Introduction to Modern Set Theory

Judith Roitman

December 6, 2011

[1]

for my father, who loved mathematics

Table of Contents

Preface

When, in early adolescence, I first saw the proof that the real numbers were uncountable, I was hooked. I didn't quite know on what, but I treasured that proof, would run it over in my mind, and was amazed that the rest of the world didn't share my enthusiasm. Much later, learning that set theorists could actually prove some basic mathematical questions to be unanswerable, and that large infinite numbers could effect the structure of the reals — the number line familiar to all of us from the early grades — I was even more astonished that the world did not beat a path to the set theorist's door.

More years later than I care to admit (and, for this revision, yet another twenty years later), this book is my response. I wrote it in the firm belief that set theory is good not just for set theorists, but for many mathematicians, and that the earlier a student sees the particular point of view that we call modern set theory, the better.

It is designed for a one-semester course in set theory at the advanced undergraduate or beginning graduate level. It assumes no knowledge of logic, and no knowledge of set theory beyond the vague familiarity with curly brackets, union and intersection usually expected of an advanced mathematics student. It grew out of my experience teaching this material in a first-year graduate course at the University of Kansas over many years. It is aimed at two audiences — students who are interested in studying set theory for its own sake, and students in other areas who may be curious about applications of set theory to their field. While a one-semester course with no logic as a prerequisite cannot begin to tell either group of students all they need to know, it can hope to lay the foundations for further study. In particular, I am concerned with developing the intuitions that lie behind modern, as well as classical, set theory, and with connecting set theory with the rest of mathematics.

Thus, three features are the full integration into the text of the study of models of set theory, the use of illustrative examples both in the text and and in the exercises, and the integration of consistency results and large cardinals into the text when appropriate, even early on (for example, when cardinal exponentiation is introduced). An attempt is made to give some sense of the history of the subject, both as motivation, and because it is interesting in its own right.

The first chapter is an introduction to partial orders and to well-ordered sets, with a nod to induction on $\mathbb{N}$, filters, and ideals. The second chapter is about first-order theories and their models; this discussion is greatly extended from the first edition. Without becoming too formal, this chapter carefully examines a number of theories and their models, including the theory of partially ordered sets, in order to provide a background for discussion of models of the various axioms of set theory.

The third chapter introduces all of the axioms except regularity and choice, formally defines the natural numbers, and gives examples of models of the axioms, with an emphasis on standard models (in which the symbol "$\in$" is interpreted by the real relation $\in$). The fourth chapter discusses transitive sets, regularity, ordinals, and choice and, in its last section, gives a taste of how to embed standard mathematics within set theory.

Cardinal and ordinal numbers are the subject of chapter five. Chapter six discusses V, V_κ where κ is inaccessible, and L. Chapter seven introduces infinite combinatorics: partition calculus, trees, measurable cardinals, CH. Martin's axiom, stationary sets and $\diamondsuit$, and cardinal invariants of the reals.

A brief word about formality. The first chapter is written in ordinary mathematical style

without set-theoretical formality (compare the definition of partial order in section 2.1 with the formal definition in section 3.3). The reader is assumed to be familiar with set-theoretic notation as found in most advanced mathematical texts, and we will make use of it throughout. The reader is also assumed to be familiar with the standard body of basic mathematics, e.g., the basic properties of the natural numbers, the integers, the rationals, and the reals.

1 Preliminaries

The reader is probably used to picking up a mathematical textbook and seeing a first chapter which includes a section with an approximate title of "Set-theoretical prerequisites." Such a chapter usually contains a quick review or an overview of the relevant set theory, from something as simple as the definition of the union of two sets to something as complicated as the definitions of countable and uncountable sets. Since this is a set theory text, we reverse the usual procedure by putting in the first chapter some mathematics that will prove essential to the serious study of set theory: partially ordered and linearly ordered sets, equivalence relations, well-ordered sets, induction and recursion, filters and ideals.. Why these topics?

The spine of the set-theoretic universe, and the most essential class of objects in the study of set theory, is the class of ordinals. One of the basic properties of an ordinal is that it is a well-ordered set. An acquaintance with various examples and properties of well-ordered sets is essential to the study of ordinals.

Two of the basic techniques of set theory are transfinite induction and transfinite recursion, which are grounded in induction and recursion on the natural nubmers.

When set theory is applied to the rest of mathematics, the methodology often used is to reduce the original question to a question in the area known as infinite combinatorics. The combinatorics of partially ordered sets (especially those known as trees, see chapter 7) are particularly important, and partially ordered sets are crucial to the major technique of proving consistency results.

Finally, filters and ideals are not only important combinatorial objects, but are essential to the theory of large cardinals.

Thus the choice of topics in this chapter.

1.1 Partially ordered sets

Partially ordered sets underlie much of the combinatorics we will use.

Definition 1.1. $\leq$ is a partial order on a set X iff for all $x, y, z \in X$

P1 (Reflexive axiom) $x \leq x$.

P2 (Antisymmetric axiom) If $x \leq y$ and $y \leq x$ then $x = y$.

P3 (Transitive axiom) If $x \leq y$ and $y \leq z$ then $x \leq z$.

As shorthand, we say $x < y$ (x is strictly less than y) if $x \leq y$ and $x \neq y$. If $\leq$ partiallly orders X, we call X a partially ordered set under $\leq$ and say $<$ strictly orders X. As will become clear from the examples, a set can have many different partial orders imposed upon it.

Here are some examples.

Example 1.2. Let $\mathbb{N}$ be the set of natural numbers $\mathbb{N} = \{0, 1, 2, 3...\}$. For $n, k \in \mathbb{N}$ we define $n \leq_D k$ iff n divides k.

Is this a partial order?

Check for P1: Every n divides n, so each $n \leq_D n$.

Check for P2: If n divides k then $n \leq k$ (where $\leq$ is the usual order). If k divides n, then $k \leq n$. We know that if $n \leq k$ and $k \leq n$ then $n = k$. Hence if n divides k and k divides n, then $n = k$.

Check for P3: $n \leq_D k$ iff $k = in$ for some i. $k \leq_D m$ if $m = jk$ for some j. So if $n \leq_D k \leq_D m$ there are i, j with $m = jk = jin$. Hence $n \leq_D m$.

Notice that the k, j whose existence was needed for the proof of P3 must come from $\mathbb{N}$. The fact that $2 = \frac{2}{3}(3)$ does not imply that $2 \leq_D 3$. This will be important when we discuss models.

Example 1.3. Let X be any set and, for $x, y \in X$, define $x \leq_E y$ iff $x = y$ for all $x, y \in X$.

Is $\leq_E$ a partial order?

Check for P1: Since each $x = x$, each $x \leq_E x$.

Check for P2: If $x \leq_E y \leq_E x$ then $x = y$ and (redundantly) $y = x$.

Check for P3: If $x \leq_E y \leq_E z$ then $x = y$ and $y = z$, so $x = z$. Hence $x \leq_E z$.

Example 1.3 is instructive. Partial orders may have very little structure. In this case, there is no structure.

Example 1.4. Let X be any collection of sets, and for all $x, y \in X$ define $x \leq_S y$ iff $x \subseteq y$.[2]

The proof that example 4 is a partial order is left to the reader.

Example 1.5. Consider the set of real numbers $\mathbb{R}$. For $x, y \in \mathbb{R}$ we define $x \preceq y$ iff $\exists z\ z^2 = y - x$.

In fact, in example 1.5, $x \preceq y$ iff $x \leq y$ in the usual sense. We will come back to this when we discuss models of first order theories.

Example 1.6. Let X be the set of (possibly empty) finite sequences of 0's and 1's. For $\sigma, \tau \in X$ we define $\sigma \leq_e \tau$ iff τ extends σ or $\tau = \sigma$.

For clarity, we enclose each sequence in parentheses. I.e., the number 0 is not the same as the sequence (0).

To flesh out example 1.6 a bit, $(0) \leq_e (01) \leq_e (011) \leq_e (0110) \leq_e (01101)...$ We define $l(\sigma)$ to be the length of σ, i.e., $l(011) = 3$. We define $\sigma^\frown\tau$ (read: σ followed by τ) to mean, not surprisingly, σ followed by τ. I.e., $(01)^\frown(10) = (0110)$. Note that $\sigma \leq_e \tau$ iff $\exists \rho$ with $\tau = \sigma^\frown\rho$.[3]

Is $\leq_e$ a partial order?

Check for P1: Immediate from the definition.

Check for P2: If $\sigma \leq_e \tau$ then $l(\sigma) \leq l(\tau)$. So if $\sigma \leq_e \tau \leq_e \sigma$ then $l(\sigma) = l(\tau)$. Hence, since there is no $\rho \neq \emptyset$ with $\tau = \sigma^\frown\rho$, $\sigma = \tau$.

[2]Recall that $x \subseteq y$ iff, for all z, if $z \in x$ then $z \in y$. In particular, $x \subseteq x$.

[3]Recall that ρ could be empty.

Check for P3: Suppose $\sigma \leq_e \tau \leq_e \rho$. If $\sigma = \tau$ or $\tau = \rho$ then $\sigma \leq_e \rho$. Otherwise, there are ν, η with $\sigma^\frown\nu = \tau$ and $\tau^\frown\eta = \rho$. So $\rho = \sigma^\frown(\nu^\frown\eta)$. So $\sigma \leq_e \rho$.

Example 1.6 is one of a class of partial orders that is very useful in set theoretic combinatorics and independence proofs.

We generalize example 1.6 as follows:

Example 1.7. Let S be a set, and let X be the set of all finite sequences (including the empty sequence) whose elements are in S. For $\sigma, \tau \in X$, $\sigma \leq_e \tau$ iff there is some $\rho \in X$ with $\tau = \sigma^\frown\rho$.

That $\leq_e$ is also a partial order is left to the reader.

We now use this partial order to define another one:

Example 1.8. Let X be a partially ordered set under $\leq$. Let $FIN(X)$ be the set of finite sequences (including the empty sequence) whose members are elements of X. If $\sigma = (x_1, x_2, x_3..., x_n), \tau = (y_1, y_2, y_3, ...y_m) \in X$ we say $\sigma \leq_L \tau$ iff, either $\sigma \leq_e \tau$ or, where j is least so $x_j \neq y_j$ then $x_j \leq y_j$.

For example, if $X = \{0, 1\}$ where $0 \leq 1$ then $(00) \leq_L (1) \leq_L (10) \leq_L (100) \leq_L (11)$.

For another example, if $X = \mathbb{R}$, then $(e, \pi) \leq_L (\pi, e, \sqrt{2}) \leq_L (\pi, \sqrt{11})$.

The order in example 1.8 is called the lexicographic order, and it is also useful combinatorially. The proof that it is a partial order is an exercise.

Example 1.9. Let $X = \mathbb{N}^\mathbb{N}$, i.e., the set of all functions from $\mathbb{N}$ to $\mathbb{N}$. Define $f \leq_{ae} g$ iff $\{n : f(n) > g(n)\}$ is finite.[4]

Let's check P2: Suppose $f(3) \neq g(3)$, $f(n) = g(n)$ for all $n \neq 3$. Then $f \leq_{ae} g \leq_{ae} f$ but $f \neq g$. So P2 fails and $\leq_{ae}$ is not a partial order.

$\leq_{ae}$ does satisfy P1 and P3 (another exercise). Relations satisfying P1 and P3 are called pre-orders.

1.2 The structure of partially ordered sets

Definition 1.10. Let $\leq$ be a partial order on a set X. Two elements x, y of X are comparable iff $x \leq y$ or $y \leq x$; otherwise they are incomparable. They are compatible iff there is some $z \in X$ so $z \leq x$ and $z \leq y$; otherwise they are incompatible.

Note that if two elements are comparable then they are compatible, but not vice versa. In example 1.2, $3 \leq_D 6$ and $3 \leq_D 9$, so 6, 9 are compatible, but $6 \not\leq_D 9$ and $9 \not\leq_D 6$, so they are incomparable.

Definition 1.11. A subset B of a partially ordered set is a chain iff any two elements of B are comparable. B is an antichain iff no two elements of B are compatible. B is linked iff it is pairwise compatible.[5]

[4] "ae" is short for "almost everywhere." As we will see in chapter 7, it is a very important relation, usually denoted as $\leq^*$.

[5] Apologies for the lack of parallelism in the definitions of chain and antichain. This terminology has become standard and nothing can shake it.

Definition 1.12. A partial order $\leq$ on a set X is linear iff X is a chain.

If $\leq$ is a linear order on X, we often say that X is linearly ordered, or linear, and suppress $\leq$.

Theorem 1.13. *X is linear iff the lexicographic order on $FIN(X)$ is linear.*

Proof. Suppose X is not linear. Let x, y be incomparable elements of X. Then $(x), (y)$ are incomparable elements of $FIN(X)$.

Now suppose X is linear under the order $\leq$. Let $\sigma, \tau \in FIN(X)$. If $\sigma \leq_e \tau$ then $\sigma \leq_L \tau$, so we may assume there is some k with $\sigma(k) \neq \tau(k)$, hence a least such k, call it k^*. If $\sigma(k^*) < \tau(k^*)$ then $\sigma \leq_L \tau$. If $\sigma(k^*) > \tau(k^*)$ then $\sigma >_L \tau$. So σ, τ are comparable. □

Definition 1.14. Let X be a partial order, $x \in X$. We define

(a) x is minimal iff, for all $y \in X$, if $y \leq x$ then $y = x$.

(b) x is maximal iff, for all $y \in X$, if $y \geq x$ then $y = x$.

(c) x is a minimum iff, for all $y \in X, y \geq x$.

(d) x is a maximum iff, for all $y \in X, y \leq x$.

In example 1.3 every element is both maximal and minimal. In example 1.4, if $X = \{Y : Y \subseteq \mathbb{N}\}$ then $\mathbb{N}$ is a maximum and $\emptyset$ is a minimum.

Here are some basic facts:

Proposition 1.15. *(a) Maximum elements are maximal.*

(b) Minimum elements are minimal.

(c) If X has a minimum element then every subset is linked.

(d) There can be at most one maximum element.

(e) There can be at most one minimum element.

(f) A maximal element in a linear order is a maximum.

(g) A minimal element in a linear order is a minimum.

The proof of proposition 1.15 is left as an exercise.

Here's an example to show that a unique maximal element need not be a maximum, and a unique minimal element need not be a minimum.

In the above illustration, $X = (0, 1] \cup [2, 3)$. We define $x \preceq y$ iff $0 < x \leq y \leq 1$ or $2 \leq x \leq y < 3$ where $\leq$ is the usual order in $\mathbb{R}$. Then 1 is a unique maximal element and 2 is a unique minimal element but 1, 2 are not comparable. Hence 1 is not a maximum, and 2 is not a minimum.

Definition 1.16. Given X partially ordered by $\leq$ and $S \subseteq X$, we say that S is unbounded iff there is no $x \in X$ with $x \geq s$ for all $s \in S$.[6]

Definition 1.17. Given X partially ordered by $\leq$ and $S \subseteq X$, we say that S is dominating iff for all $x \in X$ there is $s \in S$ with $x \leq s$.[7]

For example, in the order $\prec$ on $X = (0,1] \cup [2,3)$, $\{2 + \frac{n}{n+1} : n \in \mathbb{N}\}$ is unbounded, but not dominating. $\{1\} \cup \{2 + \frac{n}{n+1} : n \in \mathbb{N}\}$ is dominating and unbounded.

Note that a dominating family without a maximum element is necessarily unbounded. A dominating family with a maximum element is necessarily bounded.

Dominating families in linear orders are also called cofinal. For example, $\mathbb{N}$ is cofinal in both $\mathbb{R}$ and $\mathbb{Q}$.

For another example, $\{(n) : n \in \mathbb{N}\}$ is cofinal in $FIN(\mathbb{N})$ under the lexicographic order.

When we look at a partial order, we often want to know what its chains, antichains, linked, unbounded and cofinal sets look like and what its minimal, maximal, minimum, and maximum elements (if any) are. We also ask general questions about how these things are related. For example, the Suslin Hypothesis (which we'll meet later) is a statement about the size of antichains in a certain kind of partial order, and an Aronszajn tree (which we'll also meet later) is an example showing that a certain kind of partial order can have large antichains but no large chains.

1.3 Equivalence relations

An equivalence relation is almost like a partial order, with one key exception: instead of antisymmetry, we have symmetry. The axioms for an equivalence relation $\equiv$ on a set X are:

For all $x, y, z \in X$,

E1 (Reflexive axiom) $x \equiv x$

E2 (Symmetric axiom) $x \equiv y$ iff $y \equiv x$

E3 (Transitive axiom) If $x \equiv y$ and $y \equiv z$ then $x \equiv z$.

Example 1.18. Example 1.3 is an equivalence relation.

Example 1.19. If X is a pre-order under $\leq$, we define $x \equiv_{\leq} y$ iff $x \leq y \leq x$.

We show that $\equiv_{\leq}$ is an equivalence relation:

E1: $x \leq x \leq x$

E2: $x \equiv_{\leq} y$ iff $x \leq y \leq x$ iff $y \leq x \leq y$ iff $y \equiv_{\leq} x$

E3: If $x \leq y \leq x$ and $y \leq z \leq y$ then, since $\leq$ is transitive, $x \leq z \leq x$.

An important example of an equivalence relation as in example 1.19 is the relation $\equiv_{\leq_{ae}}$ on $\mathbb{N}^{\mathbb{N}}$.In this example, $f \equiv_{\leq_{ae}} g$ iff $\{n : f(n) \neq g(n)\}$ is finite.

[6] A pre-order suffices.

[7] See the previous footnote.

Definition 1.20. (a) If $\equiv$ is an equivalence relation on X and $x \in X$ then $[x]_\equiv = \{y : y \equiv x\}$ is called the equivalence class of x. We often just write $[x]$.

(b) $X/_\equiv = \{[x]_\equiv : x \in X\}$.

Example 1.21. Suppose $\preceq$ is a pre-order on X For $[x], [y] \in X/_{\equiv_\preceq}$ we write $[x] \leq_\preceq [y]$ iff $x \preceq y$.

We show that example 1.21 is well-defined, i.e., that it does not depend on the choice of representative for the equivalence class: Suppose $x' \equiv_\preceq x, y' \equiv_\preceq y$. If $x \preceq y$ then $x' \preceq x \preceq y \preceq y'$, so $x' \preceq y'$.

We show that example 1.21 is a partial order: P1 and P2 are left to the reader. For P3: $[x] \leq_\preceq [y] \leq_\preceq [z]$ iff $x \preceq y \preceq z$. Hence if $[x] \leq_\preceq [y] \leq_\preceq [z]$, $x \preceq z$, so $[x] \leq_\preceq [z]$.

The structure of the partial order $\mathbb{N}^\mathbb{N}/\equiv_{\leq_{ae}}$ under $\leq_{\leq_{ae}}$ (known as the Fréchet order), especially its cofinal and unbounded sets, has consequences for many areas of mathematics. We will discuss this in chapter 7.

There is a close relation between equivalence relations and partitions, as follows.

Definition 1.22. A partition $\mathfrak{P}$ of a set X is a collection of subsets of X so that every element of X belongs to exactly one element of $\mathfrak{P}$.

For example, for $r \in \mathbb{R}$ let $P_r = \{(r, y) : y \in \mathbb{R}\}$. Then $\mathfrak{P} = \{P_r : r \in \mathbb{R}\}$ is a partition of $\mathbb{R}^2$.

For another example, for $n \in \mathbb{N}, n > 1$, let $D_n = \{m : n \text{ is the least divisor of } m \text{ which is greater than } 1\}$. For $n = 0$, define $D_n = \{0\}$; for $n = 1$, define $D_n = \{1\}$. Note that, if $n > 1$, $D_n \neq \emptyset$ iff n is prime. Let $\mathcal{D} = \{D_n : n \in \mathbb{N}\}$. Then $\mathcal{D}$ is a partition of $\mathbb{N}$.

Definition 1.23. If $\mathfrak{P}$ is a partition of X, for $x, y \in X$ we define $x \equiv_\mathfrak{P} y$ iff $\exists P \in \mathfrak{P}\ x, y \in P$.

Definition 1.24. If $\equiv$ is an equivalence relation on X, we define $\mathfrak{P}_\equiv = \{[x]_\equiv : x \in X\}$.

Proposition 1.25. *Assume $\mathfrak{P}$ is a partition of X and $\equiv$ is an equivalence relation on X.*

(a) $\equiv_\mathfrak{P}$ is an equivalence relation on X.

(b) $\mathfrak{P}_\equiv$ is a partition of X.

(c) $\equiv_{\mathfrak{P}_\equiv} = \equiv$.

(d) $\mathfrak{P}_{\equiv_\mathfrak{P}} = \mathfrak{P}$.

Proof. We only prove (a) and (d), and leave the rest to the reader.

For (a): E1: if $x \in P \in \mathfrak{P}$ then $x \in P$. E2: If $x, y \in P \in \mathfrak{P}$ then $y, x \in P$. E3: If $x, y \in P \in \mathfrak{P}$ and $y, z \in P' \in \mathfrak{P}$, then $P = P'$ and $x, z \in P$.

For (d) Given x, let P be the unique element of $\mathfrak{P}$ with $x \in P$. Then $y \in P$ iff $y \equiv_\mathfrak{P} x$ iff $y \in [x]_{\equiv_\mathfrak{P}}$. Hence $[x]_{\equiv_\mathfrak{P}} = P$. □

1.4 Well-ordered sets

Well-ordered sets are a particularly useful kind of linear order. They correspond to ordinal numbers, and ordinal numbers form the spine of the set-theoretic universe. By definition, this set-theoretic

universe will be built out of the empty set by the operations of union and power set, but this is in some sense too nihilistic to be useful. That the set-theoretic universe is built from layers arranged in a well-ordered fashion, indexed by ordinals — that will turn out to be very useful. We won't define the notion of *ordinal* until chapter 4. Here we develop the concepts we need to understand it.

Definition 1.26. A well-ordered set X is a linearly ordered set for which every nonepty subset has a minimum.

Every finite linearly ordered set is well-ordered. Since every nonempty set of natural numbers has a minimum, $\mathbb{N}$ is well-ordered.

In fact, in a very precise sense (which we'll see later in this section), all infinite well-ordered sets are built by stacking up copies of $\mathbb{N}$, with perhaps a finite set on top.

Example 1.27. $X = \{\frac{m}{m+1} : m \in \mathbb{N}\} \cup \{1 + \frac{m}{m+1} : m \in \mathbb{N}\}$.

Since example 1.27 is essentially one copy of $\mathbb{N}$ piled on top of another one, it is well-ordered by the usual order on the reals.

Example 1.28. $X = \{n + \frac{m}{m+1} : n, m \in \mathbb{N}\}$.

Let's give a formal proof that example 1.28 is well-ordered:

Suppose $Y \subseteq X$. Because $\mathbb{N}$ is well-ordered, there is a least n so that $\exists m\ n + \frac{m}{m+1} \in Y$. For this n, because $\mathbb{N}$ is well-ordered, there is a least m so that $n + \frac{m}{m+1} \in Y$. $n + \frac{m}{m+1}$ is the least element in Y.

Example 1.29. Let X be a partially ordered set, and, for $n \in \mathbb{N}$, let $L_n(X)$ be the sequences of length $\leq n$ with elements in X, under the lexicographic order.

We prove that if X is well-ordered, so is $L_n(X)$: Let $Y \subseteq L_n(X)$. Let x_1 be the least element in X so that some $(x_1 y_2 ... y_k) \in Y$. If $(x_1) \in Y$, we've found our minimum. Otherwise, there is x_2 the least element in X so that some $(x_1 x_2 y_3 ... y_j) \in Y$. If $(x_1 x_2) \in Y$, we've found our least element. And so on. Since n is finite, the process stops at some point.

Note that $L(X)$, the lexicographic order on all finite sequences from X, is not well-ordered as long as X has at least two elements: suppose $x < y$. Then $(y) > (xy) > (xxy) > (xxxy)...$

Proposition 1.30. *Every subset of a well-ordered set is well-ordered.*

Proof. Suppose $Y \subseteq X, X$ is well-ordered. If $Z \subseteq Y$ then $Z \subseteq X$, so Z has a minimum. Hence every subset of Y has a minimum. □

Another way of stating proposition 1.30 is: the property of being well-ordered is hereditary.

The definition of well-ordered sets involves a universal quantifier — you have to check *every* subset to see if it has a least element. There's a nice existential criterion for not being a well-ordering, nicer than the obvious "some set doesn't have a least element."

Theorem 1.31. *A linear order X is well-ordered iff it has no infinite descending chain.*

Proof. An infinite descending chain is a set $C = \{x_n : n \in \mathbb{N}\}$ where each $x_n > x_{n+1}$. So suppose X has such a chain C. Then C has no minimum and X is not well-ordered. For the other direction, suppose X is not well-ordered. Let $Y \subseteq X$ where Y has no minimum. Pick $x_1 \in Y$. It's not minimum, so there's some $x_2 \in Y$ with $x_2 < x_1$. And so on. In this way we construct $x_1 > x_2 > x_3...$ where each $x_n \in Y$. Thus there is an infinite descending chain of elements of X. □

Theorem 1.31 differs from the theorems we have seen so far in the sophistication of its proof. When we assumed that X was not well-ordered and constructed an infinite descending chain, no rule was given for chosing each x_i. Given x_0 there may be many candidates for x_1. Given x_1 there may be many candidates for x_2. And so on. That we can pick a path through this maze seems reasonable, but in fact we used a weak form of a somewhat controversial axiom, the axiom of choice. This axiom is defined and explored in chapter 4 and is critical to our discussion of size, or cardinality, in chapter 5. Note that only one direction of theorem 1.31 holds without at least a weak version of the axiom of choice. This will be important later, when we work with well-orderings both with and without the axiom of choice.

Theorem 1.31 tells us that, under the usual ordering as subsets of $\mathbb{R}$, the following sets are not well-ordered: the integers $\mathbb{Z}$; any non-empty interval in $\mathbb{R}$; any non-empty interval in $\mathbb{Q}$; $\{-x : x \in X\}$ where x is as in examples 1.27 or 1.28.

The final task of this section is to justify our statement that well-ordered sets are built out of copies of $\mathbb{N}$.

We begin with a definition useful for any partial order.

Definition 1.32. Let X be a linear order, $x \in X$. We say that y is the successor of x iff $y > x$ and if $z > x$ then $z \geq y$.

If y is a successor of x, we write $y = S(x)$.

Definition 1.33. Let X be a linear order, $x \in X$. We say that x is a limit iff $\forall y \in X \; x \neq S(y)$.

No x can have two distinct successors. A linear order might have no successors, e.g., in $\mathbb{R}$ and in $\mathbb{Q}$ every element is a limit. Note that a minimal element is a limit.

We will use the notions of limit and successor to see precisely what well-ordered sets must look like.

Proposition 1.34. *Let X be well-ordered, $x \in X$. Either x is a maximum, or x has a successor.*

Proof. If x is not a maximum, then $S = \{y \in X : y > x\} \neq \emptyset$. The minimum of S is the successor of x. □

Definition 1.35. Let X be a partial order. $S^0(x) = x$; for $n > 0$, if $S^n(x)$ is defined and has a successor, then $S^{n+1}(x) = S(S^n(x))$.

Theorem 1.36. *Let X be a partial order, $x \in X$. If $n < m$ and $S^m(x)$ exists, then $S^n(x) < S^m(x)$.*

Proof. Fix x. Fix n so $S^n(x)$ exists. By definition 1.32, $S^n(x) < S^{n+1}(x) < S^{n+2}(x)...$ for all k with $S^{n+k}(x)$ defined. □

Proposition 1.37. *Let X be well-ordered, and let Λ be the set of limits in X.*

(a) $X = \{y \in X : \exists n\ \exists x \in \Lambda,\ y = S^n(x)\}$.

(b) If $y \in \Lambda$ and $x < y$ then $S^n(x) < y$ for all n.

(c) If X has a maximum element, so does Λ, and the maximum element of X is some $S^n(x)$ where x is the maximum element of Λ.

Proof. (a) Suppose this fails. Let $C = \{y \in X : \forall n\ \forall x \in \Lambda\ y \neq S^n(x)\} \neq \emptyset$. Let y be minimum in C. $y \notin \Lambda$, so there is $z < y$ with $y = S(z)$. Since $z \notin C$ there is $x \in \Lambda$ and n with $z = S^n(x)$. Hence $y = S^{n+1}(x)$, a contradiction.

(b) $y > x$ so $y \geq S(x)$. Since $y \in \Lambda, y > S(x)$. $y > S(x)$ so $y \geq S^2(x)$. Since $y \in \Lambda, y > S^2(x)$. $y > S^2(x)$ so $y \geq S^3(x)$. Since $y \in \Lambda, y > S^3(x)$. And so on.

(c) Let y be maximum in X. By (a), there is $x \in \Lambda$ and n with $y = S^n(x)$. By (b), x is the maximum element in Λ. □

By theorem 1.37, every infinite well-ordered set looks like a copy of $\mathbb{N}$ followed by a copy of $\mathbb{N}$ followed by... possibly topped off by a finite set.

1.5 Mathematical induction

Proofs by induction, and the related technique of construction by recursion, are important in set theory, especially induction and recursion on infinite well-ordered sets longer than $\mathbb{N}$. So in this section we remind ourselves about induction using $\mathbb{N}$.

Theorem 1.38. The principle of mathematical induction

Version I. *If $0 \in X$ and "$n \in X$" implies "$n+1 \in X$" for all natural numbers n, then $\mathbb{N} \subseteq X$.*

Version II. *If, for every natural number n "$k \in X$ for all natural numbers $k < n$" implies "$n \in X$", then $\mathbb{N} \subseteq X$.*

Version III. *If a natural number $j \in X$ and "$n \in X$" implies "$n + 1 \in X$" for all natural numbers $n \geq j$ then every natural number $n \geq j$ is an element of X.*

These statements are easily seen to be equivalent (once you see that the hypothesis of version II implies that $0 \in X$). It turns out that theorem 1.38 is a consequence of the fact that $\mathbb{N}$ is well-ordered and that every non-zero element of $\mathbb{N}$ is a successor.

Proof. Suppose $0 \in X$ and "$n \in X$" implies "$n+1 \in X$" for all natural numbers n. Let $N = \mathbb{N} \setminus X$. By well-ordering, N has a minimum element m. By hypothesis, $m \neq 0$. Hence there is k with $m = k + 1$. Since $k \in X$, by hypothesis $m \in X$, a contradiction. So $N = \emptyset$, i.e., $\mathbb{N} \subseteq X$. □

How do you organize proofs by induction? For proofs about the natural numbers (for example: the sum of the first n odd numbers is n^2) it's fairly obvious. Your induction hypothesis is essentially what you're trying to prove (to continue the example: if the the sum of the first n odd numbers is n^2 then the sum of the first $n + 1$ odd numbers is $(n + 1)^2$).

But if the theorem is not about natural numbers, it may be more difficult to recognize the appropriate induction hypothesis. We give three examples to show how this is done.

The first is a proof of proposition 1.36: Let X be a partial order, $x \in X$ so $S^m(x)$ exists. If $n < m$ then $S^n(x) < S^m(x)$.

Proof. Fix n. We need to show that for all $x \in X$ and all $m > n$, if $S^m(x)$ exists then $S^n(x) < S^m(x)$. So we use version III. I.e., we need to prove the following for arbitrary $m > n$, arbitrary $x \in X$: given the induction hypothesis: $S^n(x) < S^m(x)$, prove that if $S^{m+1}(x)$ exists, then $S^n(x) < S^{m+1}(x)$.

