseteq A_2 \cdots, \\ B_0 &\supseteq f[A_0] \supseteq B_1 \supseteq f[A_1] \supseteq B_2 \cdots. \end{aligned}$$

We also define the intersections

$$A^* = \bigcap_{n=0}^{\infty} A_n, \; B^* = \bigcap_{n=0}^{\infty} B_n,$$

so that

$$B^* = \bigcap_{n=0}^{\infty} B_n \supseteq \bigcap_{n=0}^{\infty} f[A_n] \supseteq \bigcap_{n=0}^{\infty} B_{n+1} = B^*$$

and since f is an injection, by Problem **x1.7**,

$$f[A^*] = f[\bigcap_{n=0}^{\infty} A_n] = \bigcap_{n=0}^{\infty} f[A_n] = B^*.$$

Thus f is a bijection of A^* with B^*. On the other hand,

$$\begin{aligned} A &= A^* \cup (A_0 \setminus g[B_0]) \cup (g[B_0] \setminus A_1) \cup (A_1 \setminus g[B_1]) \cup (g[B_1] \setminus A_2) \ldots \\ B &= B^* \cup (B_0 \setminus f[A_0]) \cup (f[A_0] \setminus B_1) \cup (B_1 \setminus f[A_1]) \cup (f[A_1] \setminus B_2) \ldots \end{aligned}$$

and these sequences are separated, i.e., no set in them has any common element with any other. To finish the proof it is enough to check that for every n,

$$\begin{aligned} f[A_n \setminus g[B_n]] &= f[A_n] \setminus B_{n+1}, \\ g[B_n \setminus f[A_n]] &= g[B_n] \setminus A_{n+1}, \end{aligned}$$

from which the first (for example) is true because f is an injection and so

$$f[A_n \setminus g[B_n]] = f[A_n] \setminus fg[B_n] = f[A_n] \setminus B_{n+1}.$$

Finally we have the bijection $\pi : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$,

$$\pi(x) = \begin{cases} f(x), & \text{if } x \in A^* \text{ or } (\exists n)[x \in A_n \setminus g[B_n]], \\ g^{-1}(x), & \text{if } x \notin A^* \text{ and } (\exists n)[x \in g[B_n] \setminus A_{n+1}], \end{cases}$$

which verifies that $A =_c B$ and finishes the proof. $\dashv$

Using the Schröder-Bernstein Theorem we can establish easily several equinumerosities which are quite difficult to prove directly.

Problems for Chapter 2

x2.1. For any $\alpha < \beta$ where α, β are reals, ∞ or $-\infty$, construct bijections which prove the equinumerosities

$$(\alpha, \beta) =_c (0, 1) =_c \mathbb{R}.$$

***x2.2.** For any two real numbers $\alpha < \beta$, construct a bijection which proves the equinumerosity

$$[\alpha, \beta) =_c [\alpha, \beta] =_c \mathbb{R}.$$

x2.3. $\mathcal{P}(\mathbb{N}) =_c \mathbb{R} =_c \mathbb{R}^n$, for every $n \geq 2$.

x2.4. For any two sets A, B, $(A \to B) \leq_c \mathcal{P}(A \times B)$. HINT. Represent each $f : A \to B$ by its *graph*, the set

$$G_f = \{(x, y) \in A \times B \mid y = f(x)\}.$$

x2.5. $(\mathbb{N} \to \mathbb{N}) =_c \mathcal{P}(\mathbb{N})$.

***x2.6.** $(\mathbb{N} \to \mathbb{R}) =_c \mathbb{R}$.

***x2.7.** For any three sets A, B, C,

$$((A \times B) \to C) =_c (A \to (B \to C)).$$

HINT. For any $p : A \times B \to C$, define $\pi(p) = q : A \to (B \to C)$ by the formula

$$q(x)(y) = p(x, y).$$

x2.8. Using the definition (2-2), for every m,

$$T_m <_c T_\infty = \bigcup_{n=0}^{\infty} T_n.$$

You need to know something about continuous functions to do the last two problems.

***x2.9.** The set $C[0, 1]$ of all continuous, real functions on the closed interval $[0, 1]$ is equinumerous with $\mathbb{R}$.

***x2.10.** The set of all monotone real functions on the closed interval $[0, 1]$ is equinumerous with $\mathbb{R}$.