Here's the proof: Assume $S^n(x) < S^m(x)$. Since $S^{m+1}(x) = S(S^m(x)) > S^m(x)$, $S^n < S^m(x) < S^{m+1}(x)$ so by transitivity $S^n(x) < S^{m+1}(x)$. □

Our second proof by induction assumes naive set theory.

Definition 1.39. $\mathcal{P}(X)$ (called the power set of X) is the set of all subsets of X.

Theorem 1.40. *If X is finite, so is $\mathcal{P}(X)$. In fact, if X has size n, for some $n \in \mathbb{N}$, then $\mathcal{P}(X)$ has size 2^n.*

Proof. . We start with the base, $n = 0$: If X has no elements, $X = \emptyset$, and $\mathcal{P}(X) = \{\emptyset\}$, i.e., it has $1 = 2^0$ elements.

Our induction hypothesis is: every set of size n has a power set of size 2^n. From this we must prove that every set of size $n + 1$ has a power set of size 2^{n+1}.

So suppose every set of size n has a power set of size 2^n. Let X be a set of size $n + 1$. Let $x \in X$ and let $Y = X \setminus \{x\}$. Then Y has size n, so, by induction hypothesis, $\mathcal{P}(Y)$ has size 2^n.

Let $\mathcal{P}_0 = \{Z \subseteq X : x \in Z\}$ and let $\mathcal{P}_1 = \{Z \subseteq X : x \notin Z\}$. $\mathcal{P} = \mathcal{P}_0 \cup \mathcal{P}_1$. $\mathcal{P}_1 = \mathcal{P}(Y)$, so it has size 2^n. $\mathcal{P}_0 = \{Z \cup \{x\} : Z \in \mathcal{P}(Y)\}$, so it also has size 2^n. Hence $\mathcal{P}$ has size $2^n + 2^n = 2^{n+1}$. □

Our third proof by induction is also a formal proof of something we've already proved. First a definition:

Definition 1.41. Let X, Y be partially ordered sets under $\leq_X, \leq_Y$ respectively. X is isomorphic to Y iff there is a 1-1 onto function $f : X \to Y$ so that $x \leq_X x'$ iff $f(x) \leq_Y f(x')$.

Proposition 1.42. *If X is well-ordered, then so is each $L_n(X)$.*

Proof. L_1 is isomorphic to X, so it is well-ordered. Suppose $L_n(X)$ is well-ordered. We need to show that L_{n+1} is well-ordered.

Suppose $C = \{\vec{x}_m : m \in \mathbb{N}\}$ is an infinitely descending chain in $L_{n+1}(X)$. Since $L_n(X)$ is well-ordered, at most finitely many elements of C are in $L_n(X)$, so we may assume that all elements of C have length $n+1$. Let $\vec{x}_m = \vec{y}_m \frown k_m$. Since $\{\vec{y}_m : m \in \mathbb{N}\} \subseteq L_n(X)$ which is well-ordered, there is $\vec{y} \in L_n(X)$ and $m^* \in \mathbb{N}$ so that if $m > m^*$ then $\vec{y}_m = \vec{y}$. Hence $C = \{\vec{y} \frown k_m : m^* < m, m \in \mathbb{N}\}$. Since $\mathbb{N}$ is well-ordered, there are $k \in X, m^\dagger \in \mathbb{N}, m^\dagger \geq m^*$ with $k_m = k$ for all $m > m^\dagger$. Hence $C = \{\vec{x}_m : m \leq m^\dagger\} \cup \{y \frown k\}$, i.e., C is finite, a contradiction. □

A concept related to proof by induction is that of a recursive construction or definition, i.e., we construct or define something in stages, where what we do at stage n determines what we do at stage $n+1$. We already have seen recursive definitions, when we defined $L_n(X)$ (example 1.29) and when we defined $S^n(x)$ (definition 1.35). Now we give an example of a recursive construction.

Theorem 1.43. *Let $\{A_n : n \in \mathbb{N}\}$ be a sequence of infinite subsets of $\mathbb{N}$ so that each $A_n \supseteq A_{n+1}$. Then there is an infinite set $A \subseteq \mathbb{N}$ so that $A \setminus A_n$ is finite for each n, i.e., for each n, all but finitely many elements of A are elements of A_n.*

Proof. We construct $A = \{k_n : n \in \mathbb{N}\}$ recursively. Suppose at stage n, we have constructed $\{k_m : m \leq n\}$ and suppose $k_m \in A_m$ for all $m \leq n$.[8] Since A_{n+1} is infinite, it has some element k which equals no $k_m, m \leq n$. Let k_{n+1} be such a k. Continue.

We prove that this construction works: if $n < m$ then $A_n \supseteq A_m$, so if $n < m$ then $k_m \in A_n$. Thus each A_n contains all but finitely many elements of A. □

Using the proof of theorem 1.43 as a guide, we informally analyze how recursive constructions on the natural numbers work. We want to construct a set A. To do this, we construct a sequence of sets $\{B_n : n \in \mathbb{N}\}$ (in our example, $B_n = \{k_m : m \leq n\}$). The B_n's have to satisfy some requirement (in our example, that $B_n \setminus A_m \subset \{k_i : i < m\}$ for all $m < n$). Finally, A is some combination (in our example, the union) of the B_n's.

"Recursive construction on the natural numbers" does not mean that the set we're constructing is a subset of $\mathbb{N}$ (although in theorem 1.43 it was). It means that the *stages* of the construction are indexed by the natural numbers.

The proof that recursive constructions work is quite technical, and we won't give it. We will outline it, however. There are two parts: 1. A exists; 2. A does what it is supposed to do.

That A exists is a consequence of the axiom of replacement (see chapter 4); that A does what it is supposed to do is proved using the requirement on each B_n.

1.6 Filters and ideals

In later chapters we will need the combinatorial concepts of filter and ideal. Their definitions use the notions of set theory — union, intersection, difference, and so on — but since the reader is assumed to be familiar with these notions informally, we introduce filters and ideals here.

Definition 1.44. A filter on a set X is a family $\mathcal{F}$ of subsets of X so that

(a) if $F \in \mathcal{F}$ and $X \supseteq G \supseteq F$ then $G \in \mathcal{F}$.

(b) If $F_1, ..., F_n$ are elements of $\mathcal{F}$, so is $F_1 \cap ... \cap F_n$.

I.e., (a) says that $\mathcal{F}$ is closed under superset, and (b) says that $\mathcal{F}$ is closed under finite intersection.

If $\emptyset \notin \mathcal{F}$, we say that $\mathcal{F}$ is proper. If some $\{x\} \in \mathcal{F}$, we say that $\mathcal{F}$ is principal; otherwise it is nonprincipal. If $\mathcal{F}$ is proper and, for all $A \subseteq X$, either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$, we say that $\mathcal{F}$ is an ultrafilter.

[8] This sort of requirement serves the same function as the inductive hypothesis in a proof by induction.

Definition 1.45. A filterbase[9] on X is a family $\mathcal{B}$ of subsets of X so that $\mathcal{B}^{\supseteq} = \{A \subset X : \exists B \in \mathcal{B}\ A \supseteq B\}$ is a filter.

Example 1.46. Let $x \in \mathbb{R}$ and let $\mathcal{B}$ be the collection of open intervals $(x-r, x+r)$ where $r > 0$.

$\mathcal{B}$ is a filterbase on $\mathbb{R}$. $\mathcal{B}^{\supseteq}$ is a proper, nonprincipal filter on $\mathbb{R}$ and not an ultrafilter (neither $\{x\}$ nor $\mathbb{R} \setminus \{x\}$ are in $\mathcal{B}^{\supseteq}$).

Example 1.47. Let X be an infinite set, and let $\mathcal{F} = \{F \subseteq X : X \setminus F \text{ is finite}\}$.

$\mathcal{F}$ is a proper nonprincipal filter and not an ultrafilter (since every infinite set splits into two disjoint infinite subsets).

Example 1.48. Let X be any nonempty set, choose $x \in X$, and let $\mathcal{F} = \{Y \subseteq X : x \in Y\}$.

$\mathcal{F}$ is a proper, principal ultrafilter on X (since for every $Y \subseteq X$ either $x \in Y$ or $x \in X \setminus Y$).

If you are wondering where the examples are of proper, nonprincipal ultrafilters, the answer is that such filters need the axiom of choice for their construction. This will be done in chapter 4.

If you are wondering why we didn't define example 1.48 as $\mathcal{F} = \{Y : x \in Y\}$, the answer is because such a collection is too big to be a set. This notion of "too big to be a set" will be discussed in relation to the axiom schema of separation.

Here are some basic facts about ultrafilters.

Proposition 1.49. *Suppose $\mathcal{F}$ is an ultrafilter on a set X.*

(a) If $F \in \mathcal{F}$ and $G \subseteq F$ then either $G \in \mathcal{F}$ or $F \setminus G \in \mathcal{F}$.

(b) If $X = A_1 \cup ... \cup A_n$ then some $A_n \in \mathcal{F}$.

(c) If $\mathcal{F}$ is proper and contains a finite set, then $\mathcal{F}$ is principal.

Proof. We prove (a) and (b). For (a): If $G \notin \mathcal{F}$ then $X \setminus G \in \mathcal{F}$, so $F \cap (X \setminus G) = F \setminus G \in \mathcal{F}$.

For (b): If not, then, since no $A_i \in \mathcal{F}, \mathcal{F}$ is proper, and each $X \setminus A_i \in \mathcal{F}$. So $\bigcap_{i \leq n}(X \setminus A_i) \in \mathcal{F}$. But $\bigcap_{i \leq n}(X \setminus A_i) = \emptyset \notin \mathcal{F}$. □

The dual concept to a filter is that of an ideal.

Definition 1.50. An ideal on a set X is a family $\mathcal{J}$ of subsets of X so that

(a) If $J \in \mathcal{J}$ and $X \supseteq J \supseteq I$ then $I \in \mathcal{J}$.

(b) If $J_1, ...J_n \in \mathcal{J}$, then $J_1 \cup ... \cup J_n \in \mathcal{J}$.

I.e., $\mathcal{J}$ is closed under subset and finite union.

The connection between ideals and filters is

Proposition 1.51. *$\mathcal{J}$ is an ideal on X iff $\mathcal{F}_{\mathcal{J}} = \{X \setminus J : J \in \mathcal{J}\}$ is a filter on X.*

[9] also called a centered family

We say that a filter $\mathcal{J}$ is principal, nonprincipal or proper according to whether $\mathcal{F}_{\mathcal{J}}$ is principal, nonprincipal or proper. Similarly, $\mathcal{J}$ is a maximal ideal iff $\mathcal{F}_{\mathcal{J}}$ is an ultrafilter. $\mathcal{K}$ is a base for an ideal iff $\{I \subseteq K : K \in \mathcal{K}\}$ is an ideal.

Example 1.52. The collection of finite subsets of a given infinite set X is a proper ideal (its dual filter is example 1.47).

Example 1.53. The collection of closed intervals in $\mathbb{R}$ is a base for the ideal of bounded subsets of $\mathbb{R}$.

1.7 Exercises

1.Prove that example 1.4 is a partial order.

2. Prove that example 1.7 is a partial order.

3. Prove that example 1.8 is a partial order.

4. Prove that P1 and P3 hold for example 1.9.

5. Prove that the diagonal is dominating for $\leq_L$ in $\mathbb{R}^2$.

6. Which examples in section 1.1 are linear?

7. Define the dual R^{-1} of a relation R by $xR^{-1}y$ iff yRx.

(a) If R is an equivalence relation, what's R^{-1}?

(b) Show that $(R^{-1})^{-1} = R$.

(c) Show that R is a pre-order iff R^{-1} is, and that R is a partial order iff R^{-1} is.

8. Consider example 1.2.

(a) If $A \subseteq \mathbb{N}$, must A be linked?

(b) Is the set of primes a chain? an antichain? neither?

(c) Is the set of all multiples of 3 a chain? an antichain? neither?

(d) Is the set of all powers of 3 a chain? an antichain? neither?

9. Prove proposition 1.15.

10. Show that if $A \subset B \subset X$ a partial order, A is dominating in B and B is dominating in X, then A is dominating in X.

11. Consider the unit square $I^2 = [0,1] \times [0,1]$. For $(a,b),(c,d) \in I^2$ define $(a,b) \equiv (c,d)$ iff $a = c$ and either $b = d$ or $b = 1 - d$.

(a) Show that this is an equivalence relation.

(b) Show that $[(a, \frac{1}{2})]_{\equiv}$ has exactly one element, namely $(a, \frac{1}{2})$.

(c) If $b \neq \frac{1}{2}$, how many elements are in $[(a,b)]_{\equiv}$? What are they?

(d) Identifying equivalent points is the same as what geometric action on I^2?

12. Consider the unit square $I^2 = [0,1] \times [0,1]$. For $(a,b),(c,d) \in I^2$ define $(a,b) \equiv (c,d)$ iff $a = c$ and either $b = d$ or ($bd = 0$ and $b = 1 - d$).

(a) Show that this is an equivalence relation.

(b) Which points have equivalence classes with more than one element? How many points are in those equivalence classes? What are they?

(c) If you glue equivalent points together you get a geometric shape. What is it (up to topological equivalence)?

13. Let $Y = \mathbb{N}^{\mathbb{N}}$ and, for $f, g \in Y$ define $f \equiv_{ae} g$ iff $\{n : f(n) \neq g(n)\}$ is finite.

(a) Prove that $\equiv_{ae}$ is an equivalence relation.

(b) Define $X = \{[f]_{\equiv_{ae}} : f \in Y\}$. Let f be the identity function.Characterize all functions in $[f]_{\equiv_{ae}}$.

(c) Define X as in (b). Define $[f]_{\equiv_{ae}} \leq [g]_{\equiv_{ae}}$ iff $f \leq_{ae} g$. Show that $\leq_{ae}$ is well-defined and is a partial order on X.

14. Generalizing 13, let $\leq$ be a pre-order on a set Y. For $x, y \in Y$, define $x \equiv y$ iff $x \leq y \leq x$. Let $X = \{[x]_{\equiv} : x \in Y\}$. For $x, y \in Y$ define $[x]_{\equiv} \leq_* [y]_{\equiv}$ iff $x \leq y$. Show that $\leq_*$ is well-defined and is a partial order on X.

15. Show that a partial order in which every subset has a minimum element is a linear order (hence well-ordered).

16. Show that if $L_n(X)$ is well-ordered, so is X.

17. Find a subset of $\mathbb{Q}$ isomorphic to $L_3(\mathbb{N})$.

18. Find a well-ordered subset of $\mathbb{Q}$ with exactly 7 limit elements.

19. Find a well-ordered subset of $\mathbb{Q}$ with infinitely many limits.

20. What are the limits in examples 1.27 and 1.28?

21. Example 1.27 looks like how many copies of $\mathbb{N}$ following each other? What about example 1.28? What about each $L_n(\mathbb{N})$? How many limits does each $L_n(\mathbb{N})$ have?

22. Let X be well-ordered, Y a partial order (under $\leq_Y$), and let F be the set of functions from X to Y. For $f, g \in F$ define $f \leq_{lex} g$ iff $f = g$ or, for x the minimum element in $\{z \in X : f(z) \neq g(z)\}$, $f(x) <_Y g(x)$.[10]

(a) Show that $\leq_{lex}$ is a partial order on F.

(b) Show that if Y is linearly ordered, $\leq_{lex}$ is a linear order.

23. Prove proposition 1.37(b) by induction.

24. Let $\{f_n : n \in \mathbb{N}\} \subseteq \mathbb{N}^{\mathbb{N}}$. Show that there is a function $g \in \mathbb{N}^{\mathbb{N}}$ so that $\{k : g(k) < f_n(k)\}$ is finite for all n.

25. Prove that $\mathcal{F}$ is a proper filterbase iff for all finite nonempty $\mathcal{A} \subseteq \mathcal{F}$ $\bigcap \mathcal{A} \neq \emptyset$.

[10] This is a variant of the lexicographic order.

26. Prove proposition 1.49(c).

27. Prove proposition 1.51

2 Theories and their models

When a mathematical question outside set theory can be settled only by set-theoretic techniques, that is often because consistency results are involved. This revolutionary (to the mid-twentieth century) methodology can only be understood with reference to mathematical logic, in particular to models of first-order theories. An understanding of models is also crucial to the study of constructibility and large cardinals. The purpose of this section is to introduce the reader semi-formally to the important notions we will need. Our approach will be via examples, largely from the theory of partially ordered sets and its variations, and from the theory of groups.

2.1 First-order languages

First-order languages are very simple. Their verbs (called relations) are always in the present tense. "=" is always a verb; there may or may not be others. There may or may not be functions. There are two kinds of nouns: constants (somewhat like proper names, and optional) and infinitely many variables $x_0, x_1, x_2....$[11] Logical symbols like "$\wedge$" (and) "$\vee$" (or), "$\rightarrow$" (if... then) (these three are called logical connectives) and the negation "$\neg$" (not) are always included, as are the quantifiers "$\forall$" and "$\exists$." Parentheses are used to keep clauses properly separated (and we will be informal in our use of them). And that's it.[12]

First-order languages do not come close to ordinary language. They are missing so much — tense, adjectives, adverbs, prepositions... They are very limited. What they do have is very precise grammatical rules, i.e., syntax.

For example, in the language of partial orders, there are no functions, and the only verb besides "=" is "$\leq$." There are no constants. A sample sentence is: $\forall x\ \exists y\ \forall z((x \leq z \wedge x \neq z) \rightarrow (x \leq y \wedge y \leq z \wedge x \neq y)$. Not all strings of symbols are allowed: $x\forall \leq y = \exists$ does not have correct syntax. It is not part of the language.

For another example, in the language of the theory of groups, there is one binary function symbol, $\circ$, and one constant, e. A sample sentence is: $\forall x \forall y\ x \circ y = y \circ x$. Here's a string with bad syntax: $e \circ \vee xe$.

A sample formula in the language of the theory of groups is: $\forall y\ x \circ y = y \circ x$. This isn't a sentence because sentences have no ambiguities, that is, they have no free variables.This particular formula has one free variable, x. If you want to know whether it's true in a model, you have to know what x is. The formula $x \circ y = y \circ x$ has two free variables, x and y.

First-order languages have no meaning in themselves. There is syntax, but no semantics. At this level, sentences don't mean anything.

[11] But we'll denote variables by x, y, z etc. most of the time, for readability.

[12] I am being redundant here: Negation, one logical connective, and one quantifier suffice, since you can define the rest from this. I am being a little disengenuous: there is a way to avoid parentheses, the so-called Polish notation, but it is very difficult to read. I am also being a little sloppy: languages don't have relations, but relation symbols; they don't have functions, but function symbols. Finally, I will ignore the convention that differentiates between elements of the language, and their interpretations: $\leq, \in$ and so on can either be in the language, or relations in models, depending on context. In some sentences, they will play each role, in separate clauses.

2.2 Theories

Formally, a theory is a collection of sentences (for example, L1, L2, L3) and all its logical consequences. What's a logical consequence? It's what you can deduce from the rules of logic. What are the rules of logic? The details are in any logic text, but basically the rules of logic are what you've been using since you starting studying how to prove things in mathematics. Internally, theories are not true or false. Instead, there is the notion of logical consequence.

For example, if φ is in your theory, then $\varphi \vee \psi$ is also in your theory, for any ψ. That's because $\varphi \vee \psi$ is a logical consequence of φ. For another example[13] if φ is in your theory, and $\varphi \to \psi$ is in your theory, then ψ is in your theory. That's because ψ is a logical consequence of $\varphi \to \psi$ and φ. Finally, a third example: if $\forall x\ \varphi(x)$ is in your theory, so is $\exists x\ \varphi(x)$. That's because $\exists x\ \varphi(x)$ is a logical consequence of $\forall x\ \varphi(x)$

Thus, the *theory of partial orders* PO consists of all the first-order sentences you can deduce from L1, L2 and L3. The *theory of linear orders* LO consists of all the first-order sentences you can deduce from L1, L2, L3 and the following L4: $\forall x\ \forall y\ x \leq y$ or $y \leq x$.

That the sentences are first-order is crucial here. We originally defined a linear order as a partial order which is a chain. But the second word in the definition of "chain" was "subset," and this is not part of the first-order language. The first-order language of partial orders cannot speak of subsets. It can only speak of the elements in the order. L4 *is* in the first-order language, and that's why we have a first-order theory of linear orders.

What about well-orders? The definition of a well-ordered set was that every subset had a minimum element. Here, we can't avoid the notion of subset. There is no first-order set of sentences in the language of partial orders which captures the notion of well-order.[14]

Let's introduce two more theories: The *theory of dense linear orders* DLO consists all logical consequences of L1, L2, L3, L4 and the following L5: $\forall x \forall y\ ((x \leq y \wedge x \neq y) \to \exists z\ (x \leq z \wedge z \leq y \wedge x \neq z \wedge y \neq z))$. The *theory of discrete linear orders* DiLO consists of all logical consequences of L1, L2, L3, L4 and the following L6: $\forall x \exists y \forall z\ y \neq x \wedge y \geq x \wedge ((z \geq x \wedge z \neq x) \to y \leq z)$ — in English, L6 says that every element has a successor.

Here are the axioms of *the theory of groups* TG:

G1: (*associative axiom*) $\forall x\ \forall y\ \forall z\ (x \circ y) \circ z = x \circ (y \circ z)$.

G2: (*identity axiom*) $\forall x\ x \circ e = x$

G3: (*inverse axiom*) $\forall x\ \exists y\ x \circ y = e$.

You may remember proving, from these axioms the following statement φ: $\forall x\ e \circ x = x$. This means that φ is in the theory of groups.

We were very careful to say "the theory of groups" and not "group theory." That's because what we usually think of as group theory is very far from first-order. Consider a theorem like: "a cyclic group is commutative." No sentence in the first-order language of the theory of groups can capture the notion of "cyclic." Consider a theorem like: "the kernel of every homomorphism is a group." The first-order language can't talk about homomomorphisms or their kernels. It can't even

[13] the famous modus ponens

[14] This is not obvious. Its proof is by the compactness theorem of mathematical logic, showing that there is a non-standard model for the theory of $(\mathbb{N}, \leq)$.

talk explicitly about one group, much less two. It can only talk about the objects in one group at a time, just as the language of partial orders can only talk about the objects in one order at a time.

There is another kind of theory, the theory you get when you interpret a language in a structure. For example, using the set $\mathbb{N}$ we had two versions of $\leq$: the usual $\leq$, and $\leq_D$ (see definition 1.2). The first was a linear order. The second was not. We can talk about all the sentences that hold in $\mathbb{N}$ with the usual $\leq$. That's a theory. We can talk about all the sentences that hold in $\mathbb{N}$ with $\leq_D$. That's also a theory, a different theory.

In $\mathbb{R}$, if our version of $\circ$ is $+$, and our version of e is 0, then we'll have a group. The theory of all the sentences that hold in $\mathbb{R}$ with $+$ and 0 is a extension of the theory of groups. But if our version of $\circ$ is $\times$ and of e is 1, then we won't have a group (0 has no multiplicative inverse), and the theory of all the sentences which are true in this structure is not an extension of the theory of groups.

Theories don't give meaning, but they limit the kinds of meanings we can give. You can't assign $\circ$ to $\times$ and e to 1 and have a group. You can't assign $\leq$ to $\leq_D$ and have a linear order. Under the axioms of DLO, no element has a successor. But in DiLO, every element has a successor.

There is one kind of theory we want to avoid, and that is a theory which is inconsistent. An inconsistent theory is a theory in which every sentence in the language is a logical consequence. Every φ. Hence every not-φ. A theory which is not inconsistent is called consistent.

We say that two theories are incompatible iff their union is inconsistent. For example, one of the exercises asks you to prove that DLO and DiLO are incompatible.

To prove that a theory is inconsistent you could try to prove that every sentence φ in the language is a logical consequence. But there are infinitely many of them. This technique would take too long. It's better to derive a contradiction as a logical consequence.[15] That works by the following theorem of logic: If φ is a contradiction, then every sentence is deducible from φ.

To show that a theory is consistent you could try to find some sentence in the language that can't be proven from the axioms. But how do you know there isn't a proof out there somewhere? Again, this technique demands infinite time. It's better to use Gödel's completeness theorem. The completeness theorem says that a theory is consistent iff it has a model. To show that a theory is consistent, all you have to do is find a model.

What is a model? It's a structure in which the theory is true. Thus PO, LO, DO, DiO, GT are all consistent. We have seen models (in some cases, many models) for all of them.

Finally, we have the notion of consistency and independence. A sentence φ is consistent with a theory T iff $T \cup \{\varphi\}$ is consistent. φ is independent of T if both φ and $\neg\varphi$ are consistent with T.[16] We will give some examples in the next subsection.

Consistency and independence are very important in set theory. Consider a pretty basic claim: "the size of the set of real numbers is exactly [something specific]." For almost every version of [something specific] the claim is independent, that is, there are very few specific somethings that can be ruled out.

[15] A contradiction is something like $(\varphi \wedge \neg\varphi)$ or $\exists x\ x \neq x$. It's something that could never, under any interpretation, be true. See your nearest logic text for further discussion.

[16] But not, of course, at the same time.

2.3 Models and incompleteness

First-order languages have syntax but no semantics. Theories have something that looks a little like truth — logical consequence — but isn't exactly truth. When we get to models, we can talk about truth.

The precise definition of truth in mathematical logic is due to Tarski. It dates from the 1930's and was a very big deal. It is very technical, very precise, and is an inductive definition.[17] Here is a loose summary:

You start with a structure: a set, maybe some relations, maybe some functions, maybe some specific objects (like 0 or 1 in $\mathbb{R}$). You interpret the relation symbols, function symbols, and constant symbols of the language by the relations, functions, and objects of the structure. That's your model. Then, to see whether a sentence φ is true in the model, you apply Tarski's definition of truth. Which, luckily, corresponds to what you are already familiar with, e.g., to know whether a group satisfies $\forall x\ \forall y\ x \circ y = y \circ x$, you ask yourself whether, given any two elements in the group $a, b, a \circ b = b \circ a$.[18]

In a model, every sentence has a truth value. That means, given a sentence φ, it's either true or false in the model. To restate the obvious: if φ is not true in a model, it's false in the model. If it's not false in a model, it's true in the model.

A sentence has no free variables, but we are also interested in formulas which do have free variables.. For example, in example 1.5 we defined a relation $x \preceq y$ iff $\exists z\ z^2 = y - x$. We can recast this as follows: $\varphi(x, y)$ is the statement $\exists z\ z^2 = y - x$. Which pairs $(x, y) \in \mathbb{R}^2$ satisfy this formula?

It turns out that $\varphi(x, y)$ iff $x \leq y$, where $\leq$ is the usual partial order on $\mathbb{R}$. But the description "all pairs satisfying φ" is quite different from the description "$x \leq y$." Yet they describe exactly the same set of ordered pairs in $\mathbb{R}$.

For another example, "x is a unicorn in Africa" describes exactly the same set of animals as "x is a winged horse in Europe." Both sets are empty. But the descriptions are quite different.

In philosophy, descriptions are related to a notion called *intentionality*, which has to do with how the mind gives meaning to language. The actual thing described is related to a notion called *extensionality*. In mathematics we care only about extensionality. We don't care about intentionality.[19]

In this new meta-language, let's say things we already know.

Example 2.1. Suppose our set is $\mathbb{N}$ and our interpretation of $\leq$ is $\leq_D$. Let's call this model $\mathcal{N}_D$. Then $\mathcal{N}_D$ is a model of L1, L2, L3, but not L4, L5, L6. We write this as $\mathcal{N}_D \models \text{L1} \wedge \text{L2} \wedge \text{L3}$; $\mathcal{N}_D \not\models \text{L4} \vee \text{L5}$ *vee* L6. I.e., $\mathcal{N}_D \models \neg\text{L4} \wedge \neg\ \text{L5} \wedge \neg\ \text{L6}$. (We read the symbol "$\models$" as "models.")

[17] The induction is on linguistic complexity.

[18] So why the big deal? Tarski's definition of truth is a technical justification of what people sort of knew all along, which is what a lot of mathematical logic is. If you think about how often the obvious is false, the need for such a definition becomes clear. While Tarski meant his definition to apply to scientific methodology, in our context it's more of a definition of what it means to say that a sentence is true in a mathematical model, i.e., a definition of what a model is.

[19] Everyone loves a theorem that says two things with radically different descriptions are in fact the same. If we cared about intentionality, we'd object that, because the descriptions were radically different, they could not be describing the same thing.

The importance of staying within the model is clear from example 2.1. To repeat a statement from chapter 1, $2 = \frac{2}{3}(3)$, but, since $\frac{2}{3} \notin \mathbb{N}$, $2 \not\leq_D 3$.

Example 2.2. Suppose our set is $FIN(\mathbb{Q})$ and we interpret $\leq$ by $\leq_L$ (see example 1.8). Call this model $\mathcal{Q}_L$. Then $\mathcal{Q}_L \models$ L1 $\wedge$ L2 $\wedge$ L3 $\wedge$ L4. $\mathcal{Q}_L \not\models$ L5 $\vee$ L6. I.e., $\mathcal{Q}_L \models \neg$ L5 $\wedge \neg$ L6.

By examples 2.1 and 2.2, L4, is independent of PO. By the usual orders on $\mathbb{N}$ and $\mathbb{Q}$, L5 and L6 are independent of PO.

Example 2.3. Suppose X is the set of permutations on a set S where S has at least three elements, and we interpret $\circ$ by composition, e by the identity permutation. Call this model $\mathcal{P}_S$. Then $\mathcal{P}_S \models$ G1 $\wedge$ G2 $\wedge$ G3, but $\mathcal{P}_S \not\models \forall x \forall y\ x \circ y = y \circ x$.

I.e., $\mathcal{P}_S$ is not a commutative group. Since $\mathbb{R}$ with $+$ and 0 is a commutative group, the statement $\forall x \forall y\ x \circ y = y \circ x$ is independent of TG.

The last first order theory we introduce in this section is Peano arithmetic, PA. PA has a lot of axioms, and we won't give them all. They fall into two groups, which essentially say:

(1) Every element x has a successor $S(x)$; every element except 0 is a successor; and if $S(x) = S(y)$ then $x = y$.

(2) The principle of induction.

From (1) you can define $+$, $\times$, and 1 (= the multiplicative identity), i.e., you get arithmetic; and you can define $\leq$. Here's how to define $+$: $\forall x\ \forall y\ x + 0 = x$ and $x + S(y) = S(x + y)$.

The principle of induction should give us pause — when we stated it as theorem 1.38 we clearly used the notion of subset. How can we get around this problem?

The technique, which we'll also use for some of the set theoretic axioms, is to provide an axiom schema. That is, for every formula φ with one free variable, we add an axiom PI_φ: $(\varphi(0) \wedge \forall x\ (\varphi(x) \to \varphi(S(x))) \to \forall x\ \varphi(x)$.

Unlike our previous theories, PA has infinitely many axioms. So will set theory. But, even though they have an infinite set of axioms, both PA and set theory are what are known are axiomatizable theories: there is a cut-and-dried procedure, the kind of thing you can program on a computer, so that given any sentence we can find out, using this procedure, whether or not the sentence is an axiom of the theory.

Theorem 2.4. *(Gödel's first incompleteness theorem) Let $\mathfrak{N}$ be the structure $\mathbb{N}$ with the unary predicate S, i.e., successor under the usual interpretation. For every consistent axiomatizable theory which extends PA and is true in $\mathfrak{N}$ there is a sentence φ which is independent of the theory.*

Why is this such a big deal? After all, we already know of several sentences independent of PO, of LO, and a sentence independent of TG.

But PA has a different genesis. Rather than trying to capture commonalities among structures, it is trying to lay a foundation for the theory of one structure, $\mathfrak{N}$.[20] What theorem 2.4 is saying is that you cannot capture the theory of $\mathfrak{N}$ by a consistent, axiomatizable theory. The theory of

[20] Hence, via definitions, $\mathbb{N}$ with $+, \times, \leq, 0, 1$.

$\mathfrak{N}$ — all the sentences true in $\mathfrak{N}$ — is not axiomatizable. There is no nice procedure[21] which can churn out, given infinite time, all the true sentences of $\mathfrak{N}$.

Things get even more interesting, because PA can talk about itself. That is, you can code names in the language for its linguistic objects, and code formulas in the language whose standard interpretations are things like "this is a proof of that;" "this has a proof;" "that doesn't have a proof." You can even find a sentence whose standard interpretation is: "I have no proof." Call this sentence φ_{PA}. In fact, every consistent axiomatizable theory T which extends PA has a similar sentence φ_T.

More precisely, Gödel's first incompleteness theorem says: If T is a consistent axiomatizable theory which extends PA and holds in $\mathfrak{N}$, then φ_T is independent of T.

By the way φ_{PA} is constructed, $\mathfrak{N} \models \varphi_{\mathrm{PA}}$. So we have a sentence which we know is modeled by $\mathfrak{N}$ but which can't be proved by PA. In fact, no matter how many sentences we add to PA, as long as the theory T extends PA, is axiomatizable and $\mathfrak{N} \models T$, then $\mathfrak{N} \models \varphi_T$ but φ_T is independent of T.

Set theory is a consistent axiomatizable theory. But, since the language of set theory is not the language of PA, it doesn't strictly speaking extend PA. Instead, we say that PA is *interpreted* in set theory.[22] The interpretation of PA within set theory also holds in $\mathfrak{N}$. Gödel's theorem holds when we replace "extends" by "interprets" and the interpretation holds in $\mathfrak{N}$. So Gödel's theorem holds for set theory.

Furthermore, any consistent axiomatizable theory which interprets PA has a statement which essentially says "this theory is consistent." Which brings us to

Theorem 2.5. *(Gödel's second incompleteness theorem). A consistent axiomatizable theory which interprets PA and whose interpretation of PA holds in $\mathfrak{N}$ cannot prove its own consistency.*

I.e., PA can't prove it's own consistency. We think it's consistent because we think it holds in $\mathbb{N}$. But we have to reach outside PA to make this statement.

More to the point, set theory can't prove its own consistency. And, unlike PA, set theory doesn't have a natural model lying around for us to pick up. Whenever you do set theory, you have to accept that it might in fact be inconsistent.

Even worse, since just about all of mathematics[23] can be done within set theory (we'll do a little of this in chapter 4), we don't know if mathematics as we know it is consistent. You do differential equations at your own peril.

A final word about set theory as a theory: it can talk about subsets. So everything collapses into a first-order language. That is why we can embed essentially all of mathematics into set theory.

2.4 Exercises

1. (a) Define 0 in $\mathfrak{N}$. I.e., find a formula φ with one free variable so that $x = 0$ iff $\mathfrak{N} \models \varphi(x)$.

[21] "nice" means: can be carried out on an infinite computer

[22] This is another technical notion. The interested reader is referred to any mathematical logic text. "Interprets" is a weaker notion than "extends."

[23] The exception is category theory, but this can be embedded in a slight extension of set theory, adding the axiom that there is an inaccessible cardinal (see chapter 6).

(b) Define 1 in $\mathfrak{N}$.

(c) Define $\times$ in $\mathfrak{N}$.

(d) Define $\leq$ in $\mathfrak{N}$.

2. Which of the following (under their usual order) is a model of DO, DiO? $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

3. Let φ be the sentence: $\forall x\ \forall y\ x \leq y \leftrightarrow \exists z\ (z \times z) = y - x$.

(a) Let $\mathcal{R}$ be the structure whose underlying set is $\mathbb{R}$ where $+, \times, \leq, 0, 1$ are assigned as usual. Show that $\mathcal{R} \models \varphi$.

(b) Let $\mathcal{Q}$ be the structure whose underlying set is $\mathbb{Q}$ where $+, \times, \leq, 0, 1$ are assigned as usual. Show that $\mathcal{Q} \not\models \varphi$.

4. Show that DiLO and DLO are not compatible.

5. Interpret $\leq$ in $\mathbb{C}$ (the complex numbers) so that $(\mathbb{C}, \leq)$ is a linear order.[24]

6. Which of the following are (a) in (b) inconsistent with (c) independent of TG? Briefly justify your answers.

(i) $\exists x \exists y \forall z (z = x \vee z = y)$

(ii) $\forall x \exists z (z \neq x)$

(iii) $\forall x \forall y (x \circ y = e \rightarrow y \circ x = e)$

(iv) $\forall x \forall y \forall z (x \circ y = x \circ z \rightarrow y = z)$

(v) $\forall x \forall y \forall z (x \circ y = x \circ z \rightarrow x = z)$

(vi) $\forall x \forall y \forall z ((\exists w\ w \neq x \wedge x \circ y = x \circ z) \rightarrow x = z)$

7. The first order theory of infinite groups is axiomatizable. What is its set of axioms?

8. Define $x^0 = e, x^{n+1} = x \circ x^n$. Recall the definition of a cyclic group: $\exists x \forall y \exists n \in \mathbb{N} (y = x^n)$. Is this a first-order sentence in the language of TG? Why or why not?

[24] You may have learned that you can't put an order on the complex numbers, but what that means is that you can't put an order on $\mathbb{C}$ which extends $\leq$ and preserves the relation between $\leq$ and arithmetic that holds in $\mathbb{R}$.

3 The axioms, part I

3.1 Why axioms?

In the nineteenth and early twentieth centuries, mathematicians and philosophers were concerned with the foundations of mathematics far more urgently than they are today. Beginning with the effort in the early 19^{th} century to free calculus from its reliance on the then half-mystical concept of infinitesimals (there is a way of making infinitesimals precise, but it is modern, an unexpected fallout from formal mathematical logic), mathematicians interested in foundations were largely concerned with two matters: understanding infinity and exposing the basic principles of mathematical reasoning. The former will be dealt with later; it is with the latter that we concern ourselves now.

The basic principles of mathematical reasoning with which we will be concerned are the axioms of set theory. There are other basic matters to deal with, for example, the laws of logic that enable us to discriminate a proof from a nonproof, but these are the subject of mathematical logic, not set theory. Even restricting ourselves to deciding which statements are obviously true of the universe of sets, we find ourselves with several axiom systems to choose from. Luckily, the ones in common mathematical use are all equivalent, that is, they all prove exactly the same first-order theorems. This is as it should be if, indeed, these theories all codify our common mathematical intuition. The system we shall use is called ZF, for the mathematicians Ernst Zermelo and Abraham Fraenkel, although Thoralf Skolem did the final codification, and it was von Neumann who realized the full import of the axiom of regularity. ZF is the axiom system most widely used today.

Why should we bother with axioms at all? Isn't our naive notion of a set enough? The answer is no. Promiscuous use of the word "set" can get us into deep trouble. Consider Russell's paradox: is the set of all sets which are not members of themselves a set? If X is defined by "$x \in X$ iff $x \notin x$" then $X \in X$ iff $X \notin X$, a contradiction. No mathematical theory which leads to a contradiction is worth studying. And if we claim that we can embed all of mathematics within set theory, as we shall, the presence of contradictions would call into question the rest of mathematics. We want a minimal set of self-evident rules that don't obviously lead to a set like the one in Russell's paradox.[25]

Another reason for elucidating axioms is to be precise about what is allowed in mathematical arguments. For example, in theorem 1.31 of chapter 2, we constructed a set by finding a first element, then a second, then a third, an so on..., and then gathering everything together "at the end." Is this a reasonable sort of procedure? Most people prefer proofs to be finite objects, and this is an infinite process. The axioms of set theory will help us here. They will show that some infinite arguments are clearly sound (e.g., induction) and will also set limits on what we can do — the axioms of choice is controversial precisely because it allows a form of argument which some mathematicians find too lenient.

A third reason turns the first on its head. Instead of asking what the universe looks like, we want to get an idea of what it means to look like the universe, i.e., we want a sense of what it means to model our intuitions of what the universe is like. In order to talk about models, we need a consistent first-order theory. In order to know what the axioms are, we need the theory to be axiomatizable.

[25] Although by the preceding chapter, if those rules interpret PA their consistency is not guaranteed.

The remarkable thing about ZF is that, in just eight fairly simple axioms,[26] it accomplishes all this.

Let's examine the second axiom, the axiom of pairing, with respect to these concerns. Here's what it says: $\forall x\ \forall y\ \exists z\ z = \{x, y\}$. From the viewpoint of the first concern, self-evident rules, this is a statement about the universe of all sets — it says that, given any two sets (not necessarily distinct) in the universe, there's another set in the universe which has exactly those sets as elements. From the viewpoint of the second concern, elucidating what's allowed in mathematical arguments, it says that if you are talking about two objects, you can collect them together and talk about that collection too. From the viewpoint of the third concern, models of set theory, it says that if two objects x, y are in a model M of set theory, then there is a set $z \in M$ so $M \models z = \{x, y\}$. This third viewpoint is far more modest than the first or the second. In some sense it is the dominant viewpoint in this book.

ZF is a first-order theory, so, by the completeness theorem, if it is consistent it has models. But, by the second incompleteness theorem, we cannot show this from within ZF. The consistency of ZF and the existence of models of ZF are articles of faith.

Does ZF decide everything? By the first incompleteness theorem, if ZF is consistent, the answer is no. But the independent statement of the first incompleteness theorem is cooked up for that purpose. What about statements that mathematicians might be interested in? As it turns out, there are many independent statements of broad mathematical interest. Their associated questions are called *undecidable*. A classic undecidable question is: how many real numbers are there? Undecidable questions are found all across mathematics: algebra, analysis, topology, infinite combinatorics, Boolean algebra... even differential equations.

A note on how this chapter is written: because we want to embed all of mathematics in ZF, we have to be very picky about what we are allowed to talk about. We have to define basic things like ordered pairs and Cartesian products. We have to prove that they exist from the axioms.[27] Ideally, at each point we would use only the little bit of mathematics that has been deduced so far.

But, to be understandable, we have to illustrate what we are talking about. Many of the concepts are best illustrated with reference to examples which can only be justified later, when we know more, but which are intuitively clear because the reader is mathematically used to them. To avoid confusion, such examples will be called *informal examples*.

3.2 The language, some finite operations, and the axiom of extensionality

Any mathematical theory must begin with undefined concepts or it will end in an infinite regress of self-justification. For us, the undefined concepts are the noun "set" and the verb form "is an element of." We have an intuition of their meanings. The axioms are meant to be a precise manifestations of as much intuition as possible.

We write "$x \in y$" for "x is an element of y"; "$x \notin y$" for "x is not an element of y."

Of course, that's just our human interpretation of what we want things to say. In the first-order language, the word "set" is superfluous. We just have the relation symbol $\in$ and the logical shorthand $\notin$.

[26] not exactly, since two of them are axiom schemas

[27] hence that they exist in any model of ZF, if ZF has models...

The next step is to define basic concepts, i.e., to introduce new symbols that abbreviate longer phrases.

Definition 3.1. (a) $x \subseteq y$ iff $\forall z\ z \in x \to z \in y$; $x \subset y$ iff $x \subseteq y$ and $x \neq y$.

(b) $z = x \cup y$ iff $\forall w\ w \in z \leftrightarrow (w \in x \vee w \in y)$.

(c) $z = x \cap y$ iff $\forall w\ w \in z \leftrightarrow (w \in x \wedge w \in y)$.

(d) $z = x \setminus y$ iff $\forall w\ w \in z \leftrightarrow (w \in x \wedge w \notin y)$.

(e) $z = \emptyset$ iff $\forall w\ w \notin z$.

To remind ourselves that the symbol $\in$ doesn't have, in itself, any meaning, let's take a partially ordered set X where the partial order is $\leq$ and interpret $\in$ by $\leq$. Then $x \subseteq y$ iff $x \leq y$. Suppose, under this interpretation, $x \subseteq y$. Then $x \cup y = \{z : z \leq y\}$ and $x \cap y = \{z : z \leq x\}$. $x \setminus y$ is the interval $(y, x]$. And no z interprets $\emptyset$.

We want to prove theorems about the concepts defined in definition 3.1. So we need an axiom. The purpose of this axiom is to capture our intuition that a set is determined by its elements, not by its description. I.e., extensionality, not intensionality. This axiom will be sufficient to establish the algebraic properties of the concepts in definition 3.1.

Axiom 1. Extensionality Two sets are equal iff they have the same elements: $\forall x\ \forall y\ x = y \leftrightarrow \forall z\ (z \in x \leftrightarrow z \in y)$.

At this point, for readability, we will relax our formality a little and use words like "and," "or," etc., instead of logical symbols.

Theorem 3.2. *$\forall x\ \forall y\ x = y$ iff $(x \subseteq y$ and $y \subseteq x)$.*

Proof. $x = y$ iff $\forall z\ (z \in x$ iff $z \in y)$ iff $\forall z((z \in x \to z \in y)$ and $(z \in y \to z \in x))$ iff $(\forall z\ (z \in x \to z \in y)$ and $\forall z\ (z \in y \to z \in x))$ iff $(x \subseteq y$ and $y \subseteq x)$. □

Theorem 1 gives us a method of proof: to show two sets are equal, show each is a subset of the other.

Next, we use extensionality to give us the algebraic properties of union, intersection, and difference.[28]

Theorem 3.3. *$\forall x, y, z$*

(a) (idempotence of $\cup$ and $\cap$) $x \cup x = x = x \cap x$.

(b) (commutativity of $\cup$ and $\cap$) $x \cup y = y \cup x$; $x \cap y = y \cap x$.

(c) (associativity of $\cup$ and $\cap$) $x \cup (y \cup z) = (x \cup y) \cup z$; $x \cap (y \cap z) = (x \cap y) \cap z$.

(d) (distributive laws) $x \cup (y \cap z) = (x \cup y) \cap (x \cup z)$; $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$.

(e) (De Morgan's laws) $x \setminus (y \cup z) = (x \setminus y) \cap (x \setminus z)$; $x \setminus (y \cap z) = (x \setminus y) \cup (x \setminus z)$.

[28] Algebras with properties (a) through (e) are called Boolean algebras, after George Boole, whose book *An Investigation of the Laws of Thought* was an important 19th century investigation of basic logic and set theory.

(f) (relation of $\cup$ and $\cap$ to $\subseteq$) $x \subseteq y$ *iff* $x \cap y = x$ *iff* $x \cup y = y$.

(g) (minimality of $\emptyset$) $\emptyset \subseteq x$.

Proof. We content ourselves with proving (f) and (g), leaving the rest as an exercise.

For (f): Suppose $x \subseteq y$. If $z \in x$ then $z \in x \cap y$. Hence $x \subseteq x \cap y \subseteq x$, so $x \cap y = x$.

Suppose $x \cap y = x$. If $z \in x$ then $z \in y$. Hence $y \subseteq x \cup y \subseteq y$, so $y = x \cup y$.

Finally, if $x \cup y = y$ and $z \in x$ then $z \in x \cup y = y$, so $x \subseteq y$.

For (g): $\forall z\ z \notin \emptyset$, so the statement "$\forall z\ z \in \emptyset \to z \in x$" is vacuously true. □

Note that we still don't officially know that things like $x \cup y$, $x \cap y$, $x \setminus y$ exist.

Let's formally define curly brackets.

Definition 3.4. (a) $\forall x_0, ...\ \forall x_n\ z = \{x_0, ...x_n\}$ iff ($x_0 \in z$ and.. and $x_n \in z$) and $\forall y$ (if $y \in z$ then $y = x_1$ or... or $y = x_n$).

(b) if φ is a formula in the language of set theory, then $z = \{x : \varphi(x)\}$ iff $\forall y\ y \in z$ iff $\varphi(y)$.

Note that (a) is actually infinitely many many definitions, one for each n.

Curly brackets will also be used less formally, as in $\{2, 4, 6, 8..\}$ or $\{x_1, x_2, x_3...\}$, or, more generally, $\{x_i : i \in I\}$. These last uses will be formally defined later, but we need them earlier for examples.

3.2.1 Models of extensionality

The exercises ask what happens if you model $\in$ by $\leq$ and by $<$ in a partial order. Here we focus on what are called standard models, that is, where the symbol $\in$ is interpreted by the binary relation $\in$.

Informal example 3.5. For any x temporarily define $s^0(x) = x; s^{n+1}(x) = \{s^n(x)\}$. Let $X = \{s^n(\emptyset) : n \in \mathbb{N}\} = \{\emptyset, \{\emptyset\}\{\{\emptyset\}\}...\}$.

The claim is that $X \models$ extensionality. Why? Informally, it's because X can distinguish one element from another by looking inside. I.e., each nonempty element of X has exactly one element, which is itself an element of X, and no two elements of X have the same element. More formally, if $x \neq y$, we may assume $y \neq \emptyset$. So $y = \{z\}$ where $z \in X$ and $z \notin x$. I.e., $X \models y \setminus x \neq \emptyset$.

Informal example 3.6. Consider X as in informal example 3.5. Let $Y = X \setminus \{\{\emptyset\}\}$.[29]

The claim is that $Y \not\models$ extensionality. Why? Because Y can see no difference between $\emptyset$ and $\{\{\emptyset\}\}$. The only element of $\{\{\emptyset\}\}$ is $\{\emptyset\} \notin Y$. So, as far as Y is concerned, $\{\{\emptyset\}\}$ has no elements, i.e., is empty: $Y \models \forall x\ x \in \emptyset$ iff $x \in \{\{\emptyset\}\}$. Yet both $\emptyset$ and $\{\{\emptyset\}\}$ are distinct elements of Y. They are like identical twins — Y knows they are different, but cannot tell them apart.

This leads us to

[29] I.e., the element we omit from X is $\{\emptyset\}$. Be careful to distinguish $\{a\}$ from a.

Theorem 3.7. *$X \models$ extensionality iff, for all distinct $x, y \in X$, $X \cap ((x \setminus y) \cup (y \setminus x)) \neq \emptyset$.*

Proof. Suppose $X \models$ extensionality. Then for each distinct pair of elements $x, y \in X$ there is $z \in X$ with either $z \in x \setminus y$ or $z \in y \setminus x$. I.e., $X \cap ((x \setminus y) \cup (y \setminus x)) \neq \emptyset$.

Now suppose that, for each distinct pair of elements $x, y \in X$, $X \cap ((x \setminus y) \cup (y \setminus x)) \neq \emptyset$. Let $x \neq y, x, y \in X$. Let $z \in X \cap ((x \setminus y) \cup (y \setminus x))$. $X \models z \in x \setminus y$ or $X \models z \in y \setminus x$. I.e., $X \models \{z : z \in x\} \neq \{z : z \in y\}$. Hence, $X \models$ extensionality.

□

3.3 Pairs

In the previous section we defined $\{x, y\}$. Our second axiom justifies talking about it.

Axiom 2. Pairing $\forall x\ \forall y\ \exists z\ z = \{x, y\}$.

The pairing axiom is an example of a closure axiom: if this is in a model, so is that. In particular, the pairing axiom says that any model of set theory will think it is closed under pairs.[30] Also under singletons, since $\{x\} = \{x, x\}$, by extensionality.

The pairing axiom doesn't distinguish between elements in a pair — its pairs are pairs of socks, rather than pairs of shoes, where each pair has a left shoe and a right. Often in mathematics we need to distinguish between two objects in a pair. For example, in the Cartesian plane, the point (1,2) is not the point (2,1). There are several ways to do this within the axioms of set theory. The one used here is standard.

Definition 3.8. $\forall x\ \forall y\ (x, y) = \{\{x\}, \{x, y\}\}$.

Notice that, by axiom 2, any model of set theory is also closed under ordered pairs.

We show that definition 3.8 does what it's supposed to do.

Proposition 3.9. *$(x, y) = (z, w)$ iff $x = z$ and $y = w$.*

Proof. By extensionality, if $x = z$ and $y = w$, then $(x, y) = (z, w)$.

So suppose $(x, y) = (z, w)$. By extensionality, either $\{x\} = \{z\}$ or $\{x\} = \{z, w\}$. If $\{x\} = \{z, w\}$, then, by extensionality, $x = w = z$ and $\{x, y\} = \{z\}$, so $x = y = z$. Hence $x = z$ and $y = w$. If $\{x\} = \{z\}$, by extensionality $\{x, y\} = \{z, w\} = \{x, w\}$, so $y = w$. □

The pairing axiom seems fairly innocuous — of course any model of set theory should at least think it is closed under pairs! — but in fact it is very powerful. If you have pairs, you have ordered pairs. If you have ordered pairs, you have ordered n-tuples and finite dimensional Cartesian products (e.g., $\mathbb{R}^n$). Even more than that, you can code the language of set theory in the language of set theory. I.e., you have a major part of the machinery to prove Gödel's incompleteness theorems.

[30] This anthropomorphic language expresses informally what our definition of model of pairing says precisely. Anthropomorphic language is useful when thinking about models.

In the next section we will develop machinery that allows for both finite and infinitary tuples, and both finite and infinite dimensional Cartesian products.[31] For now, let's show that the pairing axiom allows us to define finite n-tuples and finite dimensional Cartesian products, just to show that it can be done.

Definition 3.10. (a) $(x_1, ...x_{n+1}) = \{(x_1, ...x_n), x_{n+1}\}$.

(b) $x_1 \times ...x_n = \{(y_1, ...y_n)$: each $y_i \in x_i\}$.[32]

Definition 3.10 explains why the ability to code ordered pairs within PA is the combinatorial key to Gödel's incompleteness theorems.

3.3.1 Models of pairing

Recall the set X of informal example 3.5, a model of extensionality. What do we have to add to make it a model of pairing?

Defining $s^n(\emptyset)$ as in this example, any model Y of pairing extending X would also have to contain each pair $\{s^n(\emptyset), s^m(\emptyset)\} = a_{n,m}$. And Y would also have to contain each pair $\{s^n(\emptyset), a_{m,k}\} = b_{n,m,k}$. And each pair $\{a_{n,m}, a_{i,j}\}$. And each pair $\{s^n(\emptyset), b_{m,k,j}\}$. And each pair...

We need an inductive definition.[33]

Informal example 3.11. Let $Y_0 = X$ where X is the set in informal example 3.5. Define $Y_{n+1} = \{\{x, y\} : x, y \in \bigcup_{i \leq n} Y_i\}$. Let $Y = \bigcup_{n \in \mathbb{N}} Y_n$.

I.e., Y is the closure of Y_0 under pairs.

The claim is that $Y \models$ pairing. Why? Suppose $x, y \in Y$. Then there is n with $x, y \in Y_n$. So $\{x, y\} \in Y_{n+1} \in Y$.

But the Y of informal example 3.11 is not the only extension of X satisfying the pairing axiom.

Informal example 3.12. Let Y be as in informal example 3.11. Let x^* be infinite. Let $Y_0^* = Y, Y_1^* = Y_0 \cup \{\{x, y, x^*\} : x, y \in Y_0^*\}$. Let Y^* be minimal so that $Y^* \supseteq Y_1^*$ and if $x, y \in Y^*$ and either $x \notin Y_0$ or $y \notin Y_0$ then $\{x, y\} \in Y^*$.

$Y^* \models$ pairing. Why? Note that every element of Y^* is finite. By definition, the only thing to check is that if $x, y \in Y_0^*$ then there is $z \in Y^*$ with $Y^* \models z = \{x, y\}$. Note that $z = \{x, y, x^*\}$ is as desired, since $\{w \in Y^* : w \in z\} = \{x, y\}$.

This leads to

Theorem 3.13. *$X \models$ pairing iff $\forall x \in X\ \forall y \in X\ \exists z \in X\ z \cap X = \{x, y\}$.*

Proof. Suppose $X \models$ pairing, and let $x, y \in X$. Then $\exists z \in X\ X \models (x, y \in z$ and if $w \in z$ then $w = x$ or $w = y)$. I.e., $\exists z \in X\ z \cap X = \{x, y\}$.

[31] Formal definitions of "finite" and "infinite" will be given in chapter 5.

[32] You might think that this depends on a definition of $\mathbb{N}$, but it doesn't, since the variables in our language are indexed by the natural numbers.

[33] modulo the fact that we formally don't have induction yet...

For the other direction, suppose $\forall x\ \forall y \in X\ \exists z \in X\ z \cap X = \{x, y\}$. Let $x, y \in X$ and let z be as in the hypothesis. Then $X \models (x, y \in z$ and if $w \in z$ then $w = x$ or $w = y)$. I.e., $X \models z = \{x, y\}$. $\square$

3.4 Cartesian products

In definition 3.10 we defined ordered n-tuples and Cartesian products in a way which does not generalize to infinite dimensions. The purpose of this section is to give a better, more general, definition.

Definition 3.14. (a) A relation is a set of ordered pairs.[34]

(b) A relation R is a function iff $\forall x, y, z$ if $(x, y) \in R$ and $(x, z) \in R$ then $y = z$.[35]

(c) The domain of a relation R is $\{x : \exists y\ (x, y) \in R\} = \text{dom } R$. For $A \subseteq \text{dom } R$, we write $R|_A = \{(x, y) \in R : x \in A\}$.

(d) The range of a relation R is $\{y : \exists x\ (x, y) \in R\} = \text{range } R$.

(e) The field of a relation $R = \text{dom } R \bigcup \text{ range } R = \text{field } R$.

(f) X is a domain (respectively range, respectively field) of a relation R iff $X \supseteq \text{dom } R$ (respectively range R, respectively field R).

Note that domains, ranges, fields of a relation are not uniquely defined, but *the* domain, range, or field of a relation R is uniquely defined.

Informal example 3.15. Let $f(n) = 2n$ for all $n \in \mathbb{N}$. Then dom $f = \mathbb{N}$; range $f =$ the set of even natural numbers; field $f = \mathbb{N}$; but $\mathbb{R}$ is a domain, a range, and a field of f.

Taking advantage of the shorthand supplied by our definitions, let's show how to define some of the notions from chapter 2 in the language of set theory.

Example 3.16. A relation R is an equivalence relation iff

E1. if $x \in \text{field } R$ then $(x, x) \in R$.

E2. If $(x, y) \in R$ then $(y, x) \in R$.

E3. If $(x, y) \in R$ and $(y, z) \in R$ then $(x, z) \in R$.

The exercises ask for similar definitions for various notions of order, and for groups. Note that the theory of well-orders suddenly becomes expressible in the first-order language of set theory, because we can refer to subsets in this language.

Let's remind ourselves of some familiar notation:

Definition 3.17. (a) If R is a relation, $R[A] = \{y : \exists x \in A\ (x, y) \in R\}$.[36]

(b) If f is a function and $(x, y) \in f$ we write $f(x) = y$.

[34] Strictly speaking, we are defining a *binary* relation.
[35] Strictly speaking, this is a *unary* function.
[36] This is usually used when R is a function.

(c) A relation on X is a subset of X^2.

(d) A function on X is one for which $X = \text{dom } f$. A partial function on X is one for which $\text{dom } f \subseteq X$.

(e) For any relation R, $R^{\leftarrow} = \{(y, x) : (x, y) \in R\}$.

(f) A function f is 1-1 iff $f^{\leftarrow}$ is a function.

(g) We write $f : A \to B$ iff f is a function, $\text{dom } f = A$, and $\text{range } f \subseteq B$.

(h) If $f : A \to B$ and $B = \text{range } f$, we say f is onto B.

(i) If $f : A \to B$ is 1-1 and onto B, we say f is a set isomorphism (or bijection).

If f is a function with $\text{dom } f = I$ and, for all $i \in I, f(i) = x_i$, we write $\{x_i : i \in I\}$ and use it in two ways:

Strict use: $f = \{x_i : i \in I\}$. That is, we deeply care that x_i is, informally, the i^{th} element of the list; x_i's salient feature is that it is $f(i)$. This is sometimes written as $(x_i : i \in I)$ or $< x_i : i \in I >$.

Casual use: $\{x_i : i \in I\} = \text{range } f$. That is, we do not particularly care how x_i got into the range of f, although I is used to index the elements of the set range f. In some combinatorial arguments we may require that each element be listed a specified number of times, e.g., just once (f is 1-1) or infinitely often, or...

Informal example 3.18. In casual use, $\{2n : n \in \mathbb{N}\}$ is the set of even natural numbers. In strict use, $\{2n : n \in \mathbb{N}\}$ is the function $f : \mathbb{N} \to \mathbb{N}$ for which each $f(n) = 2n$.

The ambiguity between strict use and casual use is not a problem, since the context always makes it clear what meaning we are using.

Definition 3.19. Given $\{x_i : i \in I\}$ we define the Cartesian product $\Pi_{i \in I} x_i = \{f : \text{dom } f = I \text{ and each } f(i) \in x_i\}$.

Elements of $\Pi_{i \in I} x_i$ are called choice functions (because they choose an element out of each x_i). If, for some x, each $x_i = x$, we write $\Pi_{i \in I} x_i = x^I = \{f : f \text{ is a function, } \text{dom } f = I, x \text{ is a range of } f\}$.

Once we have the natural numbers defined in the language of set theory, an ordered n-tuple will be an element of an n-dimensional Cartesian product $\Pi_{i \in \{0,1 \ldots n-1\}} x_i$.

3.5 Union, intersection, separation

Definition 3.20. (a) $\bigcup x = \{z : \exists y \in x \; z \in y\}$.

(b) $\bigcap x = \{z : \forall y \in x \; z \in y\}$.

Note that $x \cup x = \bigcup \{x, x\} = x$, but $\bigcup x$ need not equal $x \cup x$. Example: let $x = \{y\}$ where $y \neq x$. Then $x \cup x = x$ but $\bigcup x = y$.

When $X = \{x_i : i \in I\}$ we write $\bigcup_{i \in I} x_i = \bigcup X$ and $\bigcap_{i \in I} x_i = \bigcap X$.

The next two axioms say that if any sets exist, then the sets in definition 3.1 exist.

Axiom 3. Union $\forall x\ \exists y\ y = \bigcup x$.

An immediate corollary of axiom 3 is that $x \cup y$ exists, since $x \cup y = \bigcup\{x, y\}$. Similarly, given sets $x_0, ...x_n$, $x_0 \cup ... \cup x_n$ exists, since $x_0 \cup ... \cup x_n = \bigcup\{x_0, ...x_n\}$.

There is no intersection axiom: Vacuously, $\forall x\ x \in \bigcap\emptyset$. So if $\bigcap\emptyset$ were in our universe, our universe would be in our universe, which leads both to infinite regress (generally not considered to be a good thing)[37] and to paradoxes such as Russell's, as we will see below.

Axiom 4. Separation[38] Let φ be a formula with free variables $x_0, ...x_n$.. Then $\forall x\ \forall x_0 ... \forall x_{n-1} \exists y\ y = \{z \in x : \varphi(x_0, ...x_{n-1}, z)\}$.

Like the induction theorem, this separation axiom is a schema: there is one version for each formula φ.

Separation and the fact that Russell's paradox is a paradox can be used to prove

Theorem 3.21. $\nexists x\ \forall y(y \in x)\}$.

Proof. Suppose not. Let x be the universal set (i.e., $\forall y\ y \in x$) and let φ be the formula $z \notin z$. Let $y = \{z \in x : z \notin z\}$. By separation, $y \in x$. Hence $y \in y$ iff $y \notin y$, a contradiction. □

The next theorem finishes the proof that the objects in definition 3.1 exist, and also states the infinitary De Morgan laws.

Theorem 3.22. *(a)* $x \neq \emptyset$ *iff* $\exists y\ y = \bigcap x$.

(b) $\forall x\ \forall y\ \exists z\ \exists w\ z = x \setminus y$ *and* $w = x \cap y$.

(c) $\forall x\ \forall y$ *if* $y \neq \emptyset$ *then* $x \setminus \bigcap y = \bigcup\{x \setminus z : z \in y\}$.

(d) $\forall x \neq \emptyset\ \forall y\ x \setminus \bigcup y = \bigcap\{x \setminus z : z \in y\}$.

Proof. We've essentially already done (a); (b) and (d) are left as exercises. For (c): Suppose $y \neq \emptyset$. $w \in x \setminus \bigcap y$ iff ($w \in x$ and $w \notin \bigcap y$) iff ($w \in x$ and $\exists z \in y\ (w \notin z)$) iff $\exists z \in y\ (w \in x$ and $w \notin z)$ iff $w \in \bigcup\{x \setminus z : z \in y\}$. □

3.5.1 Models of union and separation

Let $X = \emptyset$. This is vacuously a model of union and of separation, because each of these axioms starts with a universal quantifier.

Now try $X = \{\emptyset\}$. $X \models$ union, because $\bigcup\emptyset = \emptyset$. And $X \models$ separation, because for all formulas φ with free variables $x_1, ...x_n$, $\forall y_1, ...y_{n-1} \in \emptyset\ \{y \in \emptyset : \varphi(y_1, ...y_{n-1}, y)\} = \emptyset \in X$.

For a third example, let $X = \{\emptyset, \{\emptyset, w\}\}$ where $w \notin X, w \not\subseteq X$. $X \models$ union. Why? For all $x \in X$ we need to find some $z \in X$ so $z \cap X = \bigcup x$. $\bigcup\emptyset \in X$. $\bigcup\{\emptyset, w\} = w$, and $X \cap w = \emptyset$: $\emptyset$ is the $z \in X$ that works for $\{\emptyset, w\}$.

I.e., what X doesn't know doesn't matter.

[37] although there are contemporary mathematicians who find uses for the theory of what are called non-regular sets

[38] also called Comprehension.

Theorem 3.23. *$X \models$ union iff $\forall x \in X\ \exists z \in X\ z \cap X = \bigcup\{y \cap X : y \in x \cap X\}$.*

Proof. $X \models$ union iff $\forall x \in X \exists z \in X$ $(u \in z \cap X$ iff $\exists y \in x \cap X\ z \in y)$ iff $\forall x \in X\ \exists z \in X\ z \cap X = \bigcup\{y \cap X : y \in x \cap X\}$. □

Now for separation. Because separation is a schema, with one axiom for each formula, this is a little tricky. For example, in the real world (whatever that is) you might have $\varphi(y)$ for some $y \in x$, but a potential model X might have $y, x \in X$ yet might not "know" that $\varphi(y)$.[39]. Because of its complexity, we won't prove the following theorem categorizing standard models of separation.

Theorem 3.24. *$X \models$ separation iff for every φ a formula with free variables $x_0, ...x_n$, every $x \in X$ and every $y_0 \in X ... y_{n-1} \in X$ there is $z \in X$ so that $z \cap X = \{y \in x \cap X : X \models \varphi(y_0, ...y_{n-1}, y)\}$.*

Theorem 3.24 is a little difficult to parse. But it has an easy consequence.

Corollary 3.25. *If $\forall x \in X\ \mathcal{P}(x) \subseteq X$ then $X \models$ separation.*

Proof. Suppose φ is a formula with free variables $x_0, ...x_n$. Let $y_0, ...y_{n-1} \in X$. Whatever $\{y \in x \cap X : X \models \varphi(y_0, ...y_{n-1}, y)\}$ is, it's in X. □

3.6 $\mathbb{N}$ at last

Back in 1884 the great German logician and philosopher Gottlob Frege published his groundbreaking *Die Grundlagen der Arithmetik* (*The Foundations of Arithmetic*) in the course of which, after approximately 90 or so pages, he defined the number one. His fundamental question was: "what is number?"

Our motivation is not as profound. We simply want to find a definition of each natural number n within the language of set theory, and then a way to define $\mathbb{N}$, so that we can talk about $\mathbb{N}$ and its elements. I.e., we want to *represent* the natural numbers in set theory. This is a modest task. It makes no claims of deciding either the ontological or epistemological status of numbers.

Back in chapter 1, we showed that once we have $\mathbb{N}$ we can define its arithmetic and its order relation. Having represented $\mathbb{N}$ within set theory, we will, in chapter 4 find a way to represent $\mathbb{Q}$ and $\mathbb{R}$ along with their arithmetic and order relations, to bolster our claim that essentially all of mathematics can be embedded into set theory.

The representation of natural numbers we will use is due to von Neumann. It has two guiding principles:

A Each n should have n elements.

B For each n and m, $m < n$ iff $m \in n$.

These principles, whose serendipity will become clear when we discuss ordinals and cardinals, give us no choice about our definitions.

[39] For example, consider the formula "y is the size of $\mathbb{R}$". Different models disagree.

By principle A, $0 = \emptyset$.

Suppose we have represented n. How do we represent $n+1$? By principle B, if $m \leq n$ then $m \in n+1$. By principle A, $n+1 = \{0, ...n\}$.

Knowing this, we have $\forall n > 0$ $n = \{0, ...n-1\}$. So each $n+1 = n \cup \{n\}$. I.e., $1 = \{0\} = \{\emptyset\}, 2 = \{0,1\} = \{\emptyset, \{\emptyset\}\}, 3 = \{0,1,2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, and so on.

Once we know that $\emptyset$ exists (wait just a little longer), each n will have been defined.

Now for $\mathbb{N}$. This is a thorny problem. We defined each natural number separately, in terms of its predecessors. This took infinitely many sentences. But we cannot say "$x \in \mathbb{N}$ iff $x = 0$ or $x = 1$ or $x = 2$ or..." because infinite sentences are now allowed. How can be capture our intuition about $\mathbb{N}$ in a finite sentence?

There are many ways to do this. The one we use looks at each natural number as an ordered set via principle B. What properties does each natural number have?

C Each x is well ordered by $\in$, and if $z \in y \in x$ then $z \in x$ (i.e., each natural number is transitive).

D Each nonempty x has exactly one limit element, 0.

E Each nonempty x has exactly one maximum, which we'll call $x-1$; $x = x-1 \cup \{x-1\}$.

Any set x satisfying properties C, D and E will be called a finite ordinal. We define $\mathbb{N}$ to be the collection of finite ordinals. That almost looks like we're applying separation, but we're not, since so far we haven't established any set of which $\mathbb{N}$ is a subset. We need an axiom to tell us that $\mathbb{N}$ is a set. We do this by restating in set theoretic language something with which we are already familiar.

Definition 3.26. The successor of x is defined as $x \cup \{x\}$ and is denoted by $S(x)$.

I.e., if n is a finite ordinal, $n+1 = S(n)$.

Definition 3.27. x is inductive iff $\forall y \in x$ $S(y) \in x$.

Axiom 5. Infinity $\exists x$ $\emptyset \in x$ and x is inductive.

This is the first axiom which baldly asserts that sets exist.

Proposition 3.28. $\exists x$ $x = \emptyset$.

Proof. Let x be inductive.[40] $\emptyset = \{y \in x : y \neq y\}$. □

Proposition 3.29. *Let x be inductive. Then every finite ordinal is an element of x.*

Proof. Suppose there is a finite ordinal $n \notin x$. Then $n+1$ is well ordered and nonempty, so $\{k \in n+1 : k \notin x\}$ has a least element n^* . Since $0 \in x, \exists m$ $n^* = m+1$. By definition of $n^*, m \in x$. But then, since x is inductive, $n^* \in x$, a contradiction. □

[40] The existence of such an x is guaranteed by the axiom of infinity.

Corollary 3.30. *$\exists z\ \forall y\ y \in z$ iff y is a finite ordinal.*

Proof. Let x be an inductive set. By the axiom of separation, $z = \{y \in x : y \text{ is a finite ordinal}\}$. □

The z of corollary 3.30 is, of course, $\mathbb{N}$.

Set theorists usually refer to $\mathbb{N}$ as ω. We will do this from now on. The notation $\mathbb{N}$ will refer to the natural numbers together with their order and arithmetic structure.[41]

3.6.1 Models of infinity

So far we have been able to characterize standard models of each axiom, that is, if we interpret the symbol $\in$ by the real relation $\in$, we have characterized exactly which sets are models of each axiom.

The reader is invited to attempt this for the axiom of infinity, but the complications are not worth our while. Instead we'll characterize models of infinity for transitive sets.

Definition 3.31. x is transitive iff $\forall y \in x\ y \subseteq x$.

Since every standard model of ZF is isomorphic to a standard transitive model, this is not too great a loss. Transitive sets will be explored extensively in chapter 4.

Theorem 3.32. *Let X be transitive. X is a model of infinity iff $\exists x \in X\ \emptyset \in x \cap X$ and $\forall y \in x \cap X\ S(y) \in x \cap X$.*[42]

Proof. Let X be transitive. $\forall x \in X$ (($\emptyset \in x \cap X$ and $\forall y \in x \cap X\ S(y) \in x \cap X$) iff $X \models (\emptyset \in x$ and $\forall y \in x\ S(y) \in x)$). □

3.7 Power sets

Recall definition 1.39: $\mathcal{P}(X)$ (called the power set of X) is the set of all subsets of X. Our axioms so far do not guarantee that $\mathcal{P}(X)$ exists. So we need an axiom.

Axiom 6. Power set $\forall x\ \exists y\ z \in y$ iff $z \subseteq x$.

Corollary 3.33. *Given a set $\{x_i : i \in I\}$, there is a set $y = \Pi_{i \in I} x_i$.*

Proof. First, note that $(x, y) = \{\{x\}, \{x, y\}\}$, so if $x, y \in z$ then $(x, y) \subseteq \mathcal{P}(\mathcal{P}(z))$, hence

$$(x, y) \in \mathcal{P}(\mathcal{P}(\mathcal{P}(z))).$$

Second, note that a function is a set of ordered pairs, so if $f \in \Pi_{i \in I} x_i$ then

$$f \subseteq \mathcal{P}(\mathcal{P}(\mathcal{P}(I \cup \bigcup_{i \in I} x_i)))$$

[41] Some authors use $\mathbb{N}$ to denote $\{1, 2, 3...\}$.

[42] Transitivity enters into this theorem by not loosening "$S(y) \in x \cap X$" to "$\exists z \in x \cap X\ X \models z = S(y)$...

hence

$$f \in \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathcal{P}(I \cup \bigcup_{i \in I} x_i)))).$$

Finally, note that

$$\Pi_{i \in I} x_i = \{f \in \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathcal{P}(I \cup \bigcup_{i \in I} x_i)))) : \text{ each } f(i) \in x_i\}$$

which is a set by substitution. □

But are any of these sets whose existence is guaranteed by corollary 3.33 nonempty? if I is infinite, we will need the axiom of choice for this. But if I is finite, we have

Proposition 3.34. *Let $n \in \omega$. If, for each $i \in n$, $x_i \neq \emptyset$, then $\Pi_{i \in n} x_i \neq \emptyset$.*

Proof. For each $i < n$ pick $y_i \in x_i$. Then $\{(i, y_i) : i \in n\} \in \Pi_{i \in n} x_i$. □

Combining power sets with finite Cartesian products gives us the set of all functions from x to y, denoted y^x: $y^x = \{f \subset \mathcal{P}(x \times y) : f \text{ is a function}\}$.

The axiom of regularity will allow us to build up the (or perhaps just a) set theoretic universe from $\emptyset$ and power sets alone.

3.7.1 Models of power set

Again, we restrict ourselves to transitive models.

Theorem 3.35. *Let X be a transitive set. $X \models$ power set iff $\forall x \in X \ \exists y \in X \ y \cap X = X \cap \mathcal{P}(x)$.*

Proof. Let X be transitive, and suppose $x \in X$. $\exists y \in X \ y \cap X = X \cap \mathcal{P}(x)$ iff $\exists y \in X \ z \in y \cap X$ iff $z \in X \cap \mathcal{P}(x)$ iff $\exists y \in X \ \forall z \subseteq x$ if $z \in X$ then $z \in y$ iff $X \models$ power set. □

Clearly if $\forall x \in X \ \mathcal{P}(x) \in X$, then $X \models$ power set. What is consequential is that this sufficient condition is not necessary. It is exactly this fact that opens the door to most independence results of mathematical interest.

For example, consider $\mathcal{P}(\omega)$. Suppose $\omega \in X$, X is transitive, and $X \models$ power set. $\mathcal{P}(\omega) \cap X \in X$, but many subsets of ω may be left out. This is why two models of ZF can disagree on how big $\mathcal{P}(\omega)$ is.

To give you a sense of the power of theorem 3.35, here are four independent statements from four fields of mathematics. You do not have to know what any of the definitions mean to get a sense of the broad sweep of independence results. In all cases, it is the ambiguity about power sets that makes each statement independent.

Example 3.36. Examples of independent statements

(a) From general topology: There is a perfectly normal nonmetrizable manifold.

(b) From functional analysis: Every algebraic homomorphism from the Banach space $C[0, 1]$ to an arbitrary Banach space is continuous.

(c) From algebra: Every uncountable Whitehead group is free.

(d) From measure theory: The union of fewer than continuum many measure zero sets has measure zero. (Here, "continuum many" means "the same size as $\mathbb{R}$.")

Statements (a) and (b) are free from any set-theoretic terminology, and (c) and (d) have only the barest trace (in talking about the sizes of infinite objects). The last chapter of this book discusses some combinatorial objects associated with independent statements.

3.8 Replacement

The final axiom of this chapter is the axiom of replacement.[43] This is our last elementary way of constructing sets. Essentially what it tells us is that if a function should exist, it does.

The context that necessitates replacement is the following. For any x, we've already defined $S(x)$ in definition 3.26. So we can define $S(\omega), S^2(\omega)$, and so on. We'd like to have a set consisting of $\omega \cup \{S^n(\omega) : n \in \omega\}$. But nothing we've done so far lets us do this.

Like the axiom of separation, replacement is, in fact, an axiom schema, one axiom for each of a particular type of formula.

Definition 3.37. A formula φ with two free variables is a functional iff $\forall x\ \forall y\ \forall z$ if $\varphi(x, y)$ and $\varphi(x, z)$ then $y = z$.

Some formulas are functionals because of logical consequence. For example the formula $y = x \cup \{x\}$ is a functional in any theory which includes the axioms of extensionality, pairing, and union.

Other formulas are functionals only in the context of a model. For example, the formula $y^2 = x$, defining a set of ordered pairs of the form (x, y), is a functional describing a partial function from $\mathbb{N}$ to $\mathbb{N}$, but is no longer a functional when the model is $\mathbb{Q}$.

Before stating the axiom of replacement, we need to complicate things a little further. Functionals may also have parameters, a sort of temporary name.

For example, consider the following formula $\varphi_r : y = \frac{x}{r}$ where r is a given (but not specified) element of $\mathbb{R}$ — r is the parameter. Each φ_r defines a different function from any other one, with φ_0 defining the empty function. But most real numbers aren't definable in the usual first order language of $\mathbb{R}$. So you can't talk about them directly.

To make things more complicated, we might have a formula ψ so that some choices of parameter create a function, and others don't, for example $\psi_r : y^2 = rx$. This is a partial function from $\mathbb{R}$ into $\mathbb{R}$ (constantly equal to 0) exactly when $r = 0$.

To avoid the technicalities of precisely defining a parameter, we will instead add variables. For example, consider $\varphi(x_0, x, y) : y = \frac{x}{x_0}$. If we "assign" x_0 to r we get φ_r. For another example, consider $\psi(x_0, x, y) : y^2 = x_0 x$. Again, "assigning" x_0 to r gets us ψ_r. But what does "assign" mean? We avoid that by the following locution which implicitly generalizes the notion of functional.

Axiom 7. Replacement Let φ be a formula with $n+2$ free variables $x_0, ...x_{n-1}, x, y$. $\forall x_0, ...\forall x_{n-1}$ (if $(\varphi(x_0, ...x_{n-1}, x, y)$ and $\varphi(x_0, ...x_{n-1}, x, z))$ implies $y = z$ then $\forall w\ \exists z\ y \in z$ iff $\exists x \in w\ \varphi(x_0, ...x_{n-1}, x, y)\}$).

[43] Also known as the axiom of substitution. All our other axioms are due to Zermelo. This is the one due to Fraenkel.

The form of the axiom of replacement isn't quite what's promised in the first paragraph. Instead, the form is the following: given a domain, and given a functional that should define a function on that domain, the range of the function that should exist exists. So the next step is to show that the function exists.

Proposition 3.38. *Let φ be a formula with $n+2$ free variables $x_0, ...x_{n-1}, x, y$. $\forall x_0, ...\forall x_{n-1}$ if $\forall x \forall y \forall z \varphi(x_0, ...x_{n-1}, x, y)$ and $\varphi(x_0, ...x_{n-1}, x, z))$ implies $y = z$ then $\forall w\ \exists f$ dom $f = w$ and $((x, y) \in f$ iff $\varphi(x_0, ...x_{n-1}, x, y)\}))$.*

Proof. Let z be as in replacement. Define $f = \{(xy) \in w \times z : \varphi(xy)\}$. □

Example 3.39. Define $\omega + \omega = \omega \cup \{S^n(\omega) : n \in \omega\}$.

We show that $\omega + \omega$ exists. Let $\varphi(x, y)$ be defined as $y = S^n(\omega)$ if $x = n$ for some $n \in \omega$; $y = 0$ otherwise. φ is a functional By replacement, $\{S^n(\omega) : n \in \omega\}$ exists. By union, $\omega + \omega$ exists.

Example 3.40. Define $V_0 = \emptyset$; $V_{n+1} = \mathcal{P}(V_n); V_\omega = \bigcup_{n<\omega} V_n$.

We show that V_ω exists. Define $\varphi(n, y)$ iff $\exists(x_0, ...x_n)\ x_0 = \emptyset, y = x_n$ and for each $i < n\ x_{i+1} = \mathcal{P}(x_i)$. φ is a functional. Let $z = \{y : \exists n \in \omega\ \varphi(n, y)\}$. By replacement, z exists. By union, V_ω exists.

Example 3.41. Let $\{a_n : n < \omega\}$ be a family of infinite subsets of ω so each $a_n \supseteq a_{n+1}$. There is an infinite set a so that $a \subseteq_{ae} a_n$ for each n.

Here $a \subseteq_{ae} b$ iff $a \cap b$ is infinite and $a \setminus b$ is finite.

We met this example before, in theorem 1.43. Now we carefully show how the recursive construction works within set theory.

We start with $\{a_n : n \in \omega\}$ and, by proposition 3.38, a function f defined by $f(n) = a_n$. As before, we recursively define k_n to be some element in a_n with $k_n \neq k_i$ for each $i < n$. But not just any element. We define k_n to be the least element of $a_n \setminus \{k_i : i < n\}$. This is because we need a functional: $\varphi(n, k)$ iff $k = \inf(a_n \setminus \{j : \exists i \in n \varphi(i, j)\}$. φ is a functional when its domain is restricted to ω. Finally, by replacement, $a = \{k : \exists n\ \varphi(n, k)\}$.

Later, when we have the axiom of choice, we will be able to do recursive constructions less restrictively.

3.8.1 Models of replacement

Theorem 3.42. *$X \models$ replacement iff for any φ a formula with $n+1$ free variables, $\forall x_0 ... \forall x_{n=1} \in X$ if $\forall x\ \forall y\ \forall z \in X\ X \models$ if $(\varphi(x_0, ...x_{n-1}, x, y)$ and $\varphi(x_0, ...x_{n-1}, x, z))$ then $y = z)$ then $\forall u \in X \exists w \in X\ w \cap X = \{y \in X : \exists x \in u \cap X\ \varphi(x_0, ...x_{n-1}, x, y)\}$*

The proof is left as an exercise.

Proposition 3.43. *Let V_ω be as in example 3.40. Then $V_\omega \models$ extensionality, pairing, union, separation, power set, and replacement.*

Proof. We will prove pairing, and replacement, leaving the rest to the exercises.

First we prove, by induction, that each $V_n \subseteq V_{n+1}$. Clearly $V_0 \subseteq V_1$. Suppose we know that $V_0 \subseteq V_1 \subseteq ...V_n, n \geq 1$. Let $y \in V_n \in V_{n+1}$. Then $y \subseteq V_{n-1}$. By induction hypothesis $V_{n-1} \subseteq V_n$. So $y \subseteq V_n$, i.e., $y \in V_{n+1}$.

Second we prove, by induction, that each V_n is transitive. Clearly V_0 is transitive. If $x \in V_{n+1}$ and $y \in x$ then $y \in V_n \subseteq V_{n+1}$, so $y \in V_{n+1}$.

For pairing: Suppose $\forall i < n$ if $x, y \in V_i$ then $\{x, y\} \in V_n$. Since each $V_n \subseteq V_{n+1}$, it suffices to show that if $x, y \in V_n$ then $\{x, y\} \in V_{n+1}$. If $x, y \in V_n$ then $\{x, y\} \subseteq V_n$, so $\{x, y\} \in \mathcal{P}(V_n) = V_{n+1}$.

For replacement: Let φ be a formula with $n+2$ free variables $x_0, ...x_{n-1}, x, y$. Fix $x_0, ...x_{n-1} \in V_\omega$ so $V_\omega \models \forall x \forall y \forall z$ if $\varphi(x_0, ...x_{n-1}, x, y)$ and $\varphi(x_0, ...x_{n-1}, x, z)$ then $y = z$. There is $k \in \omega$ with $x_0, ...x_{n-1} \in V_k$. Let $w \in V_\omega$. There is $m \in \omega$ with $w \in V_m$, $m \geq k$. Consider $Y = \{y : \exists x \in w : \varphi(x_0...x_{n-1}, x, y)\}$. Since w is finite, Y is finite, so there is $n \geq m$ with $Y \subseteq V_n$. Hence $Y \in V_{n+1} \subseteq V_\omega$. □

We still have two axioms to go, but let's stop for a moment to note that we have a model of everything so far except infinity. Let's also note that what we have so far is sufficient to embed PA, hence sufficient for Gödel's second incompleteness theorem. There's no hope of proving, in ZF, that there is a model of even the the axioms given in this chapter, much less all of ZF.

3.9 Exercises

1. Let X be a partial order under $\leq$ and interpret $\in$ by $<$. What are the interpretations of the symbols introduced in definition 3.1?

2. (a) Let X be a partial order under $\leq$. Show that if $\in$ is interpreted by $\leq$, then the structure models extensionality.

(b) Find a partially ordered set so that if you interpret $\in$ by $<$, the structure does not model extensionality.

3. Let X be as in informal example 3.5.

(a) Which infinite subsets of X model extensionality?

(b) Which finite subsets of X model extensionality?

4. Let $X = \omega$.[44]

(a) Show that if $x_k \in x_n \in x_m$ then $x_k \in x_m$.

(b) Show that $n < m$ iff $x_n \in x_m$, hence that $n = m$ iff $x_n = x_m$.

(c) Show that each x_n has exactly n elements.

(d) Show that $X \models$ extensionality.

(e) Show that every subset of X satisfies extensionality.

[44] I.e., the set of von Neumann natural numbers; see section 3.6.

5. Assume that $\forall x\ x \notin x$.[45] Show that if $X \neq \emptyset$ and $X \models$ pairing, then X is infinite.

6. Assume that $\forall x \forall y \neg(\ x \in y \in x$. Define $[x, y] = \{x, \{x, y\}\}$. Show that $[x, y] = [w, z]$ iff $x = w$ and $y = z$.

7. Prove the rest of theorem 3.3.

8. Following example 3.16, embed the theories of (a) partial orders; (b) linear orders; (c) well-orders into the language of set theory.

9. Following example 3.16, define a group in the language of set theory. You will need the following: a ternary relation is a set of ordered triples; a binary function is a set of triples R so if $(x, y, z) \in R$ and $(x, y, w) \in R$ then $z = w$; if R is a ternary relation, then field $R = \{x : \exists y\ \exists z\ ((x, y, z) \in R$ or $(y, x, z) \in R$ or $(y, z, x) \in R\}$.

10. Prove the rest of theorem 3.22

11. Prove the following:

(a) $\bigcup y \subseteq x$ iff $\forall z \in y\ z \subseteq x$.

(b) $x \subseteq \bigcup y$ iff $\forall z \in x\ \exists w \in y\ z \in w$

(c) $\bigcup\{x\} = x$.

12. A set X is closed under union iff $\forall x \in X\ \bigcup x \in X$.

(a) Find a set with exactly seven elements which is closed under union.

(b) FInd a set X which is linearly ordered by $\subseteq$ which is not closed under union.

(c) If $\forall x, y \in X\ x \cup y \in X$ must X be closed under union? Give a counterexample or a proof.

13. (a) Show that if x is a finite set of natural numbers, then $\bigcup x = \sup x$.

(b) Show that if x is an infinite set of natural numbers, then $\bigcup x = \omega$

14. Let X be as in informal example 3.5.

(a) Does $X \models$ pairing?

(b) Does $X \models$ union?

(c) Does $X \models$ separation?

15. (a) Characterize $\bigcup\bigcap x$ in the form $\{y : ...\}$.

(b) Characterize $\bigcap\bigcup x$ in the form $\{y : ...\}$.

(c) Find a set x where $\bigcup\bigcap x = \bigcap\bigcup x$.

(d) Find a set x where $\bigcup\bigcap x \neq \bigcap\bigcup x$.

16. Prove theorem 3.24.

17. Prove corollary 3.25 directly, without citing theorem 3.24.

18. Show that $\mathbb{N}$ is transitive, as is each $n \in \mathbb{N}$.

[45] This is a consequence of the axiom of regularity.

19. Characterize all inductive subsets of $\mathbb{N}$.

20. Let x be a set, and define f as follows: $f(0) = x; f(n+1) = \bigcup f(n)$ for all $n \in \omega$. Show that $\bigcup_{n\in\omega} f(n)$ is a set.

21. Prove theorem 3.42.

22. Prove the rest of theorem 3.43

23. For all $n \in \omega$ let V_n be as in example 3.40. Prove that if $n < m$ then $V_n \in V_{n+1}$.

24. Let $n \in \omega$. Which axioms does V_n model?

25. Show that if there is a nonempty finite model of power set, then there is a set $\{x_0...x_n\}$ with $x_0 \in x_1 \in ... \in x_n \in x_0$.

4 The axiom: part II: regularity and choice

The axioms of chapter 4 are generally considered the elementary, noncontroversial axioms of set theory. All of them except extensionality give us a way to construct sets: pairs, unions, definable subsets, power sets, infinite inductive sets, definable functions and ranges. I.e., they give us a way to build up the universe.

In this chapter we consider two axioms. The axiom of regularity essentially says that the sets constructed, beginning with $\emptyset$, via the axioms of chapter 4 are all that is. This axiom had a long development, beginning with Mirimanov in 1917 and ending with von Neumann in 1925. Adding this axiom essentially completes ZF.

The other axiom in this chapter, the axiom of choice, enlarges the kinds of mathematical arguments we can make within set theory. It was first articulated and defended by Zermelo in 1908. 62 years later, it remains controversial. Some mathematicians do not grant it the same clearly intuitive label that the other axioms have, so when we add it to our list of axioms we rename the list ZFC, to warn people that we are using choice. Even if you believe that the axiom of choice is true in the universe, there are many interesting models of ZF in which choice does not hold.[46]

4.1 Regularity, part I

The axiom of regularity says, in effect, that every nonempty set has a minimal element with respect to $\in$:

Axiom 8. Axiom of Regularity $\forall x \neq \emptyset\ \exists y \in x\ x \cap y = \emptyset$ (y is called an $\in$-minimal element of x).

Another way to express the axiom of regularity is to say that $\in$ is *well-founded*, where a relation R is well-founded iff $\forall x\ \exists y\ yRx$ and if zRy then $\neg zRx$. Regularity is not as self-evident as the other axioms. One the one hand the notion of a non-well-founded set is strange. It's not immediately obvious how to talk about one. On the other hand, who is to say for sure that there is none? And, in fact, the notion of non-well-founded sets is of some interest in computer science, where it is used to model non-terminating processes. Non-well-founded sets also have a place in philosophy and in linguistics.[47] Most of mathematics would go on the same with or without regularity. Yet it seems fundamental to our intuition about sets, and in fact (see section 4.5) it is the axiom of regularity that gives us a template for constructing the set-theoretic universe.

Regularity has a number of immediate consequences.

Proposition 4.1. $\forall x\ x \notin x$.

Proof. If $x \in x$ then $\forall y \in \{x\}\ \{x\} \cap y \neq \emptyset$. □

Proposition 4.2. *The axiom of regularity is equivalent to: there is no infinite descending $\in$-chain.*

[46] assuming that ZF is consistent

[47] Peter Aczel is a good reference for non-well-founded sets.

Proof. If there is a set $\{x_n : n \in \omega\}$ so each $x_{n+1} \in x_n$, then regularity would fail.

For the other direction, suppose regularity fails. Let x_0 be a set with no $\in$-minimal element. Let $x_1 \in x_0$. Since x_1 is not $\in$-minimal, there is $x_2 \in x_1 \cap x_0$. Since x_2 is not minimal, there is $x_3 \in x_2 \cap x_0$. And so on. $\{x_n : n \in \omega\}$ is an infinite descending $\in$-chain. □

This argument is familiar. It is similar to the proof of theorem 1.43. And, like the proof of theorem 1.43, we need the axiom of dependent choice, a weak form of the axiom of choice.

There are more global consequences of the axiom of regularity. To understand these, and to understand the axiom of choice, we need understand a basic understanding of ordinals. Ordinals cannot be understood without understanding transitive sets (see definition 3.31). So we turn to transitive sets.

4.2 Transitive sets

Recall that x is transitive iff $\forall y \in x\ y \subseteq x$. To anthropomorphize, a transitive set keeps no secrets from itself. It knows all there is to know about its elements. What it sees is all anyone gets.

We already have several examples of transitive sets: each natural number n; ω; V_ω.

Proposition 4.3. *Let X be transitive. Then $X \models$ extensionality.*

Proof. Suppose X is transitive. If $x, y \in X$ and $x \neq y$ then there is some $z \in (x \setminus y) \cup (y \setminus x)$. Since X is transitive, $z \in X$. □

Anthropomorphizing again, proposition 4.3 says that if a transitive set thinks two sets have the same elements then they really do, i.e., they are the same set. The next proposition says that if a transitive set thinks one of its elements is empty, it really is; if it thinks one of its elements is a pair, it really is; and so on.

Proposition 4.4. *Let X be transitive.*

(a) if $y \in X$ and $y \cap X = \emptyset$ then $y = \emptyset$.

(b) If $x, y, z \in X$ and $z \cap X = \{x, y\}$ then $z = \{x, y\}$.

(c) If $z, x, y \in X$ and $z \cap X = (x, y)$ then $z = (x, y)$.

(d) If $z, x \in X$ and $z \cap X = \bigcup x$ then $z = \bigcup x$.

(e) If $z \in X$ and $z \cap X$ is a relation, then z is a relation.

(f) if $z \in X$ and $z \cap X$ is a function, then z is a function.

Proof. All parts of this lemma are corollaries of the following fact: if X is transitive and $z \in X$ then $z \cap X = z$. □

Proposition 4.5. *X is transitive iff $\forall x \subseteq X\ \bigcup x \subseteq X$.*

Proof. If X is transitive, $x \subseteq X$ and $y \in \bigcup x$ then there is $z \in x$ with $y \in z$. $z \in X$ and X is transitive, so $y \in X$. Hence $\bigcup x \subseteq X$.

Suppose $\forall x \subseteq X \bigcup x \subseteq X$. Let $y \in X$. Then $\{y\} \subseteq X$ and $y = \bigcup\{y\} \subseteq X$, so X is transitive. □

Proposition 4.5 gives us a standard way to construct transitive sets.

Definition 4.6. Given x, define $TC_0(x) = x; TC_{n+1}(x) = \bigcup TC_n(x)$ for all $n \in \omega$. $TC(x) = \bigcup_{n\in\omega} TC_n(x)$.

I.e., $TC(x) = x \cup \bigcup x \cup \bigcup\bigcup x \cup ....$ $TC(x)$ is called the transitive closure of x. The next theorem says that it deserves this name.

Theorem 4.7. *$\forall x$ $TC(x)$ is transitive, and if $x \subseteq y$ and y is transitive, then $TC(x) \subseteq y$.*

Proof. Suppose $w \in TC(x)$ and $z \in w$. Then $w \in TC_n(x)$ for some $n \in \omega$. Hence $z \in TC_{n+1}(x) \subseteq TC(x)$.

Suppose $x \subseteq y$ and y is transitive. By induction, each $TC_n(x) \subseteq y$, so $TC(x) \subseteq y$. □

Corollary 4.8. *x is transitive iff $x = TC(x)$.*

Proof. $TC(x)$ is transitive, and $x \subseteq TC(x)$, so we need to prove that if x is transitive, $TC(x) \subseteq x$. Since x is transitive, each $TC_n(x) \subseteq x$, so $TC(x) \subseteq x$. □

4.3 A first look at ordinals

Definition 4.9. A relation $R \subseteq X^2$ is a strict well-order on X iff no xRx and the relation $R^=$ is a well-order on X, where $xR^=y$ iff xRy or $x = y$.

Definition 4.10. An ordinal is a transitive set strictly well-ordered by $\in$.

Note that each natural number is an ordinal. So is ω.

We usually reserve Greek letters for infinite ordinals.

Definition 4.11. Let α, β be ordinals.

(a) For all ordinals α, β, $\alpha < \beta$ iff $\alpha \in \beta$.

(b) For all ordinals α, $\alpha + 1 = S(\alpha)$.

(c) An ordinal α is a limit ordinal iff for all ordinals β, $\alpha \neq \beta + 1$.

E.g., ω is a limit ordinal.

Proposition 4.12. *Let α be an ordinal.*

(a) If $x \in \alpha$ then x is an ordinal.

(b) $S(\alpha)$ is an ordinal.

(c) $\alpha \notin \alpha$.[48]

(d) If β *is an ordinal then* $\beta \subset \alpha$ *iff* $\beta \in \alpha$.

Proof. For (a): If $x \in \alpha$ then, by transitivity, $x \subseteq \alpha$, hence, since well-ordering is hereditary, x is strictly well-ordered by $\in$. For x transitive, suppose $z \in y \in x$. Since α is transitive, $y \in \alpha$, hence $z \in \alpha$. Since α is linearly ordered by $\in$, either $z \in x$ or $x \in z$. If $z \in x$ we are done. If not, we would have an infinite descending chain $x \ni y \ni z \ni x \ni y....$ Hence α would not be well-ordered.

For (b): For transitive: If $y \in x \in S(\alpha)$ then either $y \in x \in \alpha$ which is transitive, or $y \in x = \alpha \subseteq S(\alpha)$. For well-ordered: if $y \subseteq S(\alpha)$ and $y = \{\alpha\}$ then y has a least (its only) element, so we may assume $y \cap \alpha \neq \emptyset$. Since α is an ordinal, y has a least element.

For (c): Otherwise $S(\alpha)$ would have the infinite descending chain $\alpha \ni \alpha \ni ....$

For (d): Suppose $\beta \subset \alpha$. Let γ be the least ordinal in $\alpha + 1$ with $\beta \subseteq \gamma$. By extensionality, $\beta = \gamma$. If $\gamma = \alpha$ then $\beta \not\subset \alpha$. So $\beta \in \alpha$.

□

The next proposition extends proposition 4.12(d).

Definition 4.13. Let L be linearly ordered by $\leq$, $A \subseteq L$. A is an initial segment iff $\forall l \in A$ if $k < l$ then $k \in A$.

Proposition 4.14. *If* A *is an initial segment of an ordinal* α *then either* $A \in \alpha$ *or* $A = \alpha$.

Proof. A is well-ordered and transitive, so it is an ordinal. Since it is an initial segment of α, $A \subseteq \alpha$. □

Thus, the ordinals are themselves well-ordered:

Proposition 4.15. *If* α, β *are ordinals, then either* $\alpha < \beta$ *or* $\alpha > \beta$ *or* $\alpha = \beta$.

Proof. Let $A = \alpha \cap \beta$. A is an initial segment of both α and β, so A is an ordinal and $A \in \alpha \cap \beta$. But then $A \in A$, contradicting proposition 4.12(c). □

The next proposition tells us about sets of ordinals.

Proposition 4.16. *Let* x *be a set of ordinals.*

(a) $\bigcup x$ *is an ordinal.*

(b) $\exists \alpha$ *an ordinal with* $x \subseteq \alpha$.

Proof. For (a): Since every element of a ordinal is an ordinal, $\bigcup x$ is a set of ordinals, hence strictly well-ordered. For transitive: If $\delta \in \gamma \in \bigcup x$ then there is $\beta \in x$ with $\gamma \in \beta$, hence $\delta \in \beta$, so $\delta \in \bigcup x$.

For (b): Let $\bigcup x = \alpha$ an ordinal. Then $x \subset \alpha + 1$. □

[48] Of course this follows from the axiom of regularity. But we don't need regularity to prove it.

An immediate corollary of proposition 4.16 is that every set of ordinals x has a least upper bound, denoted $\sup x$.

Now we generalize induction.

Theorem 4.17. Induction on ordinals I. *Let α be an ordinal, φ a formula with one free variable, and suppose we know that $\forall \beta \in \alpha$ if $\forall \gamma < \beta\ \varphi(\gamma)$, then $\varphi(\beta)$. Then $\forall \beta \in \alpha\ \varphi(\beta)$.*

If you substitute ω for α, this is just theorem 1.38, version II.

Proof. Suppose not. Then $\{\beta \in \alpha : \neg\varphi(\beta)\} \neq \emptyset$. Let β be its least element. Then $\forall \gamma < \beta\ \varphi(\gamma)$. So, by hypothesis, $\varphi(\beta)$, a contradiction. □

Theorem 4.18. Induction on ordinals II. *Let φ be a formula, and suppose for all ordinals β if $\forall \gamma < \beta\ \varphi(\gamma)$ then $\varphi(\beta)$. Then for all ordinals $\beta, \varphi(\beta)$.*

Proof. If not, let β be an ordinal with $\neg\varphi(\beta)$. Let $\alpha = \beta + 1$. Proceed as above. □

The difference between theorem 4.17 and theorem 4.18 is that the latter doesn't stop anywhere.

Closely related to induction are recursive constructions on the ordinals. Not surprisingly this generalizes recursive constructions on ω. Again, there are two versions, depending on whether we stop at some point (as in theorem 4.17) or don't (as in theorem 4.18). How does this work? For each ordinal $\beta < \alpha$ (respectively, for each ordinal β) a set X_β is constructed via some functional which depends on $\{X_\gamma : \gamma < \beta\}$. By substitution, X_β is a set. Recursive constructions will be key in the next section.

ON denotes the collection of all ordinals. We have to be very careful how we use it, since

Theorem 4.19. *ON is not a set.*

Proof. If ON were a set it would be well-ordered and transitive, so it would be an ordinal α with $\alpha \in \alpha$, contradicting proposition 4.12(c). □

Legal uses of ON include things like "Let $\alpha \in$ ON" (i.e., let α be well-ordered and transitive), or "Let $x \in \bigcup_{\alpha \in ON} \mathcal{P}(\alpha)$" (i.e., let x be a subset of some ordinal). Illegal uses are things like "Let ON $\in x$" or "let $y = \bigcup_{\alpha \in ON} \mathcal{P}(\alpha)$."

ON is an example of what are called classes. A class is defined by a formula, but is too big to be a set. Classes are best thought of as shorthand. Whenever we use class notation, we can always remove it by substituting a formula, as we did in the preceding paragraph.

There are versions of set theory which incorporate classes as an integral notion, the most popular one being Gödel -Bernays set theory (GB): the basic idea is that sets can be elements of classes, while classes can't be elements of anything. But GB proves exactly the same first order theorems as ZF, and carries a more cumbersome logical baggage. We lose nothing by restricting ourselves to ZF.[49]

[49]There are also type theories, which was Russell's approach to avoiding Russell's paradox: there is a hierarchy of types, and an object on one level can only be an element of objects on higher levels. These theories are more complicated, and they are more relevant to philosophy than to mathematics.

4.4 Regularity, part II

In section 4.1 we discussed some of the local consequences of regularity, that is, consequences for individual sets. Now we discuss the global consequence, i.e., the structure of the universe of all sets, V. Which, you will remember, is not a set. So it's rather remarkable that we can discuss its structure.

First, we need a recursive construction, over all ordinals.

Definition 4.20. $V_0 = \emptyset$. If α is an ordinal, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$. If α is a limit ordinal, $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$.

Note that we have already met the first ω stages of this construction in example 3.40, whose structure we explored in proposition 3.43 and several exercises in chapter 4. So the next proposition is no surprise; its proof is left as an exercise.

Proposition 4.21. *(a) Each V_α is a set.*

(b) Each V_α is transitive.

(c) If $\alpha < \beta$ then $V_\alpha \subset V_\beta$ and $V_\alpha \in V_\beta$.

Definition 4.22. If $x \in V_\alpha$ for some α, we define rank x = the least α with $x \in V_\alpha$.

Proposition 4.23. *(a) rank x is not a limit.*

(b) If $y \in x$ and rank $x = \alpha$ then rank $y < \alpha$.

Proof. For (a): If $x \in V_\alpha$ and α is a limit, then, by definition, $x \in V_\beta$ for some $\beta < \alpha$.

For (b): Suppose rank $x = \beta + 1$. Then $x \subseteq V_\beta$. So, for all $y \in x$, rank $y \leq \beta$. □

The next theorem says that every set is in some V_α. I.e., to use class shorthand, $V = \bigcup_{\alpha \in \text{ ON}} V_\alpha$.

Theorem 4.24. $\forall x\ \exists \alpha \in ON\ x \in V_\alpha$.

Proof. Let's say that x is small if $\exists \alpha \in \text{ON}\ x \in V_\alpha$; otherwise, x is big.

Note that if x is big then x has at least one big element: otherwise, if $\alpha = \sup\{\text{rank } y : y \in x\}$, then $x \in V_{\alpha+1}$. But then, if there is a big element there is an infinite decreasing $\in$-chain, which contradicts regularity.

□

In fact, theorem 4.24 is equivalent to the axiom of regularity.

Theorem 4.25. *If $V = \bigcup_{\alpha \in \text{ ON}} V_\alpha$ then every set has an $\in$-minimal element.*

Proof. [50] Suppose $V = \bigcup_{\alpha \in \text{ ON}} V_\alpha$. Fix $x \neq \emptyset$. Let $R = \{\text{rank } y : y \in x\}$. Let β be the minimum element in R, and let $y \in x$ with rank $y = \beta$. If $z \in y$ then rank $z < \beta$ by proposition 4.23(b). So $y \cap x = \emptyset$. □

[50] We can't use regularity in the proof, since that's what we're trying to prove.

The realization that regularity is equivalent to $V = \bigcup_{\alpha \in ON} V_\alpha$ is due to von Neumann.

Example 4.26. Suppose rank x = rank $y = \alpha$. Then rank $\{x, y\} = \alpha + 1$, rank $(x, y) = \alpha + 2$, rank $x \cup y = \alpha$, and rank $\bigcup x \leq \alpha$.

4.5 Choice

Axiom 9. Choice. A product of nonempty sets is nonempty: Given $\{x_i : i \in I\}$ where $I \neq \emptyset$ and each $x_i \neq \emptyset$, $\Pi_{i \in I} x_i \neq \emptyset$.

Earlier we mentioned that an element of $\Pi_{i \in I} x_i \neq \emptyset$ is called a choice function, because it chooses something from each x_i. If we think of each x_i as a basket, we may simultaneously dip into each basket and pull something out of each. We have as many hands as needed, at least one for each basket.

We abbreviate the axiom of choice as AC. It guarantees that something exists, but it does so in a highly nonconstructive fashion —- you may be able to say that $\Pi_{i \in I} x_i \neq \emptyset$, but you might not be able to point to a specific element in it.

Consider the difference between pairs of shoes and pairs of socks. If each x_i is a pair of shoes, we can easily define a function: $f(i)$ = the left shoe in the pair x_i. But if each x_i is a pair of socks, you don't have a definable function.[51] Saying that something exists without being able to say what exactly it is, is deeply suspicious to some mathematicians. "Choose?" they say. "Just *how* do you choose?"[52] At the same time, AC is embedded deeply in standard mathematics, and is simply assumed by most mathematicians.

Aside from theorem 4.25, here are some things that "ought" to be true which you can't prove without AC (or, without some weaker, but still non-constructive, version of it): Every field has a completion. Every vector space has a basis. The product of compact spaces is compact. If a linear order has no infinite decreasing sequence, then it is well-ordered (half of our theorem 1.31). We will give more examples in this chapter. AC is so deeply embedded in how we do mathematics, that when set theory was being axiomatized a little over a hundred years ago, it was almost overlooked, until Zermelo codified it sometime between 1904 and 1908.[53]

AC is also the source of some counter-intuitive results, the most famous of which is the Banach-Tarski paradox: you can decompose a solid sphere into finitely many pieces and recompose the pieces via rigid motions into two solid spheres, each with the same volume as the original sphere.[54]

AC holds the same relation to ZF as the parallel postulate holds to the other Euclidean axioms. It seems unnecessary, as if it ought to follow from the other axioms. It bothers people. And for a long time (although it took much less than thousands of years) people tried to show that it either followed from or was negated by ZF.

[51] We're assuming the socks in each pair can't be distinguished — no holes, no cute designs.

[52] The intuitionists, in particular, do not accept AC. They don't even accept the law of the excluded middle: since you don't know whether or not the sky will be blue above Lawrence Kansas at noon on July 4, 3000, the statement "At noon on July 4, 3000 the sky above Lawrence Kansas will either be blue or not blue" has no truth value to an intuitionist.

[53] The process of making the implicit explicit is not always easy.

[54] The pieces necessarily are not measurable, and in fact the existence of a non-measurable set is also a consequence of AC.

In 1935 Gödel showed that ZFC is consistent if ZF is. In 1963 Cohen shows that ZF + ¬AC is consistent if ZF is. So AC is independent from ZF.

We will, as most mathematicians do, assume AC. But set theorists are also interested in models of ZF in which choice fails, for three good reasons: 1. You can prove theorems about ZFC by starting with models of ZF in which strong combinatorial principles hold which negate choice. 2. Some of these combinatorial principles that negate choice (e.g., the axiom of determinacy, AD) are interesting in themselves. 3. Given strong enough assumptions about large cardinals, there are interesting natural models of ZF in which choice fails spectacularly. So, while we will assume AC, we will not be complacent about it.

4.6 Equivalents to AC

A large number of mathematical statements turn out to be equivalent to AC. We consider only three of them. These three essentially give methods of proof. After we prove they are equivalent, we will use them to prove various mathematical theorems.

Definition 4.27. The principle of well ordering (WO). Every set can be well-ordered: $\forall x\ \exists \alpha \in \text{ON}\ \exists f : \alpha \to x$ f is a 1-1 onto function.

The next two principles are about partial orders.

Definition 4.28. Zorn's lemma (ZL). If every nonempty chain in a nonempty partial order P has an upper bound, then P has a maximal element.

Definition 4.29. Hausdorff's maximal principle (HMP). Every nonempty chain in a partially ordered set can be extended to a maximal chain.

The condition in ZL that every nonempty chain has an upper bound is crucial. Consider example 1.7, the collection of finite sequences from some underlying set, under the partial order of extension. There can be no maximal element, since any finite chain sits below some longer finite chain. HMP gives you a maximal chain in this partial order, not a maximal element.

Theorem 4.30. *AC iff WO iff ZL iff HMP.*

Proof. We will show that AC $\to$ WO $\to$ ZL $\to$ HMP $\to$ AC.

Before beginning the proof, note that if a collection of chains in a partial order P is linearly ordered by $\subseteq$, then its union is a chain in P. The proof is left as an exercise.

(i) For AC $\to$ WO: Fix x. Let $I = \mathcal{P}(x) \setminus \{x\}$. For $i \in I$ let $y_i = x \setminus i$. By AC, let $g \in \Pi_{i \in I} y_i$ and define $f : \mathcal{P}(x) \to x \cup \{x\}$ as follows: $f(i) = g(i)$ for $i \in I$; $f(x) = x$. We define a function h by induction: $h(0) = f(\emptyset)$. If we know $h|_\alpha$ then $h(\alpha) = f(h[\alpha])$ if $h(\alpha) \neq x$; otherwise $h(\alpha) = x$.

Note that if $h(\alpha) \neq x$ then $h(\alpha) \in x \setminus h[\alpha]$, hence range $h \subseteq x \cup \{x\}$.

If $\beta < \alpha$ then $h(\beta) \in h[\alpha]$, so $h(\beta) \neq h(\alpha) \notin h[\alpha]$. Hence h is 1-1.

Define $z = \{w \in x : \exists \beta\ h(\beta) = w\}$. By separation, z is a set. So $\exists \alpha\ \alpha$ is the least ordinal with $\alpha > \beta$ for all β such that $\exists w \in x\ h(\beta) = w$. By definition, $h|_\alpha$ is onto z.

Suppose $x \setminus z \neq \emptyset$. Then $h[\alpha] \neq x$ and $h(\alpha) = y \in x$. But then $\alpha > \alpha$, a contradiction. So $x = z$ and $h|_\alpha$ is 1-1 onto x.

(ii) For WO $\to ZL$: Let P satisfy the hypothesis of ZL under the partial order $\leq_P$, and by WO let $P = \{p_\beta : \beta < \alpha\}$ for some ordinal α.

Define $f : \alpha \to P$ by induction as follows: $f(\beta) = p_\delta$ if δ is the least ordinal with $p_\delta >_P p$ for all $p \in f[\beta]$. (Hence $f(0) = p_0$). Otherwise $f(\beta) = p_0$. Note that if $f(\beta) = p_\delta$ then either $\delta = 0$ or $\delta \geq \beta$.

$f[\alpha]$ is a chain. By hypothesis it has an upper bound, p. If p is not maximal in P, then there is some $p_\gamma > p$, hence $p_\gamma > q$ for all $q \in f[\gamma]$. So $f(\gamma) = p_\gamma \leq p$, a contradiction.

(iii) For ZL $\to$ HMP: Let P be a partial order under $\leq$, let C be a nonempty chain in P, and let $\mathcal{S} = \{D : D \text{ is a chain in } P \text{ and } D \supseteq C\}$. Order $\mathcal{S}$ by $\subseteq$. A chain in $\mathcal{S}$ is a collection of chains of P, all extending C, linearly ordered by $\subseteq$.

If $\mathcal{C}$ is a chain in $\mathcal{S}$ then $\bigcup \mathcal{C}$ is a chain. So every chain in $\mathcal{S}$ has an upper bound. Hence, by ZL, $\mathcal{S}$ has a maximal element, $\mathcal{C}$. I.e., $\mathcal{C}$ is a maximal chain in $\mathcal{S}$. And $\bigcup \mathcal{C}$ is a chain extending C.

We show that $\bigcup \mathcal{C}$ is a maximal chain in P. If not, there is $p \in P$ so $\{p\} \cup \bigcup \mathcal{C}$ is a chain. But then $\mathcal{C}$ would not have been maximal in $\mathcal{S}$, since $\mathcal{C} \cup \{\{p\} \cup \bigcup \mathcal{C}\})$ would also be a chain in $\mathcal{S}$.

(iv) For HMP $\to$ AC: Given $\{x_i : i \in I\}$ where I and each x_i are nonempty, let $\mathcal{F}$ be the set of partial functions whose domains are subsets of I so if $f \in \mathcal{F}$ and $i \in \text{dom } f$ then $f(i) \in x_i$, i.e., each $f \in \mathcal{F}$ is a partial choice function.

Order $\mathcal{F}$ by inclusion, i.e., by $\subseteq$. Fix some $y \in x_0$. $\{(\{0, y)\}\} \in \mathcal{F}$, so there is a maximal chain F extending it. Let $f = \bigcup F$.

We show that f is a function: if $i \in \text{dom } f$ there is $g \in F$ $g(i) = f(i)$. Suppose $i \in \text{dom } g'$ with $g' \in F$. Either $g' \subseteq g$ or $g \subseteq g'$, so $g'(i) = g(i)$.

We show that dom $f = I$. If not, there is $i \in I \setminus \text{dom } f$. Let $z \in x_i$. Then $f \cup \{(i, z)\} \in \mathcal{F}$, so F was not maximal.

Hence $f \in \Pi_{i \in I} x_i$ as required. □

The next task is to use AC in its several guises. We will prove two theorems using AC, each in three different ways: from WO, from ZL, and from HMP.[55]

Theorem 4.31. *Every linear order has a well-ordered cofinal subset.*

Proof. Method I, from WO: Let X be linearly ordered by $\leq_X$. By WO, for some ordinal α, $X = \{x_\beta : \beta < \alpha\}$. Let $C = \{x_\gamma : \forall \delta < \gamma \; x_\delta <_X x_\gamma\}$. $x_0 \in C$, so C is nonempty.

C is well-ordered: $\{\gamma : x_\gamma \in C\}$ is a subset of α, hence well-ordered. And if $x_\gamma, x_{\gamma'} \in C$ and $\gamma < \gamma'$ then $x_\gamma <_X x_{\gamma'}$.

[55] AC itself is hardly ever used directly to prove things. Its main purpose is to appear intuitive, so as to disguise its strangeness. WO, ZL, and HMP are not so obviously true. Here's a story. A long time ago a physicist friend asked me why there was a Hamel basis for $\mathbb{R}$. (In the context of his question, this is a base for $\mathbb{R}$ considered as a vector space over $\mathbb{Q}$.) "Well," I said, "let's well-order $\mathbb{R}$." "Well-order! Well-order!"my friend exclaimed. "How can you *do* that?" But if I'd told him that every product of nonempty sets is nonempty I'm sure he would have accepted it.

C is cofinal: Given $x_\beta \in X$, let γ be the least ordinal $\geq \beta$ with $x_\gamma \in C$. Either $\beta = \gamma$ or $x_\beta <_X x_\gamma$.

Method II, from ZL: Let $\mathcal{W}$ be the collection of all well-ordered subsets of X ordered by end-extension, i.e., if $W, W' \in \mathcal{W}$, $W \leq_{\mathcal{W}} W'$ iff W is an initial segment of W'.

We show that $\mathcal{W}$ has no infinite decreasing chains: If $W_0 >_{\mathcal{W}} W_1 >_{\mathcal{W}} ...$ then there are $x_i \in W_i \setminus W_{i+1}$ for all $i \in \omega$. Since W_{i+1} is an initial chain of W_i, each $x_i >_X x_{i+1}$. So W_0 would not be well-ordered, a contradiction.

A chain $\mathcal{C}$ in $\mathcal{W}$ is a collection of well-ordered subsets of X so, given $W, W' \in \mathcal{C}$, either W is an end-extension of W' or W' is an end-extension of W. To apply ZL we must show that every nonempty chain in $\mathcal{W}$ has an upper bound.[56]

Let $\mathcal{C}$ be a nonempty chain in $\mathcal{W}$. $\bigcup \mathcal{C}$ is linearly ordered, and $\bigcup \mathcal{C}$ has no infinite descending chains, so it is well-ordered. Hence $\bigcup \mathcal{C} \in \mathcal{W}$, and $\bigcup \mathcal{C}$ is an upper bound for $\mathcal{C}$.

Hence, by ZL, $\mathcal{W}$ has a maximal element W. Since W is maximal, if $x \in X$ there is $y \in W$ with $x \leq y$. I.e., W is cofinal in X.

Method III, from HMP: Let $\mathcal{W}$ be as in method II. By HMP let $\mathcal{C}$ be a maximal chain in $\mathcal{W}$. By the same proof as in method II, $\bigcup \mathcal{C}$ is a well-ordered cofinal subset of X. □

Note that the proof from WO has a slightly constructive flavor: once we're given the well-ordering of X, we proceed inductively. Set theorists tend to prefer proofs from WO because of this. Proofs from ZL and from HMP seem to obscure what's going on.

Recall the definition of filter (definition 1.44) from chapter 2.

Theorem 4.32. *If F is a proper filter on a set x then it extends to an ultrafilter on x.*

Proof. Method I, from WO: Let F be a proper filter on x. By WO, $\mathcal{P}(x) = \{y_\beta : \beta < \alpha\}$ for some ordinal α. We construct a sequence $\{F_\beta : \beta < \alpha\}$ of subsets of $\mathcal{P}(x)$ so that:

1. $F_0 = F$
2. If $\beta < \delta$ then $F_\beta \subseteq F_\delta$
3. If A is a finite subset of some F_β then $\emptyset \neq \bigcap A \in F_\beta$
4. If $\beta < \alpha$ then either $y_\beta \in F_{\beta+1}$ or $x \setminus y_\beta \in F_{\beta+1}$.
5. Each F_β is closed under superset.

These conditions ensure that $G = \bigcup_{\beta<\alpha} F_\beta$ is an ultrafilter.

So it suffices to construct such a sequence.

Suppose we know F_γ for all $\gamma < \beta$. If β is a limit ordinal, $F_\beta = \bigcup_{\gamma<\beta} F_\gamma$. The reader can easily check that conditions (2) through (5) hold.

[56] The biggest mistake students make in using ZL is to forget to check this condition.

If $\beta = \gamma + 1$ for some γ and $y_\gamma \cap z \neq \emptyset$ for all $z \in F_\gamma$, let $F_\beta = F_\gamma \bigcup \{u : \exists z \in F_\gamma \ u \supseteq y_\gamma \cap z\}$. Again, the reader can easily check that conditions (2) through (5) hold.

Suppose $\beta = \gamma + 1$ for some γ and $y_\gamma \cap z' = \emptyset$ for some $z' \in F_\gamma$. Let $w = x \setminus y_\gamma$. Then $w \supseteq z'$, so $w \in F_\gamma$. Define $F_\beta = F_\gamma$.

Method II, from ZL: Let F be a proper filter on x. Let $\mathcal{F} = \{G : G$ is a proper filter on x, $F \subseteq G\}$. Partially order $\mathcal{F}$ by extension, i.e., $\subseteq$.

Suppose $\mathcal{C}$ is a nonempty chain in $\mathcal{F}$. We show that $\bigcup \mathcal{C} \in \mathcal{F}$. Clearly $F \subseteq \bigcup \mathcal{C}$. Suppose $y \in G \in \mathcal{C}$ and $z \supseteq y$. Since G is a filter, $z \in G$, so $z \in \bigcup \mathcal{C}$. Suppose $y_0 \in G_0, ...y_n \in G_n$ where each $G_i \in \mathcal{C}$. Without loss of generality, each $G_i \subseteq G_{i+1}$. So each $y_i \in G_n$. Hence $y_0 \cap ... \cap y_n \in G_n \subset \bigcup \mathcal{C}$.

Clearly $\bigcup \mathcal{C}$ is an upper bound for each $G \in \mathcal{C}$, so the hypothesis of ZL is met.

Hence $\mathcal{F}$ has a maximal element, G. Suppose there is some $y \subseteq x$ with $y \notin G$ and $x \setminus y \notin G$. Either $y \cap z \neq \emptyset$ for all $z \in G$, or there is $z \in G$ with $y \cap z = \emptyset$. In the former case, $G \cup \{y \cap z : z \in G\}$ is a filter extending G, a contradiction. In the latter case, $x \setminus y \in G$, a contradiction. So G is an ultrafilter.

Method III, from HMP: Let $\mathcal{F}$ be as in method II, and let $\mathcal{C}$ be a maximal chain in $\mathcal{F}$. By previous methods, $\bigcup \mathcal{C}$ is an ultrafilter. □

The reader might well ask: are there proofs of theorem 4.31 and theorem 4.32 that don't use AC? The answer is no. The statement that every linear order has a well-ordered cofinal subset, and the statement that every filter on a set is contained in an ultrafilter, each imply at least a weak version of AC.

Finally, we reprove the interesting part of theorem 4.2, officially using AC (actually HMP):

Theorem 4.33. *If there are no infinite descending $\in$-chains, then every set has an $\in$-minimal element.*

Proof. Suppose x has no $\in$-minimal element. Let $b = \{y \in x : y \cap x$ has no $\in$-minimal element$\}$. Note that $b \neq \emptyset$. Let $\mathcal{C}$ be the set of finite descending $\in$-chains in b where the order is $C \leq D$ iff D is an end-extension of C, i.e., $D = \{y_0, ...y_n\}$ where each $y_{i+1} \in y_i$, and $C = \{y_0, ...y_m\}$ where $m \leq n$.

By HMP, $\mathcal{C}$ has a maximal chain. If this chain were finite, it would have a last element $D = \{y_0, ...y_n\}$ where $y_n \in b$, hence $\mathcal{C}$ would not be maximal. So $\mathcal{C}$ is infinite. But then $\bigcup \mathcal{C}$ is an infinite descending $\in$-chain. □

4.6.1 Models of regularity and choice

Models of regularity are easy.

Theorem 4.34. $\forall X \ X \models$ *regularity.*

Proof. Given $x \in X$ with $x \cap X \neq \emptyset$, let $y \in x \cap X$ be minimal with respect to $\in$. Then $X \models \ \forall z \in x \ z \notin y$, i.e., $X \models \ x$ has an element minimal with respect to $\in$. □

Note that in proving theorem 4.34 we rather blatantly assumed that regularity is true in the universe.

Models of choice are trickier: we've got an index set I in the model, a family of sets $\{x_i : i \in I\}$ in the model (not all of which need be in the model), and those x_i in the model to keep track of. To make things reasonable, we'll restrict ourselves to transitive sets.

Theorem 4.35. *Let X be transitive. $X \models$ choice iff if $I \in X$ and $\{x_i : i \in I\} \in X$ where each $x_i \neq \emptyset$ then there is $f \in X$ with dom $f = I$ and $f(i) \in x_i$ for all $i \in I$.*

The proof is left to the reader. More useful is

Theorem 4.36. *Let X be transitive and closed under power set, union, and finite product (i.e., if $x \in X$ then $\mathcal{P}(x) \in X$ and $\bigcup x \in X$, and if $x, y \in X$ then $x \times y \in X$). Then $X \models$ choice.*

Proof. Suppose $I \in X, \{x_i : i \in I\} \in X$. Then $I \times \bigcup\{x_i : i \in I\} \in X$. By transitivity, $\bigcup\{x_i : i \in I\}^I \subseteq X$. By AC, $\exists f \in \{x_i : i \in I\}^I$ with $f(i) \in x_i$ for all $i \in I$. By assumption, $\mathcal{P}(I \times \bigcup\{x_i : i \in I\}) \in X$. Hence $f \in X$. □

Note that the proof of theorem 4.36 needs only weak closure under power set: if $x \in X$ then $\mathcal{P}(X) \subseteq X$.

A word of caution: theorem 4.35 and theorem 4.36 only provide models of one version of each axiom. E.g., without the other axioms of ZFC a model of AC need not be a model of WO.

4.7 Embedding mathematics into set theory

We have claimed that all of mathematics can be done in the language of set theory, i.e., that every mathematical statement can be translated into a formula whose only nonlogical symbol is $\in$.[57]

Our chapter on theories and models showed how to axiomatize classes of structures such as partial orders and groups; all that's needed to embed this into set theory is to note that relations and functions are set-theoretic objects. So any class of structures which is defined by a first-order theory is embedded into set theory. Since, in set theory, we can talk about substructures, we not only have, e.g., the first order theory of groups, we also have full group theory.

But not everything that mathematicians care about is first order. There are canonical structures that we care about: $\mathbb{N}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, along with the whole panoplies of number theory, analysis, and so on, that grow out of these structures. We have already represented $\mathbb{N}$, along with its arithmetic and order structure, within set theory. The purpose of this section is to represent $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ along with their arithmetic and order structures, and show that $\mathbb{R}$ has the desired geometric property of "no holes."

Fifty years ago, when the impact of modern set theory on the rest of mathematics was just becoming noticeable, this material was standard in foundations of mathematics courses, where it was taught in some detail in lieu of more set theory. Now, when there is so much set theory to learn, this older material is often left out entirely, making the root of set theory's profound impact on mathematics obscure. This section attempts a middle ground. To do all of this material in full

[57] with the caveat that category theory, which relies on classes, needs special treatment.

detail would be tedious and would provide only the insight that it can be done. But this insight is itself a revolutionary one, and the reader deserves a demonstration of its reasonableness. That is the purpose of this chapter. Nearly all proofs are left to the exercises.

Since $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are distinct structures, we will refer to, e.g., $+_{\mathbb{N}}$ or $\leq_{\mathbb{Q}}$ to make the structure we're working in clear. Recall that $0_{\mathbb{N}}, 1_{\mathbb{N}}, +_{\mathbb{N}}, \times_{\mathbb{N}}$, and $\leq_{\mathbb{N}}$ have already been defined, either in the text or in the exercises.

4.7.1 $\mathbb{Z}$

A first instinct might be to represent the elements of $\mathbb{Z}$ by ordered pairs (n, m) where $n, m \in \omega$ and (n, m) represents the integer $n - m$. But then the representation of integers is not unique: is -2 represented by (0,2), (1,3), (57, 59)...? So instead we use equivalence classes of ordered pairs of natural numbers: $(n, m) \equiv (n', m')$ iff $n +_{\mathbb{N}} m' = n' +_{\mathbb{N}} m$. $\mathbb{Z} = \{[(n, m)]_{\equiv} : n, m \in \omega\}$.

Here is the definition of $0_{\mathbb{Z}} : 0_{\mathbb{Z}} = [(0, 0)]$.

Here is the definition of $+_{\mathbb{Z}} : [(n, m)] +_{\mathbb{Z}} (n', m')] = [(n +_{\mathbb{N}} n', m +_{\mathbb{N}} m')]$.

Here is the definition of $\times_{\mathbb{Z}} : [(n, m)] \times_{\mathbb{Z}} (n', m')] = [((n \times_{\mathbb{N}} n') +_{\mathbb{N}} (m \times_{\mathbb{N}} m'), (m \times_{\mathbb{N}} n') +_{\mathbb{N}} (m \times_{\mathbb{N}} m'))]$.

The reader is invited to define $1_{\mathbb{Z}}$ and $[(n, m)] \leq_{\mathbb{Z}} [(n', m')]$ in the exercises.

There are two kinds of questions about these definitions: are they well-defined? and do they do what they are supposed to do? I.e., -2 is represented by both [(1,3)] and [(57, 59)], while 5 is represented by both [(8,3)] and [(100,95)]. Does [(1,3)] $+_{\mathbb{Z}}$ [(100,95)] give the same equivalence class as [(57, 59)]$+_{\mathbb{Z}}$ [(8,3)]? And is this the equivalence class that represents the integer 3? The reader will be asked to check this kind of thing in the exercises.

4.7.2 $\mathbb{Q}$

Again we need equivalence classes of ordered pairs. This time the ordered pairs are in $\mathbb{Z}^2$, and $[(a, b)]$ represents $\frac{a}{b}$. So let $a, a', b, b' \in \mathbb{Z}$. We say $(a, b) \equiv (a', b')$ iff $a \times_{\mathbb{Z}} b' = a' \times_{\mathbb{Z}} b$. $\mathbb{Q} = \{[(a, b)] : a, b \in \mathbb{Z}$ and $b \neq 0_{\mathbb{Z}}\}$.

Here is the definition of $0_{\mathbb{Q}} : 0_{\mathbb{Q}} = [(0_{\mathbb{Z}}, 1_{\mathbb{Z}})]$.

Here is the definition of $+_{\mathbb{Q}} : [(a, b)] +_{\mathbb{Q}} [(a', b')] = [(((a \times_{\mathbb{Z}} b') +_{\mathbb{Z}} (a' \times_{\mathbb{Z}} b)), b \times_{\mathbb{Z}} b')]$.

Here is the definition of $\leq_{\mathbb{Q}}$: $[(a, b)] \leq_{\mathbb{Q}} [(a', b')]$ iff $a \times_{\mathbb{Z}} b' \leq_{\mathbb{Z}} a' \times_{\mathbb{Z}} b$.

The rest is left to the reader.

Note the progress of 0: $0_{\mathbb{N}} = \emptyset$. $0_{\mathbb{Z}} = [(\emptyset, \emptyset)]$. $0_{\mathbb{Q}} = [[(\emptyset, \emptyset)], [(\{\emptyset\}, \emptyset)]]$.

4.7.3 $\mathbb{R}$

It was relatively easy to extend $\mathbb{N}$ to $\mathbb{Z}$ to $\mathbb{Q}$, since each is an algebraic extension of the previous. But $\mathbb{R}$ is different. It adds elements that no simple equation with coefficients in $\mathbb{Q}$ can acount for. And it has a crucial geometric property as an order that cannot be derived from its algebraic

structure: $\mathbb{R}$ is a continuum, that is, it has no holes. This property of having no holes is formalized as the least upper bound property: every bounded subset of $\mathbb{R}$ has a least upper bound. How can we represent $\mathbb{R}$ in terms of $\mathbb{Q}$ and still capture our intuition of $\mathbb{R}$'s geometric completeness? This was an important question in nineteenth century mathematics, one of the keys to making the notions of calculus precise. More than one formal solution was offered. The one which is easiest to work with (all of the solutions are provably equivalent) is the method of Dedikind cuts.

Definition 4.37. A Dedekind cut is a proper initial segment of $\mathbb{Q}$ with no largest element.

Note that a bounded union of Dedekind cuts is a Dedekind cut, and an unbounded union of Dedekind cuts is $\mathbb{Q}$.

Definition 4.38. $\mathbb{R}$ is the set of Dedekind cuts.

The instinct here is that $\sqrt{2}$ is represented by $\{q \in \mathbb{Q} : q < \sqrt{2}\}$. The requirement that a Dedekind cut has no largest element is to ensure that each real number has only one representative: $\{q \in \mathbb{Q} : q < 17\} \in \mathbb{R}$ but $\{q \in \mathbb{Q} : q \leq 17\} \notin \mathbb{R}$.

Here is the definition of $+_{\mathbb{R}} : x +_{\mathbb{R}} y = \{a +_{\mathbb{Q}} b : a \in x \text{ and } b \in y\}$.

Here is the definition of $\leq_{\mathbb{R}}$: $x \leq_{\mathbb{R}} y$ iff $x \subseteq y$.

The definition of $\times_{\mathbb{R}}$ is left to the exercises.

We show that $\mathbb{R}$ has no holes.

Theorem 4.39. *A set of Dedekind cuts with an upper bound has a least upper bound.*

Proof. Let B be a set of Dedekind cuts with an upper bound. $\bigcup B$ is a Dedekind cut, and $\bigcup B \neq \mathbb{Q}$. So $\bigcup B$ is also an upper bound for B. We show $\bigcup B$ is a least upper bound: suppose x is also an upper bound for B. I.e., $x \supseteq y$ for all $y \in B$. Hence $x \supseteq \bigcup B$. □

Once we have $\mathbb{R}$ we have pretty much everything, since all areas of mathematics either are about models of an axiomatizable system; or are derived from the standard structures $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ (left for the exercises), or can be expressed directly in set theory (such as some areas of topology and of combinatorics).

4.8 Exercises

1. Show the following
   (a) If x is transitive, so are $S(x)$ and $x \cup \mathcal{P}(x)$.
   (b) If x_i is transitive for all $i \in I$ so is $\bigcup_{i \in I} x_i$.
   (c) If x, y are transitive, so is $x \cap y$.
   (d) There are transitive x, y with $x \setminus y$ not transitive.
2. (a) What is the transitive closure of $\{0, 3, \{5, 7\}\}$?
   (b) What is the transitive closure of $\{x_n : n \in \omega\}$ where $x_0 = \emptyset$ and each $x_{n+1} = \{x_n\}$?

3. Show that $\emptyset \in x$ for every transitive x.

4. Show that if x is transitive, so is $\mathcal{P}(x)$.

5. Show that each V_α is transitive; if $\alpha < \beta$ then $V_\alpha \in V_\beta$; and that if $\alpha < \beta$ then $V_\alpha \subseteq V_\beta$.

6. What is V_3? V_4?

7. Prove proposition 4.21.

8. If rank $x = \alpha \leq$ rank $y = \beta$,

   (a) what's rank (x, y)?

   (b) what's rank $x \cup y$?

9. (a) Prove that if rank $x = \alpha$ then rank $\bigcup x \leq \alpha$.

   (b) Find x so rank $x = \alpha$ and rank $\bigcup x < \alpha$.

10. Show that if a collection of chains in a partial order is linearly ordered by $\subseteq$, then its union is a chain. (See the proof of theorem 4.30,)

11. In the proof of theorem 4.30, what other axioms of set theory did we use, and where did we use them?

12. Complete the proof of theorem 4.32.

13. Which axioms are modeled by V_ω?

14. Suppose X satisfies the hypothesis of theorem 4.36.

   (a) Which axioms does X model?

   (b) Which versions of AC hold in X? [Hint: see problem 12]

15. A *selector* for $\{x_{\cdot} i \in I\}$ is a 1-1 choice function in $\Pi_{i \in I} x_i$.

   (a) Show that if, for each $n < \omega$, x_n has at least $n + 1$ elements, then $\{x_n : n \in \omega\}$ has a selector.

   (b) Let $\delta \in ON$. Show that if, for each $\alpha < \delta, \bigcup_{\beta < \alpha} x_\beta \not\supseteq X_\alpha$, then $\{x_\alpha : \alpha < \delta\}$ has a selector.

   (c) Did you use AC?

16. Without the axiom of choice, show that if $x, I \neq \emptyset$ then $x^I \neq \emptyset$.

17. Show that every vector space has a basis, where a basis is a maximal linearly independent set.

18. Let $\mathbb{F}$ = the set of polynomials with coefficients in $\mathbb{R}$. Consider $\mathbb{F}$ as a vector space over $\mathbb{R}$. Without using AC, show that it has a basis.

19. Show that every partial order contains a maximal incompatible set (called a maximal antichain).

20. Show that every family of sets contains a maximal centered subset (i.e., a maximal filterbase).

21. A family of sets is linked iff any two members of the family have nonempty intersection. Show that every family of non-empty sets has a maximal linked subfamily.

22. A subset of a partial order is linked iff any pair of elements has a lower bound, and centered iff

any finite set of elements has a lower bound.

23. (a) Show that every partial order has a maximal linked family and a maximal centered family.

    (b) Find a partial order with a linked family that is not centered.

24. Prove the following version of theorem 1.43: Let $\{x_n : n \in \omega\}$ be a sequence of infinite sets so that each $x_n \supseteq x_{n+1}$. Show there is an infinite set x so $x \setminus x_n$ is finite for each n.

25. Define $1_{\mathbb{Z}}$ and $\leq_{\mathbb{Z}}$.

26. a) Show that $+_{\mathbb{Z}}$ is well-defined.

    (b) Show that $\times_{\mathbb{Z}}$ is well-defined.

    (c) Show that $+_{\mathbb{Z}}$ is the correct definition.[58]

    (d) Show that $\times_{\mathbb{Z}}$ is the correct definition.[59]

27. Define $\times_{\mathbb{Q}}$.

28. Show that $+_{\mathbb{Q}}$ is well-defined.

29. Define $\times_{\mathbb{R}}$ and show that it is the correct definition.[60]

30. (a) Define the set of complex numbers $\mathbb{C}$.

    (b) Define $+_{\mathbb{C}}$, and $\times_{\mathbb{C}}$

[58] Note: to do this you'll need to go outside set theory to our usual understanding of $\mathbb{Z}$.
[59] See the previous footnote.
[60] Note: to show it's the right definition you'll need to go outside set theory to our usual understanding of $\mathbb{R}$.

5 Infinite numbers

Before Zermelo, before Fraenkel, before Frege defined the number zero in nearly ninety closely argued pages, there was Cantor.[61]

Cantor's achievement was monumental. He made the study of infinity precise. He discovered infinite ordinals. He discovered the variety of infinite cardinals. He invented three of the most useful tools in mathematical logic: dove-tailing (the ancestor of time-sharing), diagonalization (which he used to show that there are more reals than natural numbers, and which Gödel used to prove the first incompleteness theorem), and the back-and-forth argument (used in model theory). His work went so counter to ideas predating Aristotle — infinity is out there where you cannot get at it, there is only only one infinity, and nothing infinite can truly be said to exist anyway — that he was bitterly fought by philosophers, mathematicians, and even theologians (who associated infinity with God and hence saw Cantor as sacrilegious). But even his greatest critics had to admit that the only way to seriously deny the validity of his work, so rooted as it is in mathematical practice, is to deny even that $\mathbb{N}$ is a set, i.e., to insist that only finite objects exist, as Aristotle had. Cantor's great enemy Kronecker tried to convince other mathematicians that the infinite could not be spoken of, but the effort failed. In 1925 David Hilbert acknowledged the general acceptance of Cantorian infinity by saying "no one can drive us from the paradise that Cantor created for us."

We have already talked about infinite ordinals a bit. In this chapter we give some basic facts about cardinality and then do some ordinal arithmetic, and some more advanced work with cardinals. For the first two sections of this chapter we will point out all uses of AC. After that we simply assume that we are working in full ZFC — otherwise there is little that can be proved.

5.1 Cardinality

Rather than defining a cardinal number (a task fraught with difficulties) we will define cardinality. We write $|x| \leq |y|$ for "the cardinality of x is no bigger than the cardinality of y", and $|x| = |y|$ for "x and y have the same cardinality." We define these forthwith:

Definition 5.1. $|x| \leq |y|$ iff there is a 1-1 function $f : x \to y$. $|x| = |y|$ iff there is a 1-1 function from x onto y; $|x| < |y|$ iff $|x| \leq |y|$ and $|x| \neq |y|$.

In particular, if $x \subseteq y$ then $|x| \leq |y|$ (hence each $|x| \leq |x|$), and if $|x| \leq |y| \leq |z|$ then $|x| \leq |z|$.

It's important to note that, without AC, you may have two sets whose sizes cannot be compared. I.e., there may be x, y so $|x| \not\leq |y|$ and $|y| \not\leq |x|$.

Definition 5.1 defines a relation between x and y. It does not define an object called "the cardinality of x." This phrase can be considered just a quirk of grammar,[62] a kind of shorthand. There is no need to reify it[63].

Why is definition 5.1 reasonable? We have a room with people and chairs. Suppose everyone in the room is sitting down. If there are some chairs left over, we know without counting that there

[61] Cantor's pioneering work was done in the 1870's and 1880's; Frege's *Die Grundlagen der Arithmetik* was written at the turn of the century; Zermelo's seminal paper was in 1908.

[62] or it can be taken seriously, but there is no mathematical necessity to doing that.

[63] although assuming AC we can and will.

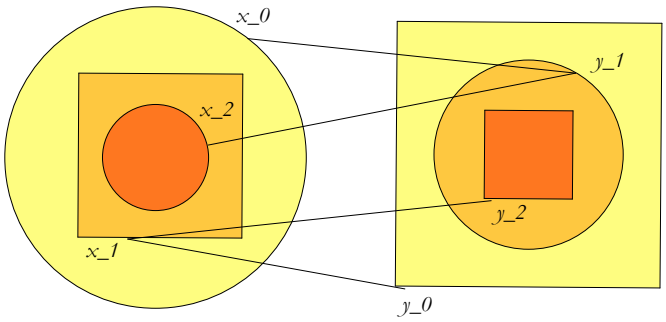
are more chairs than people. If there are no empty chairs, we know without counting that there are the same number of chairs and people. And if, instead of everyone being seated, every chair is full and there are some people left standing, we know without counting that there are more people than chairs.[64] The chairs can be in rows, in a single line, in a circle, arranged as in a concert hall — it doesn't matter. We do not care about the arrangement, only whether every person has a chair or every chair a person.

Suppose for every x, we had in fact defined a set $|x|$ so that definition 5.1 became a theorem, rather than a definition.[65] Then the relation $\leq$ would be reflexive and transitive. In this loose way of speaking the next theorem shows antisymmetry.

Theorem 5.2. The Schröder-Bernstein theorem *If* $|x| \leq |y|$ *and* $|y| \leq |x|$ *then* $|x| = |y|$.

Proof. We are given a 1-1 function $f : x \to y$ and a 1-1 function $g : y \to x$. We must find a 1-1 function h from x onto y.

We start by letting $x_0 = x, y_0 = y$. By induction, $x_{n+1} = g[y_n]$ and $y_{n+1} = f[x_n]$. I.e., in the diagram below, the square y_0 is the preimage of the square x_1 which is the preimage of the square y_2 which is... and the circle x_0 is the preimage of the circle y_1 which is the preimage of the circle x_2 which is ...



Clearly

Subclaim 5.2.1. Each $f|_{x_n}$ is 1-1 onto y_{n+1} and each $g|_{y_n}$ is 1-1 onto x_{n+1}.

Subclaim 5.2.2. Each $x_{n+1} \subseteq x_n$ and each $y_{n+1} \subseteq y_n$.

Proof. By induction. Clearly $x_1 \subseteq x_0$ and $y_1 \subseteq y_0$. Suppose $x_n \subseteq x_{n-1}$ and $y_n \subseteq y_{n-1}$. Then $x_{n+1} = g[y_n] \subseteq g[y_{n-1}] = x_n$ and $y_{n+1} = f[x_n] \subseteq f[x_{n-1}] = y_n$. □

Define $x^*_n = x_n \setminus x_{n+1}$ and $y^*_n = y_n \setminus y_{n+1}$. Define $x^\dagger = \bigcap_{n<\omega} x_n$ and $y^\dagger = \bigcap_{n<\omega} y_n$. By definition,

[64] This assumes, of course, that at most one person is sitting in each chair.
[65] We can and will do this under AC.

Subclaim 5.2.3. $x = x^\dagger \cup \bigcup_{n<\omega} x^*_n$, $y = y^\dagger \cup \bigcup_{n<\omega} y^*_n$, $\{x^\dagger\} \cup \{x^*_n : n < \omega\}$ is pairwise disjoint, and $\{y^\dagger\} \cup \{y^*_n : n < \omega\}$ is pairwise disjoint.

Subclaim 5.2.4. $f|_{x^*_n}$ is 1-1 onto y^*_{n+1}; $g|_{y^*_n}$ is 1-1 onto x^*_{n+1}; $f|_{x^\dagger}$ is 1-1 onto $y^\dagger$; and $g|_{y^\dagger}$ is 1-1 onto $f^\dagger$.

Proof. By 1-1, $f[x^*_n] = f[x_n \setminus x_{n+1}] = f[x_n] \setminus f[x_{n+1}] = y_{n+1} \setminus y_{n+2} = y^*_{n+1}$. A similar proof works for $g[x^*_n]$. By definition, $y^\dagger = f[x^\dagger]$ and $x^\dagger = g[y^\dagger]$. □

Now we construct h. $h|_{x^\dagger} = f$. $h|_{x^*_{2n}} = f|_{x^*_{2n}}$. $h|_{x^*_{2n+1}} = g^{\leftarrow}[x^*_{2n+1}]$.

By the subclaims, $h : x \to y$ is 1-1 onto.

□

The complexity of this proof is because we are not using AC. Under AC, theorem 5.2 is trivial, and we'll show this later.

The Schröder-Bernstein theorem justifies the use of the word "size" when we discuss cardinality. It says that, modulo the caveat about reification, $\leq$ is a partial order. Later we'll show that, under AC, it is a linear order (no caveats).

The next proposition captures two intuitions about size: (a) if the size of one set is less than or equal to the size of another set, then taking one element away from each preserves the relative sizes; and (b) a small set can't be sent onto a big set. Interestingly, this second, very basic, intuition needs AC for its proof.

Proposition 5.3. *(a) If $|x| \leq |y|, a \in x$, and $b \in y$, then $|x \setminus \{a\}| \leq |y \setminus \{b\}|$.*

(b) Assume AC. If $\exists f : x \to y$ with f onto, then $|x| \geq |y|$.

Proof. For (a): Suppose $f : x \to y$ is 1-1. Let $c = f^{\leftarrow}(b)$ and $d = f(a)$. Define $f^* : x \setminus \{a\} \to y \setminus \{b\}$ as follows: If $z \neq c$ then $f^*(z) = f(z)$; if $c \neq a$ then $f^*(c) = d$. f^* is 1-1.

For (b): Let $f : x \to y$, f onto. For all $b \in y$ let $x_b = f^{\leftarrow}(b)$. Let $h \in \Pi_{b \in y} x_b$. $h : y \to x$ is 1-1. □

Definition 5.4. x is finite iff there is $n \in \omega$ with $|x| = |n|$. Otherwise x is infinite.

A property of infinite sets is that they can have the same size as proper subsets. For example, $|\omega| = |\{2n : n \in \omega\}|$, via the map which takes n to $2n$.[66] Finite sets don't have this property. Under AC, this property is equivalent to definition 5.4.

Proposition 5.5. *Assume AC. x is finite, iff $|x| \neq |y|$ for any $y \subset x$.*

Proof. For necessity, it suffices to show that if $m < n$ then $|m| < |n|$. This proof is by induction and does not use AC. So suppose we know that $\forall i < j \leq n\ |i| < |j|$. Let $x = n \setminus \{0\}, y = n+1 \setminus \{0\}$. $|x| = |n-1|, |y| = |n|$, so by hypothesis $|x| < |y|$. Hence, by proposition 5.3(a), $|n| < |n+1|$.

For the other direction, suppose x is not finite. Then, by WO, $x = \{a_\alpha : \alpha < \delta\}$ where $\delta \geq \omega$. Let $y = x \setminus \{a_{2n} : n \in \omega\}$. Then $|y| = |x|$. □

[66] And then there's the Hilbert Hotel, with infinite many rooms, which can always squeeze in infinitely more guests...

A set satisfying "$|x| \neq |y|$ for any $y \subset x$" is called Dedekind-finite. Without AC there can be Dedekind-finite sets which are not finite.

Cardinalities can be classified in many ways. Our first, crude, attempt, was to classify them as finite or infinite. Our second, slightly more subtle, attempt is

Definition 5.6. x is countable iff $|x| \leq |\omega|$; otherwise it is uncountable. x is denumerable iff $|x| = |\omega|$.

Cantor's two great early theorems on cardinality were that the rationals are countable but the reals are not.

Theorem 5.7. $\mathbb{Q}$ *is countable.*

Proof. We give two proofs, the first quick, and the second (due to Cantor) insightful.

Proof I. To each $q \in \mathbb{Q}$ we associate the unique pair (n_q, m_q) $q = \frac{n_q}{m_q}$, $m_q > 0$, and m_q is the smallest possible such denominator. Let r, s, t be distinct prime numbers, and define

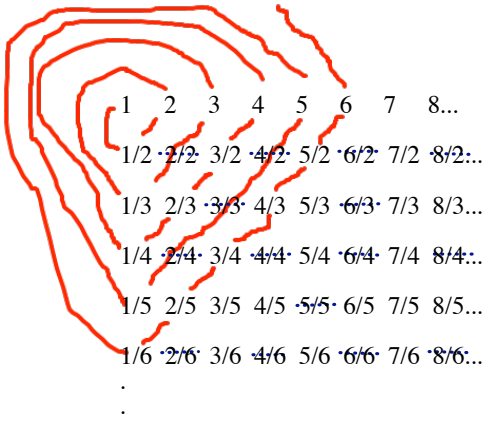
$$f(q) = \begin{cases} r^{n_q} s^{m_q} & \text{if } q \geq 0 \\ t^{n_q} s^{m_q} & \text{if } q < 0 \end{cases}$$

$f : \mathbb{Q} \to \omega$ is 1-1, so $|\mathbb{Q}| \leq |\omega|$. By the Schröder-Bernstein theorem, $|\mathbb{Q}| = |\omega|$.

Proof II. When Cantor proved theorem 5.7, there was no definition of cardinality; that came later, due to Frege, essentially making explicit what Cantor had left implicit. What Cantor proved, in modern terminology, was the existence of a 1-1 function from $\mathbb{N}$ to $\mathbb{Q}$.

So it's no surprise that his proof was not as slick as the one above. Slickness, however, does not equal insight. The technique of the proof we'll give below, in which there are infinitely many tasks, and you shuttle back and forth doing a little on this one, a little on that one, a little more on the first one, etc., is known as dove-tailing, and is one of the basic techniques in computer science.

First, note that $|\mathbb{Q}| = |\{q \in \mathbb{Q} : q > 0\}|$. Then list all quotients of positive integers so the n^{th} row consists of all $\frac{m}{n}$, crossing out all expressions not in lowest terms.



Finally, we list according to the antidiagonals, as in the diagram, so $f(0) = 1, f(1) = \frac{1}{2}, f(2) = 2, f(3) = \frac{1}{3}, f(4) = 3$ (because we crossed out $\frac{2}{2}$), $f(5) = \frac{1}{4}$, and so on.

More formally, we recursively describe f: $f(0) = 1$. Suppose we know $f|_{n+1}$ and $f(n) = \frac{i}{k}$ where $i + k = m$. Let $j > 0, j < k$ be least so $\forall s < n\ f(s) \neq \frac{n+j}{k-j}$. If just a j exists, $f(n+1) = \frac{i+j}{k-j}$. Otherwise $f(n+1) = \frac{1}{m}$.

□

As a corollary to the method of proof we have

Corollary 5.8. *A countable union of countable sets is countable.*

Proof. $\forall n < \omega$ let x_n be countable. We may assume that the x_n's are pairwise disjoint (if necessary, by replacing x_n by $x_n \setminus \bigcup_{i<n} x_i$). Let $f_n : x_n \to \{\frac{m}{n+1} : m \in \mathbb{N}; m \neq 0; m, n+1$ have no common divisors$\}$. Let $x = \bigcup_{n<\omega} x_n$ and let $g = \bigcup_{n<\omega} f_n$. Let f be as in the proof of theorem 5.7. Let $h = f \circ g$. Then $h : x \to \mathbb{N}$, h is 1-1.

□

The generalization of corollary 5.8 is the statement that if Y is infinite, and each $|y| = |Y|$ for $y \in Y$, then $|\bigcup Y| = |Y|$. This statement is not a theorem of ZF. There are models of ZF + ¬AC in which it fails.

Now for Cantor's second great theorem on cardinality, that $\mathbb{R}$ is uncountable. His first proof was analytic and did not generalize. We give his second proof, whose method is called diagonalization. Diagonalization is a powerful technique in set theory and logic — it was used by Gödel to prove the first incompleteness theorem, and is frequently invoked in recursion theory and computer science. We state the generalized theorem, and then show how it applies to $\mathbb{R}$.

Theorem 5.9. $\forall x\ |x| < |\mathcal{P}(x)|$.

Proof. It's trivial to show that $|x| \leq |\mathcal{P}(x)|$: just let $f(y) = \{y\}$. We need to show that $|x| \neq |\mathcal{P}(x)|$.

Suppose $f : x \to \mathcal{P}(x)$. We must show that f is not onto. So construct a set $y \subseteq x$ with $y \notin$ range f as follows: $\forall z \in x\ z \in y$ iff $z \notin f(z)$. If, for some $z \in x$, $y = f(z)$, then we essentially have Russell's paradox: $z \in f(z)$ iff $z \notin f(z)$.[67] □

We apply this to $\mathbb{R}$.

Theorem 5.10. $|\mathbb{R}| = |\mathcal{P}(\omega)|$.

Proof. To show that $|\mathbb{R}| \geq |\mathcal{P}(\omega)|$, let $C = \{x \in \mathbb{R} : 0 \leq x < 1$ and the ternary expansion of x contains no 2's$\}$.[68] Define $f : C \to \mathcal{P}(\omega)$ as follows: $f(x) = \{n :$ the n^{th} digit in the ternary expansion of x is 1$\}$. The reader is asked in the exercises to show that f is 1-1 and onto. Since $C \subseteq \mathbb{R}$, $|\mathcal{P}(\omega)| \leq |\mathbb{R}|$.[69]

For the other direction, for each $x \in \mathbb{R}$ define $D_x = \{q \in \mathbb{Q} : q < x\}$.[70] Define $g : \mathbb{R} \to \mathcal{P}(\mathbb{Q})$ by: $g(x) = D_x$. g is 1-1. Since $|\mathcal{P}(\mathbb{Q})| = |\mathcal{P}(\omega)|$, $|\mathbb{R}| \leq |\mathcal{P}(\omega)|$.[71] □

[67] This proof is brief, but ingenious, even revolutionary, as its mixed reception — consult the historical literature — makes clear.

[68] This is almost the famous Cantor set — to get the Cantor set, include 1.

[69] This is all that's needed to show $\mathbb{R}$ uncountable, but the second half of the proof is useful.

[70] i.e., the Dedekind cut that defines x in section 4.7.

[71] This assumes that if $|x| = |y|$ then $|\mathcal{P}(x)| = |\mathcal{P}(y)|$, which you are asked to prove in the exercises.

Hence, as a corollary to theorem 5.9, $\mathbb{R}$ is uncountable.

And as an immediate corollary to theorem 5.9

Corollary 5.11. *There is no biggest set:* $\forall x\ \exists y\ |x| < |y|$.

To finish what we will do without AC

Proposition 5.12. $\forall x\ |\mathcal{P}(x)| = |2^x|$.

Proof. We define $\chi : \mathcal{P}(x) \to 2^x$ as follows: for $y \subseteq x$, define $\chi_y : x \to 2$ by: $\chi_y(z) = 0$ iff $z \in y$. (χ_y is called the characteristic function of y.) Define $\chi(y) = \chi_y$. χ is easily seen to be 1-1 and onto. □

5.2 Cardinality with choice

It's time to face the question "what is a cardinal number?" The reply of the early set theorists was "$|x|$ is the class of all y with $|y| = |x|$." This is highly unsatisfactory, since numbers should be canonical objects, i.e., sets, and not classes. Furthermore, a number should be in some sense a single object, not a bunch of things equivalent to each other.[72] The notion of infinite number should extend, in some way, the notion of finite number. But without AC there is no good solution.

With AC things are different. Finite cardinals were ordinals; so infinite cardinals should be ordinals. But not every ordinal can be a cardinal. After all, $|\omega| = |\omega + 1|$. What ordinal, out of all the denumerable ordinals (and there are uncountably many) should be pick to call a cardinal number? The answer is not surprising.

Definition 5.13. An ordinal κ is an initial ordinal iff $\forall \alpha < \kappa\ |\alpha| < |\kappa|$.

Note that definition 5.13 does not need AC. Even without AC, some ordinals will be initial ordinals and others won't.

By definition 5.13, the following is immediate:

Proposition 5.14. $\forall \alpha$ *an ordinal,* $\exists \kappa$ *an initial ordinal with* $|\alpha| = |\kappa|$.

For example, ω is an initial ordinal, and we can define the first uncountable ordinal, ω_1, which by definition is an initial ordinal. And we can define the first ordinal whose cardinality is bigger than ω_1 (not surprising, called ω_2), and so on. (We will do this below.)

Neither definition 5.13 nor proposition 5.14 depend on AC. What does depend on AC is the following:

Theorem 5.15. *AC iff* $\forall x\ \exists \kappa\ \kappa$ *is an initial ordinal and* $|x| = |\kappa|$.

Proof. Assume AC. Then $\exists \alpha$ an ordinal, $|x| = |\alpha|$. Let κ be the least ordinal with $|\kappa| = |\alpha|$. Then κ is an initial ordinal and $|x| = |\kappa|$.

Assume AC fails. Then there is some x which cannot be well-ordered, i.e., $|x| \neq |\alpha|$ for all ordinals α. □

[72] even if it is reduced to a set — e.g., $\{y : y$ of minimal rank among all those y with $|y| = |x|\}$

Hence, under AC, we can define the cardinality of a set in a very nice way.

Definition 5.16. Assume AC. $\forall x\ |x| =$ the least ordinal κ with $|x| = |\kappa|$.

I.e., under AC, cardinals are initial ordinals. From now on, we will use κ, λ to denote cardinals.

The proof of the Schröder-Bernstein theorem is now trivial: If $|x| \leq |y| \leq |x|$ and $|x| = \kappa, |y| = \lambda$, then $\kappa \leq \lambda \leq \kappa$, so $\kappa = \lambda$.

Let's define some notation:

Definition 5.17. $\omega_0 = \omega$; given ω_α, we define $\omega_{\alpha+1}$ to be the least cardinal $> \omega_\alpha$; if α is a limit, then $\omega_\alpha = \bigcup_{\beta<\alpha} \omega_\beta = \sup\{\omega_\beta : \beta < \alpha\}$.

When its cardinal nature is being emphasized, or for typographical reasons (e.g., the awkwardness of "ω_{ω_n}") ω_α is sometimes written as $\aleph_\alpha$.[73] We will use both notations.

Sometimes we can avoid the cumbersome footnoted notation:

Definition 5.18. (a) If $\kappa = \aleph_\alpha$ and $\lambda = \aleph_{\alpha+1}$ we write $\lambda = \kappa^+$ and say that λ is a successor cardinal.

(b) A limit cardinal is a cardinal which is not a successor cardinal.

Definitions 5.17 and 5.18 did not depend on AC, but the next definition does.

Definition 5.19. (a) $2^\kappa = |2^\kappa|$, i.e., 2^κ is how we denote the cardinal of the set of all functions from κ to 2.[74]

(b) Assume AC. $\beth_0 = \omega$; given $\beth_\alpha, \beth_{\alpha+1} = 2^{\beth_\alpha}$; if α is a limit, $\beth_\alpha = \bigcup_{\beta<\alpha} \beth_\beta$.[75]

It need not be the case, even with AC, that every $|\mathcal{P}(x)|$ is some $\beth_\alpha$. It is consistent, for example, to have $\beth_1 = \omega_4, \beth_2 = 2^{\beth_1} = \omega_{42}$, but $2^{\omega_1} = \omega_{17}$.

[73] $\aleph$ is the Hebrew letter pronounced "aleph."

[74] This is why some authors use ${}^\kappa 2$ to denote the set of all functions from κ to 2; we prefer to use the more conventional notation and trust readers to parse any potential ambiguities.

[75] $\beth$ is the Hebrew letter pronounced "bet."

5.3 Ordinal arithmetic

Cardinals are ordinals, but cardinal arithmetic and ordinal arithmetic are distinct. For example, we will see that in ordinal arithmetic, $\omega+1 > 1+\omega = \omega$, but in cardinal arithmetic $\omega+1 = 1+\omega = \omega$. This section is about ordinal arithmetic; the next section about cardinal arithmetic.

Line up three tortoises. Then line up five more behind them. Eight tortoises are now lined up. This is how we define $\alpha + \beta$: a copy of α followed by a copy of β.

Line up three tortoises. Line up three more behind them. And another three behind the second three. And another three behind the third three. Three times four tortoises are now lined up. This is how we define $\alpha \cdot \beta$: α followed by α followed by α... β times.

The formal definitions are inductive.

Definition 5.20. For α, β ordinals: $\alpha + 0 = \alpha$; $\alpha + (\beta + 1) = (\alpha + \beta) + 1$; if β is a nonzero limit, then $\alpha + \beta = \bigcup_{\gamma<\beta}(\alpha + \gamma) = \sup\{\alpha + \gamma : \gamma < \beta\}$.

Definition 5.21. For α, β ordinals: $\alpha \cdot 0 = 0; \alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$; if β is a limit, then $\alpha \cdot \beta = \bigcup_{\gamma<\beta} \alpha \cdot \gamma = \sup\{\alpha \cdot \gamma : \gamma < \beta\}$.

For completeness we include

Definition 5.22. For α, β ordinals: $\alpha^0 = 1; \alpha^{\beta+1} = \alpha^\beta \cdot \alpha$; if β is a nonzero limit, then $\alpha^\beta = \bigcup_{\gamma<\beta} \alpha^\gamma = \sup\{\alpha^\gamma : \gamma < \beta\}$.[76]

Note that the non-limit clauses exactly parallel the definitions in $\mathbb{N}$.

It is immediate from the definitions that

Proposition 5.23. *Let α, β, γ be ordinals with $\beta < \gamma$. Then $\alpha + \beta < \alpha + \gamma; \alpha \cdot \beta < \alpha \cdot \gamma$; and $\alpha^\beta < \alpha^\gamma$.*

Example 5.24. Addition does not commute: $1 + \omega = \sup\{1 + n : n < \omega\} = \omega < \omega + 1$.

Example 5.25. Multiplication does not commute: $2 \cdot \omega = \sup\{2 \cdot n : n < \omega\} = \omega < \sup\{\omega + n : n < \omega\} = \omega + \omega = \omega \cdot 2$.

Example 5.26. Ordinal exponentiation $\neq$ cardinal exponentiation: in ordinal exponentiation, $2^\omega = \sup\{2^n : n < \omega\} = \omega$. But in cardinal exponentiation, $\omega < 2^\omega$.[77]

Commutativity does not hold in ordinal addition and multiplication, but associativity and left-distributivity do.

Theorem 5.27. *Let α, β, γ be ordinals.*

(a) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

(b) $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$.

(c) $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$.

[76] We will say very little about ordinal exponentiation.

[77] Yet a third ambiguity in the notation 2^κ.

Proof. We prove (a); (b) and (c) are left to the reader.

The proof of (a) is by induction on γ. The induction hypothesis is that, for all $\delta < \gamma$, $(\alpha+\beta)+\delta = \alpha + (\beta + \delta)$.

Case 1. $\gamma = \delta + 1$ for some δ. Then

$$(\alpha + \beta) + \gamma = (\alpha + \beta) + (\delta + 1) = ((\alpha + \beta) + \delta) + 1 = (\alpha + (\beta + \delta)) + 1 =$$

$$= \alpha + ((\beta + \delta) + 1) = \alpha + (\beta + (\delta + 1)) = \alpha + (\beta + \gamma)$$

Case 2. γ is a limit. First we need

Subclaim 5.27.1. Let α be an ordinal and A a set of ordinals. Then $\sup\{\alpha + \beta : \beta \in A\} = \alpha + \sup A$.

The proof of the subclaim is left to the exercises.

By subclaim 5.27.1

$$(\alpha + \beta) + \gamma = \sup\{(\alpha + \beta) + \delta : \delta < \gamma\} = \sup\{\alpha + (\beta + \delta) : \delta < \gamma)\} =$$

$$= \alpha + \sup\{\beta + \delta : \delta < \gamma)\} = \alpha + (\beta + \gamma)$$

□

Note that right distributivity fails for ordinal arithmetic. For example, $(\omega + 1) \cdot \omega = \omega \cdot \omega$. Why? $(\omega + 1) \cdot 2 = (\omega + 1) + (\omega + 1) = \omega + (1 + \omega) + 1 = \omega + \omega + 1 = \omega \cdot 2 + 1$. Similarly (or by induction) each $(\omega + 1) \cdot n = \omega \cdot n + 1$. So $(\omega + 1) \cdot \omega = \sup\{\omega \cdot n + 1 : n < \omega\} = \omega \cdot \omega$. But $(\omega \cdot \omega) + \omega > \omega \cdot \omega = (\omega + 1) \cdot \omega$.

Let's give concrete pictures of some infinite ordinals.

Definition 5.28. A set x ordered by $\leq$ is said to have order type α iff x is order-isomorphic to α.

Example 5.29. Consider the usual ordering on $\mathbb{Q}$.

(a) $\{\frac{m}{m+1} : m \in \mathbb{N}\}$ has order type ω.

(b) $\{n + \frac{m}{m+1} : n, m \in \mathbb{N}\}$ has order type $\omega \cdot \omega$.

Example 5.30. Fix ordinals α, β. Under the lexicographical ordering, the following sets have the following order types:

(a) $(\{0\} \times \alpha) \cup (\{1\} \times \beta)$ has type $\alpha + \beta$.

(b) $\beta \times \alpha$ has order type $\alpha \cdot \beta$.

(c) $\gamma \times \beta \times \alpha$ has order type $\alpha \cdot \beta \cdot \gamma$.

The assertions in example 5.30 are easily proved by induction.

Example 5.31. For $n, m, k \in \mathbb{N}$, define real numbers $x_n, x_{n,m}, x_{n,m,k}$ as follows:

$x_n = 0.i_1 i_2 i_3 ...$ where $i_1 = ... = i_n = 1$ and all other $i_j = 0$.

$x_{n,m} = 0.i_1 i_2 i_3 ...$ where $i_1 = ... = i_n = 1, i_{n+1} = 0, i_{n+2}, ... = i_{n+m+2} = 1$, and all other $i_j = 0$.

$x_{n,m,k} = 0.i_1 i_2 i_3 ...$ where $i_1 = ... = i_n = 1, i_{n+1} = 0, i_{n+2} = ... = i_{n+m+2} = 1, i_{n+m+3} = 0, i_{n+m+4} = ... = i_{n+m+4+k} = 1$; and all other $i_j = 0$.

I.e., x_n is a decimal beginning with n 1's followed by 0's; $x_{n,m}$ is a decimal beginning with n 1's, followed by a 0, followed by m 1's, followed by 0's; and so on.

(a) $\{x_n : n \in \omega\}$ has order type ω.

(b) $\{x_{n,m} : n, m \in \omega\}$ has order type $\omega \times \omega$.

(c) $\{x_{n,m,k} : n, m, k \in \omega\}$ has order type $\omega \times \omega \times \omega$.

(a) is obvious; we prove (b) and leave (c) to the reader.

By example 5.30 it suffices to show that $\{x_{n,m} : n, m \in \omega\}$ is order-isomorphic to $\omega \times \omega$ under the lexicographic order. So suppose $(n, m) \leq_L (k, s)$. If $(n, m) = (k, s)$ we are done. If not, denote $x_{n,m} = 0.i_1 i_2 i_3 ...$ and $x_{k,s} = j_1 j_2 j_3 ....$ We consider two cases:

Case 1 $n < k$. Then $i_r = j_r$ for all $r \leq n$ and $i_{n+1} = 0 < 1 = j_{n+1}$. So $x_{n,m} < x_{k,s}$.

Case 2 $n = k$ and $m < s$. Then $i_r = j_r$ for all $r \leq n + m + 2$, $i_{n+m+3} = 0 < 1 = j_{n+m+3}$. So $x_{n,m} < x_{k,s}$.

Hence $x_{n,m} \leq x_{k,s}$ iff $(n, m) \leq_L (k, s)$, which concludes the proof.

Are there well-ordered sets not order-isomorphic to some ordinal? The next theorem says no.

Theorem 5.32. *Every well-ordered set is order-isomorphic to an ordinal.*

Proof. Let x be well-ordered by $\leq$. We define f a function whose range is x by induction. $f(0)$ is the minimal element of x. For each α, if $f[\alpha] = \{f(\beta) : \beta < \alpha\} \neq x$ then $f(\alpha)$ is the least element in $x \setminus f[\alpha]$. If $f[\alpha] = x$ then $f : \alpha \to x$ is the desired 1-1 onto order-preserving map. □

So ordinals are canonical well-ordered sets.

Our next goal is to characterize countable ordinals in terms of $\mathbb{R}$. Before doing this we need

Proposition 5.33. *Every non-zero countable limit ordinal is some* $\bigcup_{n<\omega} \beta_n$ *where each* $\beta_n < \beta_{n+1}$.

Proof. Suppose α is a countable limit ordinal. Since α is countable there is a 1-1 onto function $h : \omega \to \alpha$. Define an increasing function $f : \omega \to \alpha$ as follows: $f(0) = h(0)$. Give $f|_{n+1}$, let k be the least natural number $> n + 1$ so $h(k) > f(n), h(n + 1)$. Define $f(n + 1) = h(k)$. Clearly f is 1-1, and $f(n) < f(n + 1)$ for all n, and each $f(n) \geq h(n)$. Hence $f[\omega]$ is cofinal in α. Letting $\beta_n = f(n), \alpha = \bigcup_{n<\omega} \beta_n$. □

We characterize countable ordinals as follows:

Theorem 5.34. *An ordinal* α *is countable iff some subset of* $\mathbb{R}$ *has order type* α.

Proof. Suppose $A \subseteq \mathbb{R}$ has order type α. Then $A = \{x_\beta : \beta < \alpha\}$ where if $\beta < \gamma < \alpha$ then $x_\beta < x_\gamma$. Consider the open intervals $I_\beta = (x_\beta, x_{\beta+1}) \subset \mathbb{R}$. The I_β's are pairwise disjoint, and each contains a rational, so there are only countably many of them. Hence A is countable.

Now suppose α is countable and we know that every $\beta < \alpha$ is order-isomorphic to a subset of $\mathbb{R}$. If $\alpha = \beta + 1$ then some subset $A \subset (0,1)$ is order isomorphic to β, so $A \cup \{1\}$ is order isomorphic to α. If α is a limit, then $\alpha = \bigcup_{n<\omega} \beta_n$ as in proposition 5.33. Let $\delta_0 = \beta_0$ and let δ_{n+1} be the order type of $\beta_{n+1} \setminus \beta_n$. Let $A_n \subset (n, n+1) \subset \mathbb{R}$ so A_n has order type δ_n. Then $\bigcup_{n<\omega} A_n$ has order type α. □

Theorem 5.34 tells us why uncountable ordinals are so hard to picture. They cannot be embedded in $\mathbb{R}$.

So our next task is to give a few examples of ordinal arithmetic involving specific uncountable ordinals.

Example 5.35. (a) $\omega + \omega_1 = \omega_1 < \omega_1 + \omega$.

(b) $\omega \cdot \omega_1 = \omega_1 < \omega_1 \cdot \omega$.

(c) $\omega^{\omega_1} = \omega_1 < \omega_1{}^{\omega}$

(d) $\omega_1 < \omega_1 + \omega < \omega_1 \cdot \omega < \omega_1^\omega$.

For the proof of (a): $\omega + \omega_1 = \sup\{\omega + \alpha : \alpha < \omega_1\}$. Since, if $\alpha < \omega_1$ then $\omega + \alpha$ is countable, $\omega_1 \leq \sup\{\omega + \alpha : \alpha < \omega_1\} \leq \omega_1$.

For (d): $\omega_1 < \omega_1 + 1 < \omega_1 + \omega < \omega_1 + \omega_1 < \omega_1 \cdot \omega < \omega_1 \cdot \omega_1 < \omega_1{}^{\omega}$.

(b) and (c) are left to the reader.

The next theorem says that ordinal arithmetic has division and remainder.

Theorem 5.36. *Let $0 < \alpha \leq \beta$ be ordinals. Then there are ordinals δ, γ with $\gamma < \alpha$ and $\beta = \alpha \cdot \delta + \gamma$.*

Proof. Let $\delta = \sup\{\rho : \alpha \cdot \rho \leq \beta\}$. By definition of δ, $\alpha \cdot \delta \leq \beta$ and $\alpha \cdot (\delta + 1) > \beta$. Let γ be the order type of $\beta \setminus \alpha \cdot \delta$. Then $\alpha \cdot \delta + \gamma = \beta$ and, since $\beta < \alpha \cdot (\delta + 1)$, $\gamma < \alpha$. □

The next theorem says that right-cancellation fails spectacularly.

Theorem 5.37. *For all ordinals α, β, $\alpha + \beta = \beta$ iff $\beta \geq \alpha \cdot \omega$.*

Proof. If $\alpha + \beta = \beta$ then $\beta \geq \alpha$ so there are δ, γ with $\beta = \alpha \cdot \delta + \gamma, \gamma < \alpha$.

$$\alpha + \beta = \alpha + (\alpha \cdot \delta) + \gamma = \alpha(1 + \delta) + \gamma.$$

Note that $\beta \geq \alpha \cdot \omega$ iff $\delta \geq \omega$. If $\delta < \omega$ then $\alpha(1 + \delta) + \gamma > \alpha \cdot \delta + \gamma = \beta$. If $\delta \geq \omega$ then $\alpha(1 + \delta) + \gamma = \alpha \cdot \delta + \gamma = \beta$. □

Fix α. Theorem 5.37 says that the functional $f_\alpha(\beta) = \alpha + \beta$ has many fixed points.[78] The functional $g_\alpha(\beta) = \alpha \cdot \beta$ also has many fixed points (see below). But by proposition 5.23, the functionals $f(\beta) = \beta + \alpha$ (for $\alpha > 0$) and $g(\beta) = \beta \cdot \alpha$ (for $\alpha > 1$) have no fixed points.

[78] Here the word "functional" is short for: a class function defined by means of a formula.

There is a general theorem about fixed points.

Definition 5.38. A class (respectively set) function f from ON (respectively δ) to ON is continuous iff it is nondecreasing and, for all limit α (respectively, limit $\alpha < \delta$), $f(\alpha) = \sup\{f(\beta) : \beta < \alpha\}$.

Every constant function is continuous. By definition, f_α and g_α as above are continuous, as is h_α where $h_\alpha(\beta) = \alpha^\beta$.

Theorem 5.39. *A continuous strictly increasing class function defined on ON has arbitrarily high fixed points, i.e., for every α there is some $\beta > \alpha$ with $f(\beta) = \beta$.*

Proof. Note that, by strictly increasing, each $f(\alpha) \geq \alpha$. Let $\beta_0 = f(\alpha)$ and define $\beta_{n+1} = f(\beta_n)$ for all finite n. Let $\beta = \sup\{\beta_n : n < \omega\}$. By continuity, $f(\beta) = \beta \geq \alpha$. □

Corollary 5.40. *For each ordinal α there are arbitrarily high limit ordinals with $\alpha + \beta = \beta = \alpha \cdot \beta$.*

Proof. Let β be a fixed point for the function $f(\beta) = \alpha \cdot \beta$ where $\beta \geq \alpha \cdot \omega$. By theorem 5.39, $\alpha + \beta = \beta$. □

Finally, for completeness, we close this section by defining infinite ordinal operations.

Theorem 5.41. *For each ordinal β and each set of ordinals $\{\alpha_\gamma : \gamma < \beta\}$:*

(a) If $\beta = \eta + 1$ then $\Sigma_{\gamma<\beta}\alpha_\gamma = \Sigma_{\gamma<\eta}\alpha_\gamma + \alpha_\eta$; if β is a limit, then $\Sigma_{\gamma<\beta}\alpha_\gamma = \sup\{\Sigma_{\gamma<\delta}\alpha_\gamma : \delta < \beta\}$.

(b) If $\beta = \eta + 1$ then $\Pi_{\gamma<\beta}\alpha_\gamma = (\Pi_{\gamma<\eta}\alpha_\gamma) \cdot \alpha_\eta$; if β is a limit, then $\Pi_{\gamma<\beta}\alpha_\gamma = \sup\{\Pi_{\gamma<\delta}\alpha_\gamma : \delta < \beta\}$.

For example, in ordinal arithmetic, $1 + 2 + ... + \omega = \omega + \omega; 1 \cdot 2 \cdot ... \cdot \omega = \omega \cdot \omega$.

5.4 Cardinal arithmetic

Without the axiom of choice there is almost nothing you can say about cardinal arithmetic, so, reversing our previous convention, from now on we assume AC unless we explicitly say that we don't.

To distinguish between cardinal and ordinal arithmetic notations, we introduce the following conventions. Lowercase Greek letters in the first part of the alphabet ($\alpha, \beta, \gamma, \delta$ and so on) represent ordinals, and arithmetic operations involving them (e.g., $\alpha+\beta$) are always ordinal operations. Later lowercase Greek letters (e.g., κ, λ, ρ and so on) represent cardinals, and operations involving only them (e.g., $\kappa + \lambda$) are cardinal operations unless noted otherwise. Mixed notation (e.g., $\lambda + \alpha$) are ordinal operations, as are all $\lambda + n$ where n is finite, unless otherwise noted.

Cardinal arithmetic is not interested in arrangement, only quantity. So if you have three trees in your front yard and four trees in your back yard, you have seven trees all together. That's addition: finding the size of the union of disjoint sets. And if you have seven rows of chairs, with five chairs in each row, you have 35 chairs. That's cardinal multiplication, the size of a Cartesian product.

Definition 5.42. (a) $|x| + |y| = |(x \times \{0\}) \cup (y \times \{1\})|$

(b) $|x| \cdot |y| = |x \times y|$.

(c) $|x|^{|y|} = |x^y|$.

Definition 5.43. (a) $\kappa + \lambda = \rho$ iff there are disjoint sets x, y with $|x| = \kappa, |y| = \lambda$, and $|x \cup y| = \rho$.

(b) $\kappa \cdot \lambda = |\kappa \times \lambda|$.

(c) $\kappa^\lambda = |\{f : f \text{ a function from } \kappa \text{ to } \lambda\}|$.

Neither definition 5.42 nor definition 5.43 need AC for their statement, but we need AC to show that definition 5.43 is equivalent to definition 5.42, which is left to the reader.

Cardinal arithmetic on addition is very easy, due to the following definition and proposition:

Definition 5.44. An ordinal α is even iff it has the form $\beta + 2n$ where β is a limit ordinal and $n < \omega$; otherwise it is odd.

Proposition 5.45. *Let κ be an infinite cardinal. $\kappa = |\{\alpha < \kappa : \alpha \text{ even}\}| = |\{\alpha < \kappa : \alpha \text{ odd}\}|$.*

Proof. Define $f(\alpha + 2n) = \alpha + n; g(\alpha + 2n) = \alpha + 2n + 1$. f shows that $\kappa = |\{\alpha < \kappa : \alpha \text{ even}\}|$ and g shows that $|\{\alpha < \kappa : \alpha \text{ even}\}| = |\{\alpha < \kappa : \alpha \text{ odd}\}|$. □

(Note that back in theorem 1.37 we essentially proved that every ordinal is either even or odd.) From proposition 5.45 it's immediate that $\kappa + \kappa = \kappa$.

Theorem 5.46. *Let κ, λ be infinite cardinals. Then*

(a) $\kappa + \kappa = \kappa = \kappa \cdot \kappa$

(b) $\kappa + \lambda = \sup\{\kappa, \lambda\} = \kappa \cdot \lambda$

Proof. For (a): By proposition 5.45, $\kappa + \kappa = \kappa$. We need to show that $\kappa = |\kappa \times \kappa|$.

Define the following order $\leq$ on $\kappa \times \kappa$: for $(\alpha, \beta), (\gamma, \delta) \in \kappa \times \kappa$, we say $(\alpha, \beta) \leq (\gamma, \delta)$ iff $\alpha + \beta < \gamma + \delta$, or $\alpha + \beta = \gamma + \delta$ and $\alpha \leq \gamma$.[79]

$\leq$ is easily seen to be well-ordered.

Suppose $\kappa = \aleph_\delta$ and, for all $\lambda = \aleph_\beta$ with $\beta < \delta$, $\lambda \cdot \lambda = \lambda$. For each $\alpha < \kappa$, define $B_\alpha = \{(\gamma, \varepsilon) \in \kappa \times \kappa : \gamma + \varepsilon = \alpha\}$. For $\beta < \kappa$, let $C_\beta = \bigcup_{\alpha < \beta} B_\alpha$. $C_\beta \subseteq \beta \times \beta$ so, by induction hypothesis, $|\beta| \leq |C_\beta| < \kappa$.

Each C_β is an initial segment of $\kappa \times \kappa$ under $\leq$, so the order type of $\kappa \times \kappa$ is at least κ. Since any initial segment of $\kappa \times \kappa$ under $\leq$ is contained in some C_β, the order type of $\kappa \times \kappa$ is at most κ.

Hence $|\kappa \times \kappa| = \kappa$.

For (b): Suppose $\kappa < \lambda$. Then $\lambda \leq \kappa + \lambda \leq \lambda + \lambda = \lambda$ and $\lambda \leq \kappa \cdot \lambda \leq \lambda \cdot \lambda = \lambda$.

□

Note that the statement $\kappa^2 = \kappa$ is equivalent to the statement that the arbitrary union of κ many sets, each of size κ, has size κ.

Addition and multiplication were easy, but exponentiation is much trickier.

First, a basic theorem:

Theorem 5.47. *(a) For κ infinite, $2 \leq \lambda \leq \kappa$, $\lambda^\kappa = \kappa^\kappa$.*

(b) $(\kappa^\lambda)^\rho = \kappa^{\lambda \dot{\rho}}$

Proof. For (a) Let φ be a 1-1 function from $\kappa \times \kappa$ onto κ. A function $f : \kappa \to \kappa$ is a subset of $\kappa \times \kappa$. So we define $\psi(f) = \{\varphi(\alpha, f(\alpha)) : \alpha < \kappa\}$. Then ψ is a 1-1 function from $\{f : f : \kappa \to \kappa\}$ to $\mathcal{P}(\kappa)$ and $2^\kappa = \kappa^\kappa$. Since $\{f : f : 2 \to \kappa\} \subseteq \{f : f : \lambda \to \kappa\} \subseteq \{f : f : \kappa \to \kappa\}$, $2^\kappa \leq \lambda^\kappa \leq \kappa^\kappa = 2^\kappa$.

For (b), we define a function φ as follows: if $f : \rho \to \kappa^\lambda$ and $\alpha < \lambda, \beta < \rho$, then $\varphi(f)(\alpha, \beta) = f(\beta)(\alpha)$. φ is the desired 1-1 onto function. □

Theorem 5.47 is almost all we can say. In theorem 5.46 we said exactly what $\kappa + \lambda$ and κ^λ are. There is no similar result for exponentiation. Even 2^ω could be nearly anything in the following precise sense: If κ is a cardinal of uncountable cofinality (defined in the next section) then it is consistent that $2^\omega = \kappa$.

2^ω is not an anomaly. Most questions involving exponentiation are undecidable. When one of them is discovered not to be, it is a major breakthrough (see the discussion of the singular cardinals problem in section 5.6).

Although we cannot know what 2^ω really is,[80] we can hypothesize what it might be. The oldest of these hypotheses is the continuum hypothesis, CH: $2^\omega = \omega_1$. The general hypothesis, GCH, extends this: $\forall \kappa \geq \omega \; 2^\kappa = \kappa^+$. Cantor believed that CH was true and spent many fruitless years trying to prove it. In 1937, Gödel proved that GCH is consistent by showing that it holds in a

[79] Recall: we are speaking of ordinal sums here.

[80] whatever that means

particular model of set theory known as the constructible universe (see chapter 6). In 1963, Cohen proved that CH does not hold in all models of set theory, via the technique of forcing. In fact, forcing was invented by Cohen to show the consistency of the failure of CH and the consistency of the failure of AC, and has become the main method for showing statements to be independent of ZF or ZFC.

5.5 Cofinality

Our previous classifications of infinite cardinals — into countable and uncountable cardinals, and into successor and limit cardinals — are far too coarse. Here we develop a more useful classification, based on the concept of cofinality.

Recall (see the comment after definition 1.17) that a set A is cofinal in a linear order X iff $A \subset X$ and, for every $x \in X$ there is $a \in A$ with $x \leq a$. In particular, if α is a limit ordinal, then A is cofinal in α iff $\alpha = \bigcup A$. In an abuse of notation we say

Definition 5.48. Let α, β be ordinals. We say that α is cofinal in β iff there is a set of order type α which is cofinal in β.

Definition 5.49. Let α, β be ordinals. We define $\text{cf}(\alpha) = \beta$ iff β is the least ordinal cofinal in α.

Clearly each $\text{cf}(\alpha) \leq \alpha$.

Note that if α is a successor ordinal, then $\text{cf}(\alpha) = 1$, that $\text{cf}(\beta) = \text{cf}(\alpha + \beta) = \text{cf}(\alpha \cdot \beta)$; and that if β is a limit then $\text{cf}(\beta) = \text{cf}(\aleph_\beta) = \text{cf}(\beth_\beta)$.

What sorts of ordinals can be cofinalities?

Theorem 5.50. *(a) An infinite ordinal of cardinality κ has cofinality at most κ*

(b) For all ordinals α, cf(α) is a cardinal.

Proof. For (a): Suppose $|\alpha| = \kappa$ and $\rho = \text{cf}(\alpha)$. Let $f : \kappa \to \alpha$ be 1-1 and onto, and define g with dom $g \subseteq \kappa$ and range g cofinal in α as follows: $g(\beta) = \inf\{f(\gamma) : \gamma \geq \beta \text{ and } f(\gamma) > g(\delta) \text{ for all } \delta < \beta\}$. For some $\delta \leq \kappa$, dom $g = \delta$. Note that g is a strictly increasing function whose range is cofinal in α, so $\delta \geq \rho$.

For (b): If $|\alpha| = \kappa < \alpha = \text{cf}(\beta)$, then by (a), $\text{cf}(\alpha) \leq \kappa$. Hence some $\delta \leq \kappa$ is cofinal in α which is cofinal in β, so α is not the least ordinal cofinal in β, a contradiction. □

While every cofinality is a cardinal, not every cardinal is a cofinality, e.g., $\text{cf}(\aleph_\omega) = \omega$.

The important bifurcation of cardinals is given by

Definition 5.51. A cardinal κ is regular iff $\kappa = \text{cf}(\kappa)$; otherwise κ is singlar.

For example, $\aleph_\omega$ is singular, but ω_1 is regular. The latter follows from

Theorem 5.52. *(a) For every cardinal κ, κ^+ is regular.*

(b) If $\kappa = \text{cf}(\alpha)$ for some α, then κ is regular.

Proof. (a) By theorem 5.50, $\text{cf}(\kappa^+) = \kappa^+$ or $\text{cf}(\kappa^+) \leq \kappa$. If the former, we're done, so suppose the latter. Let $\lambda = \text{cf}(\kappa^+)$. There is an increasing cofinal 1-1 map $f : \lambda \to \kappa^+$. For each $\alpha < \lambda$ define $a_\alpha = \{\gamma < \kappa^+ : \gamma < f(\alpha)\}$. $\kappa^+ = \bigcup_{\alpha<\lambda} a_\alpha$ and each $|a_\alpha| \leq \kappa$, so $\kappa^+ = |\kappa^+| \leq \lambda \cdot \kappa \leq \kappa < \kappa^+$, a contradiction.

(b) If $\delta = \text{cf}(\kappa)$ and $\kappa = \text{cf}(\alpha)$ then, by composition of functions, $\delta \geq \text{cf}(\alpha) = \kappa$. So $\delta \leq \kappa \leq \delta$, i.e., $\kappa = \text{cf}(\kappa)$. □

By theorem 5.52, every infinite successor cardinal is regular. Are there any regular limit cardinals? Clearly ω is one. In the next chapter we will show that the existence of an uncountable regular limit cardinal implies the consistency of ZF. Hence, by Gödel's incompleteness theorems, ZF (and in fact ZFC) cannot prove the existence of a regular uncountable limit cardinal. Such a cardinal is called weakly inaccessible. (Strongly inaccessible cardinals will be defined in the next chapter.) On the other hand, the existence of an uncountable regular limit cardinal seems like little enough to ask of the universe, and inasmuch as set theorists can be said to believe in the existence of the mathematical universe (whatever "existence" means in this context) they can probably be said to believe in the existence of weakly inaccessible cardinals. Theorems which begin "if there is a weakly inaccessible cardinal, then..." are fairly innocuous. But, on the other hand, when a theorem does begin that way, an effort is always made to check that the hypothesis is necessary.

5.6 Infinite operations and more exponentiation

Generalizing section 5.4, we define infinite cardinal sum and infinite cardinal product as follows:

Definition 5.53. Given I, suppose κ_i is a cardinal for each $i \in I$. We define $\Sigma_{i\in I}\kappa_i = |\bigcup_{i\in I} x_i|$ where each $|x_i| = \kappa_i$ and the x_i's are pairwise disjoint. In an abuse of notation, we define $\Pi_{i\in I}\kappa_i = |\Pi_{i\in I}\kappa_i|$, where the first Π refers to cardinal product, and the second refers to set-theoretic product.

The connection between infinite sum and infinite product is given by a theorem known as König's theorem:

Theorem 5.54. *Suppose $I \neq \emptyset$ and $\kappa_i < \lambda_i$ for all $i \in I$. Then $\Sigma_{i\in I}\kappa_i < \Pi_{i\in I}\lambda_i$.*

Proof. Let $\{x_i : i \in I\}$ be pairwise disjoint, each $|x_i| = \kappa_i$, and let $f_i : x_i \to \lambda_i \setminus \{0\}$ be 1-1. For $y \in x_i$ we define $f_y \in \Pi_{j\in I}\lambda_i$ by $f_y(i) = f_i(y); f_y(j) = 0$ for all $j \neq i$. The map F with domain $\bigcup_{i\in I} x_i$ defined by $F(y) = f_y$ is a 1-1 map from $\bigcup_{i\in I} x_i$ into $\Pi_{i\in I}\lambda_i$.[81] So $\Sigma_{i\in I}\kappa_i \leq \Pi_{i\in I}\lambda_i$.

Why is $\Sigma_{i\in I}\kappa_i \neq \Pi_{i\in I}\lambda_i$? Let $G : \bigcup_{i\in I} x_i \to \Pi_{i\in I}\lambda_i$. We need to show that G is not onto.

Define $z_i = \{G(y)(i) : y \in x_i\}$. Each $|z_i| \leq \kappa_i < \lambda_i$ so there is $\gamma_i \in \lambda_i \setminus z_i$. Let $f(i) = \gamma_i$ for all $i \in I$. Then $f \in \Pi_{i\in I}\lambda_i$ and $f \in$ range G iff $f = G(y)$ for some i and some $y \in x_i$ iff $f(i) \in z_i$ for some i, a contradiction. So $f \notin$ range G. □

With the concept of cofinality and theorem 5.54 we can get sharper theorems about exponentiation. The basic elementary theorem is

Theorem 5.55. *For each infinite cardinal κ*

(a) $\kappa < \kappa^{\mathrm{cf}(\kappa)}$.

(b) $\kappa < cf(2^\kappa)$.

Proof. For (a): Let f be a 1-1 increasing cofinal map from $\mathrm{cf}(\kappa)$ to κ. Then $\kappa = |\bigcup_{\alpha<\mathrm{cf}(\kappa)} f_\alpha| \leq \Sigma_{\alpha<\mathrm{cf}(\kappa)}|f(\alpha)| < \kappa^{\mathrm{cf}(\kappa}$ by König's theorem.

For (b): Let $\lambda = \mathrm{cf}2^\kappa$. Then $\lambda \leq \kappa$ iff $(2^\kappa)^\lambda = 2^\kappa$, contradicting (a). □

Note that, since $\mathrm{cf}(\lambda) \leq \lambda$ for all λ, theorem 5.55(b) implies the infintary case of Cantor's theorem 5.9.

Theorem 5.55 nearly ends what can be said about exponentiation without getting into consistency results.[82] Even an elementary statement such as "If $2^\kappa \geq 2^\lambda$ then $\kappa \geq \lambda$" is independent.

The exceptions are remarkable results due to Silver, to Galvin and Hajnal, and more recently to Shelah, on the singular cardinals problem. To put these results in context, and to explain the situation for regular cardinals, we first state Easton's theorem.

[81] here Π refers to set-theoretic product

[82] The remaining theorem is Bukovsky's theorem, proved independently by Hechler, that κ^κ is determined by $\kappa^{\mathrm{cf}(\kappa)}$ — see the exercises for part of this.

Theorem 5.56. *Let F be a nondecreasing function from the class of infinite regular cardinals to the class of cardinals where $cf(F(\kappa) > \kappa$ for all κ.*[83] *Then if ZFC is consistent, so is ZFC + "for all regular $\kappa, 2^\kappa = F(\kappa)$."*

Easton's theorem is proved via forcing, the technique created by Cohen in 1963 to construct new models of set theory from old models, is a complicated generalization of Cohen's original proof that 2^ω need not be ω_1, and hence lies beyond the scope of this book. It completely settles the question of what the possibilities are for 2^κ: the only constraint is the constraint of theorem 5.56.

The singular cardinal problem, then, is: what are the constraints on 2^κ when κ is singular?

Silver was the first to notice that if κ is singular with uncountable cofinality then there are strong constraints on 2^κ. The easiest result of his to state is

Theorem 5.57. *Let $\aleph_\alpha$ be a singular cardinal of uncountable cofinality and, suppose that for some fixed $\gamma < cf(\alpha)$ and every ordinal $\beta < \alpha, 2^{\aleph_\beta} = \aleph_{\beta+\gamma}$. Then $2^{\aleph_\alpha} = \aleph_{\alpha+\gamma}$.*

In particular, letting $\gamma = 1$, this says that if GCH holds below some singular $\aleph_\alpha$ of uncountable cofinality, then it holds at $\aleph_\alpha$.

This result was surprising, since Prikry and Silver had previously shown that, modulo the consistency of a particular cardinal stronger than a measurable cardinal (see chapter 7) there could be a singular cardinal κ with countable cofinality, GCH held below κ, but $2^\kappa = \kappa^{++}$.[84]

Silver's theorem was generalized by, among others, Galvin and Hajnal. The result of theirs which is easiest to state is

Theorem 5.58. *If κ is a singular cardinal of uncountable cofinality, if $2^\lambda < \kappa$ for all $\lambda < \kappa$ and $\kappa < \aleph_\kappa$, then $2^\kappa < \aleph_\kappa$.*

Jensen, in a remarkable result which has inspired analogues in other contexts, showed that the negation of various large cardinal axioms also affect the possibilities for 2^κ when κ is singular, basically reducing the issue to the combinatorics of the constructible universe L (see chapter 6).

Shelah then proved the following remarkable theorem

Theorem 5.59. *(a) If $2^\omega < \aleph_\omega$ then $(\aleph_\omega)^\omega < \aleph_{\omega_4}$*

(b) If $\forall n\ 2^{\omega_n} < \aleph_\omega$ then $2^{\aleph_\omega} < \aleph_{\omega_4}$.[85]

This was proved using a technique Shelah developed for the proof, pcf theory, which turns out to have many applications, both in set theory and outside it.

[83] Note that F is a class, and that $F(\kappa)$ need not itself by regular.

[84] "There could be" means "it's consistent."

[85] Again, this is the simplest version to state.

5.7 Counting

Now that we've got the basic rules of cardinal arithmetic, we can calculate the size of various common mathematical objects.

Example 5.60. How big is $\mathbb{R}$? We already answered this in theorem 5.10, but here's a short proof. By section 4.7, every real corresponds to a set of rationals. So $|\mathbb{R}| \leq \omega^\omega = 2^\omega$. And, since the Cantor set is a subset of $\mathbb{R}$, $|\mathbb{R}| \geq 2^\omega$. So $|\mathbb{R}| = 2^\omega$.

Similarly

Example 5.61. If D is dense in a linear order X, then $|X| \leq 2^{|D|}$. (Here, D is dense in X iff for every $x < y \in X$ there is some $d \in D$ with $x \leq d \leq y$.) This follows from the following: if $x \neq y$ then either $\{d \in D : d < x\} \neq \{d \in D : d < y\}$ or $x \notin D, y \in D$ and $y = S(x)$. The reader is invited to fill in the details.

Example 5.62. Let $\mathbb{P}$ be the set of irrationals. What's $|\mathbb{P}|$? $\mathbb{Q}$ is countable, and $2^\omega = |\mathbb{R}| = |\mathbb{Q} \cup \mathbb{P}| = |\mathbb{Q}| + |\mathbb{P}|$, so $|\mathbb{P}| = 2^\omega$.

Example 5.63. Let K be an infinite field of size κ. How big is $P[K]$, the ring of polynomials in one variable over K? A polynomial over K has the form $\Sigma_{i \leq n} k_i x^i$ where $n \in \omega$ and each $k_i \in K$. So $|P[K]| = \bigcup_{n \in \omega} K^n$. Each $|K^n| = \kappa^n = \kappa$, so $|P[K]| = \kappa \cdot \omega = \kappa$.

As a corollary, the completion of an infinite field of size κ also has size κ.

Example 5.64. How many open sets of reals are there? Recall that an open set of reals is a union of open intervals with rational endpoints. Let $\mathcal{I}$ be the set of open intervals with rational endpoints. $|\mathcal{I}| = |\mathbb{Q}|^2 = \omega$. Given an open set u, we define $f(u) = \{I \in \mathcal{I} : I \subseteq u\}$. The function f is 1-1 and $\mathcal{I}$ is countable, so there are at most 2^ω open sets of reals. Since, for every $r \in \mathbb{R}, \mathbb{R} \setminus \{r\}$ is open, there are exactly 2^ω open sets of reals.

Example 5.65. Define $C(\mathbb{R})$ to be the set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$. How big is $C(\mathbb{R})$? Every continuous function from $\mathbb{R}$ to $\mathbb{R}$ is determined by its values on $\mathbb{Q}$, so there are at most $|\mathbb{R}|^{|\mathbb{Q}|} = (2^\omega)^\omega = 2^\omega$ many such functions. There are exactly that many, since for each $r \in \mathbb{R}$, the function $f(x) = r$ for all x is continuous.

Example 5.66. For any x, let $[x]^\kappa$ denote the subsets of x of size exactly κ. If κ is infinite, how big is $[\kappa]^\kappa$? Given $a \in [\kappa]^\kappa$, let χ_a be the characteristic function of a. There are at most 2^κ many χ_a's, so $[\kappa]^\kappa \leq 2^\kappa$. Since $\kappa \times \kappa = \kappa$, there is a pairwise disjoint family $\{a_\alpha : \alpha < \kappa\} \subset [\kappa]^\kappa$, where each $a_\alpha = \{\beta_{\gamma,\alpha} : \gamma < \kappa\}$. If $f : \kappa \to \kappa$, define $b_f = \{\beta_{(f(\alpha),\alpha)} : \alpha < \kappa\}$. Each $b_f \in [\kappa]^\kappa$, and if $f \neq g$ then $b_f \neq b_g$, so $[\kappa]^\kappa = 2^\kappa$.

5.8 Exercises

Unless told otherwise, you may use AC.

1. Use definition 5.6 to show directly that $\omega \times \omega$ and $\omega \times \omega \times \omega$ are countable. Do not use corollary 5.8.

2; Use definition 5.6 to show directly that the set of polynomials in one variable in integer coefficients is countable. Do not use corollary 5.8.

3; Use definition 5.1 to show directly that $|\mathbb{R}| = |\mathbb{R}^2| = |\mathbb{R}^n|$ for each finite n. Do not use AC.

4. Use definition 5.1 to show directly that the set of points on the surface of the unit sphere has the same cardinality as $|\mathbb{R}|$.

5. Using the two previous exercises and the Schröder-Bernstein theorem, show that $|\mathbb{R}| = |$unit ball$|$.

6. An exercise on the proof of the Schröder-Bernstein theorem: let $x = \{0,1\} \times \omega, y = \omega, f : x \to y$ where $f(0,n) = 4n, f(1,n) = 4n+3, g : y \to x$ where $g(n) = (0,n)$.

(a) What is each x_n, y_n?

(b) What are $x^\dagger, b^\dagger$?

7. Prove that, in the proof of theorem 5.10, f is 1-1 and onto.

8. Prove that if $|x| = |y|$ then $|\mathcal{P}(x)| = |\mathcal{P}(y)|$

9. Prove that, in the proof of proposition 5.12, χ is 1-1 onto.

In the following exercises, all arithmetic is ordinal arithmetic:

10. Show that, for every infinite ordinal α

(a) $0 + \alpha = \alpha$

(b) $1 \cdot \alpha = \alpha$

(c) $|\alpha| = |\alpha + 1|$.

11. Find sets of reals of order type (a) $\omega + \omega + \omega + \omega + \omega$; (b) $\omega \cdot \omega + \omega \cdot \omega$.

12. Prove subclaim 5.27.1

13. Complete the proof of theorem 5.27, including filling in the details of cases 1 and 2 in the proof of (a), i.e., justifying each step in the string of equalities.

14. Prove that example 5.30 is correct.

15. Prove that example 5.31(c) is correct.

16. Prove that example 5.35(b) and (c) are correct.

17. Find the correct order relations among $\omega_1 + \omega_2, \omega_2 + \omega_1, \omega_1 \cdot \omega_2, \omega_2 \cdot \omega_1, \omega_1$, and ω_2.

18. List the following ordinals in nondecreasing order. State which ordinals are in fact equal to each other: $\omega^{\omega_1}; 3^\omega; \omega_1^\omega; \omega; \omega_1; \omega^3; \omega \cdot 3; \omega \cdot \omega_1; \omega_1 \cdot \omega$.

19. Now do the previous exercise replacing each ω_1 by ω_{17} and each ω by ω_{15}.

20. List the following ordinals in nondecreasing order and state which ordinals are in fact equal to each other: $\Sigma_{i<\omega}\ i; \Sigma_{i<\omega}\ \omega \cdot i; \Pi_{0<i<\omega}\ i; \Pi_{0<i<\omega}\ \omega \cdot i; \omega \cdot \omega; \omega^\omega; \Sigma_{i<\omega}\ \omega^i; \Pi_{0<i<\omega}\ \omega^i$.

21. An ordinal is indecomposable iff it is not the sum of two strictly smaller ordinals. For example, both ω and $\omega \cdot \omega$ are indecomposable.

(a) Show that $\alpha \cdot \omega$ is indecomposable for all α.

(b) Show that if δ is the largest indecomposable ordinal $< \alpha$ then there is some β so that if γ is the order type of $\alpha \setminus (\delta \cdot \beta)$ then $\gamma < \delta$.[86]

(c) Show that every ordinal is a finite sum of indecomposable ordinals (one of which might be 0).

In the following exercises, all arithmetic is cardinal arithmetic:

22. (a) Show that $\kappa + \kappa = 2 \cdot \kappa$.

(b) Show that $\kappa^2 = \kappa \cdot \kappa$.

23. (a) What is $\Sigma_{n<\omega}\ \omega_n$?

(b) What is $\Pi_{n<\omega}\ \omega_n$?

24. Show that definition 5.42 is equivalent to definition 5.43.

25. In the proof of theorem 5.46(a), show that $\leq$ is a well-order.

26. Show the distributive law: $\kappa \cdot (\lambda + \rho) = \kappa \cdot \lambda + \kappa \cdot \rho$.

27. (Hausdorff) Show that $\omega_2^{\omega_1} = 2^{\omega_1}$.

28. Arrange the following cardinals in nondecreasing order, and state which cardinals are in fact equal to each other: $\omega^{\omega_1}; 3^\omega; \omega_1^\omega; \omega; \omega_1; \omega^3; \omega \cdot 3; \omega \cdot \omega_1; \omega_1 \cdot \omega$.

29. Show that $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$ for all ordinals α.

30. If I is infinite, must $\Sigma_{i \in I} \kappa_i = \Pi_{i \in I} \kappa_i$?

31. Show that κ is singular iff it is the union of fewer than κ sets, each of which has size $< \kappa$.

32. (a) Show that if κ is regular and $\lambda < \kappa$ then $\kappa^\lambda = \kappa \cdot \sup\{|\rho^\lambda| : \rho < \kappa\}$.

(b) Show that if κ is weakly inaccessible and $\lambda < \kappa$ then $\kappa^\lambda = \sup\{|\rho^\lambda| : \rho < \kappa\}$.

(c) Assuming CH, show that if κ is regular and $\lambda < \kappa$ then $\kappa^\lambda = \kappa$.

33. Suppose κ is weakly inaccessible. Using Easton's theorem, which of the following are consistent? (a) $2^\omega < \kappa$ (b) $2^\omega = \kappa$; (c) $2^\omega > \kappa$.

34. For $\gamma < \lambda$ let $\kappa_\gamma = \kappa$. Show that $\kappa \cdot \lambda = \Sigma_{\gamma<\lambda} \kappa_\gamma$ and $\kappa^\lambda = \Pi_{\gamma<\lambda} \kappa_\gamma$.

35. For $i \in I$, let each $\kappa_i \geq \omega$.

(a) Show that $\Sigma_{i \in I} \kappa_i = \sup\{\kappa_i : i \in I\} + |I|$.

[86] This is division with remainder for ordinal arithmetic.

(b) Show that $\Pi_{i\in I}\kappa_i \neq \sup\{\kappa_i : i \in I\} + |I|$.

36. Use König's theorem to give a one-line proof that each $\kappa < 2^\kappa$

37. Show that if $2^\lambda < \kappa$ for every $\lambda < \kappa$ then $2^\kappa = \kappa^{\mathrm{cf}(\kappa)}$.

38. Which of the following statements are true, which are false, and which are independent?

(a) $2^\omega = \omega_1$

(b) $2^\omega = \aleph_\omega$

(c) $2^\omega < 2^{\omega_1}$

(d) $2^\omega \leq 2^{\omega_1}$

(e) $2^\omega = \aleph_{\omega_1}$

(f) $2^\omega = \omega_{17}$

(g) $2^\lambda < \aleph_{\omega_1}$ for all $\lambda < \aleph_{\omega_1}$

(h) if $2^\lambda = \lambda^{++}$ for all $\lambda < \aleph_{\omega_1}$ then $2^{\aleph_{\omega_1}} = \aleph_{\omega_1+2}$

(i) $(\aleph_\omega)^\omega = \aleph_{\omega_{10}}$

39. How many 1-1 functions are there from κ to κ^+? From κ to κ^{++}?

40. How many (a) partial orders; (b) linear orders; (c) well-orders are there on a set of size κ?

41. Complete the details of example 5.61.

42. A G_δ set is a countable intersection of open sets. How many G_δ sets are there in $\mathbb{R}$? In each $\mathbb{R}^n$, where $n < \omega$?

43. How many 1-1 functions are there

(a) from ω into ω_2?

(b) From ω_1 into ω_{17}?

6 Two models of set theory

In this chapter we develop two models of set theory: V_κ, where κ is strongly inaccessible, and Gödel's model L. For the first we must assume more than just ZFC. For the second, since we are constructing a class and not a set model, only ZFC need be assumed.

In chapter 3, we went to some trouble to avoid having to define parameters, and ended up with some clumsy notation. In this chapter things will be unreadable without the notion of parameter, so we define it forthwith:

Definition 6.1. Let Φ be a formula, F the set of free variables in Φ. If f is a function with domain F, the elements in the range of f are called parameters. The locution "with parameters in a" means that the range of f is a subset of a.

I.e., a parameter is a temporary name. Later we will have locutions such as "$\forall x_i \in a$." When we do this, x_i is a variable in the language (see section 2.1) and a is understood to be a parameter.

6.1 A set model for ZFC

Definition 6.2. A cardinal κ is a strong limit iff $\forall \lambda < \kappa \; 2^\lambda < \kappa$.

We state without proof:

Fact 6.3. *(a) Strong limits are infinite.*

(b) Strong limits are limit cardinals.

(c) ω is a regular strong limit.

(d) Assume GCH. Then every limit cardinal is a strong limit.

In the next section we'll meet a model of GCH, i.e., a model in which strong limits are abundant. What is not abundant are uncountable regular strong limits.

Definition 6.4. A strongly inaccessible cardinal is an uncountable regular strong limit.

Why are strongly inaccessible cardinals not abundant? Because their existence implies that ZFC has a model, so, by Gödel's theorem, they cannot be proved to exist by the axioms of ZF or ZFC.

Recall (definition 4.20) the definition of the V_α's: $V_0 = \emptyset; V_{\alpha+1} = \mathcal{P}(V_\alpha)$; if α is a limit then $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$.

Theorem 6.5. *Let κ be strongly inaccessible. Then V_κ is a model of ZFC.*

Proof. Every V_α is transitive (chapter 4 exercise 5(b)), so by theorem 3.7, each $V_\alpha \models$ extensionality. $\forall x \; x \models$ regularity. And if $x \in V_\alpha$, so is $\bigcup x$, so each $V_\alpha \models$ union. Thus every $V_\alpha \models$ extensionality, regularity, and union.

If $\alpha \geq \omega + 1, \omega \in V_\alpha$, so if $\alpha \geq \omega + 1, V_\alpha \models$ infinity.

Suppose α is a limit ordinal. If $x, y \in V_\alpha$ then $\exists \beta < \alpha \; x, y \in V_\beta$, hence $\{x, y\} \in V_{\beta+1} \subset V_\alpha$ and $\mathcal{P}(x) \in V_{\beta+2} \subset V_\alpha$. Hence $V_\alpha \models$ pairing and power set. Similarly, V_α is closed under finite product. So (by theorem 4.36) $V_\alpha \models$ choice. And, as a transitive set closed under power set, by corollary 3.25, $V_\alpha \models$ separation.

What's left is replacement.

First a subclaim:

Subclaim 6.5.1. Let κ be a regular strong limit cardinal.

(a) each element of V_κ has cardinality less than κ.

(b) $|V_\kappa| = \kappa$.

Proof. For (a) it suffices to show that if $\alpha < \kappa$ then $|V_\alpha| < \kappa$. So suppose $\alpha < \kappa$ and suppose $\forall \beta < \alpha \; |V_\beta| < \kappa$. If $\alpha = \beta + 1$ for some β then, by the induction hypothesis, $|V_\alpha| = 2^{|V_\beta|} < \kappa$. If α is a limit, then $|V_\alpha| = \sup\{|V_\beta| : \beta < \alpha\}$. Each $|V_\beta| < \kappa$ and $|\alpha| < \kappa$, so by regularity, $|V_\alpha| < \kappa$.

For (b): V_κ is an increasing union of κ many sets, so $|V_\kappa| \geq \kappa$. By (a), $|V_\kappa| \leq \kappa$. □

We use subclaim 6.5.1 to show that if κ is a regular strong limit cardinal then $V_\kappa \models$ replacement.

Let κ be a regular strong limit cardinal, and let Φ be as in the hypothesis of the axiom of replacement: Φ has $n + 2$ free variables for some n; $\forall p_1, ...p_n, x, y, z \in V_\kappa$, if $\Phi(p_1, ...p_n, x, y)$ and $\Phi(p_1, ...p_n, x, z)$ then $y = z$;[87] and if $p_1, ...p_n, x \in V_\kappa$ and $\Phi(p_1, ...p_n, x, y)$ then $y \in V_\kappa$. Let $a \in V_\kappa$. By subclaim 6.5.1, $|a| < \kappa$. Fix $p_1, ...p_n$. If $b = \{y : \exists x \in a \; \Phi(p_1, ...p_n, x, y)\}$, $|b| < \kappa$ and $b \subset V_\kappa$. $\forall y \in b$ rank $(y) < \kappa$, so by regularity $\sup\{\text{rank } (y) : y \in b\} = \beta < \kappa$. Hence $b \in V_{\beta+1} \subset V_\kappa$ and, by theorem 3.42, $V_\kappa \models$ replacement. □

While strongly inaccessible cardinals are needed to model all of ZFC, other ordinals give us quite a bit. As a corollary to the proof we have

Corollary 6.6. *(a) V_ω is a model of all of the axioms of ZFC except infinity.*

(b) if α is a limit cardinal, $\alpha > \omega$ then V_α is a model of all the axioms of ZFC except replacement.

We have already mentioned that, by Gödel's theorem

Theorem 6.7. *ZFC cannot prove that there is a strongly inaccessible cardinal.*

There is a second proof, which, while longer, is instructive, introducing us to the notion of reflection in models.

Proof. If there is a strongly inaccessible cardinal, there's a smallest one, κ. $V_\kappa \models$ ZFC. If ZFC proved that there is a strongly inaccessible cardinal, then there would be $x \in V_\kappa$ with $V_\kappa \models x$ is strongly inaccessible, i.e.

1. $V_\kappa \models x$ is a regular cardinal

[87] essentially, $p_1, ...p_n$ are parameters.

2. if $V_\kappa \models y$ is a cardinal and $V_\kappa \models y < x$ then $V_\kappa \models 2^y < x$.

We show that $V_\kappa \models$ "x is a cardinal and $\forall y$ if y is a cardinal and $y < x$ then $2^y < x$" iff it's really true that x is a cardinal and $\forall y$ if y is a cardinal and $y < x$ then $2^y < x$.

First of all, suppose $V_\kappa \models z$ is an ordinal, i.e., $V_\kappa \models z$ is transitive and linearly well-ordered by $\in$. Since V_κ is transitive, that means that z really is transitive and linearly well-ordered by $\in$, i.e., z really is an ordinal. Hence V_κ accurately reflects whether a set is an ordinal or not.

Now let's look at cardinality. Suppose $V_\kappa \models |a| \leq |b|$. Then there is some $f \in V_\kappa, f : a \to b, f$ is 1-1. I.e., $|a| \leq |b|$. Now suppose $|a| \leq |b|$, with $a, b \in V_\kappa$. There is $\beta < \kappa$ with $a, b \in V_\beta$, and there is $f : a \to b$ with f 1-1. Since f is a set of ordered pairs (z, y) with $z \in a, y \in b$, $f \in V_{\beta+2} \subset V_\kappa$. Hence $V_\kappa \models |a| \leq |b|$. I.e., V_κ accurately reflects whether two sets have the same cardinality.

Since cardinals are initial ordinals, $V_\kappa \models x$ is a cardinal iff x really is a cardinal.

We leave the proof that, for any cardinal $\lambda < \kappa$ $V_\kappa \models \lambda$ is regular iff λ is regular, to the reader.

Also, if $y \in V_\kappa$ then $2^y \in V_\kappa$. So if $V_\kappa \models \forall y$ if y is a cardinal and $y < x$ then $2^y < x$ iff it's really true that $\forall y$ if y is a cardinal and $y < x$ then $2^y < x$.

So $x = \lambda < \kappa$ where λ is strongly inaccessible. But κ was the least strongly inaccessible cardinal, a contradiction. □

In the next section we'll see that the existence of a weakly inaccessible cardinal is enough to give us the consistency of ZFC. Hence ZFC cannot prove the existence of a weakly inaccessible cardinal either. A cardinal whose existence cannot be proved in ZFC but whose existence has not been shown to be inconsistent with ZFC is called a large cardinal: strongly inaccessible cardinals are large, as are weakly inaccessible cardinals. There is an exotic zoo of large cardinals: weakly compact, measurable, strongly compact, huge, not to mention Mahlo, weakly Mahlo, Woodin, n-extendibles.... Most large cardinals are defined in terms of embeddings from the universe into a proper subset of itself and cannot be understood without an extensive background in mathematical logic. For an out-of-date reference which remains excellent on the basics, the reader is referred to Drake; for a more up-to-date encyclopedic reference, the reader is referred to Kanamori. In the next chapter we'll say a little bit about the combinatorics of weakly compact cardinals and measurable cardinals.

Generally, large cardinals are those that imply the existence of a model of ZFC, and other hypotheses that imply the existence of a model of ZFC, such as "$0^\#$ exists"[88] can also be called large cardinal hypothesis.

6.2 The constructible universe

In this section we sketch the construction of and state some facts about the constructible universe L. Gödel discovered L on his way to proving that if ZF is consistent, so is ZFC. But it is Jensen's painstakingly detailed techniques that have uncovered so much of the structure not only of L but of classes constructed in similar fashion. The techniques derived from Jensen's work are grounded in concern for the level of complexity of definition of an object, and hence are grounded in fairly

[88] see the next section

advanced mathematical logic. Here we content ourselves with giving the bare definition of L and stating without proof some facts about it.

Definition 6.8. Given a set a and a formula Φ we define the formula Φ^a to be the formula derived from Φ by replacing each occurrence of "$\forall x_i$" in Φ by an occurrence of "$\forall x_i \in a$".[89]

Definition 6.9. b is definable from a iff there is a formula Φ with parameters in a so $b = \{x \in a : \Phi^a(x)\}$. Def($a$) is the set of all sets which are definable from a.

Note that the same formula may define different sets, depending on a. For example, each a is definable from itself via the formula "$x = x$".

Note that this notion of definability is not the usual one. For example, we ordinarily define ω as the least infinite ordinal. But for ω to be in Def(a) for some a, you need a formula Φ with parameters in a so that $\{\text{finite ordinals}\} = \{x \in a : \Phi(x)\}$. I.e., it's not that ω satisfies a formula, it's that its elements do.

Proposition 6.10. *(a) The set of even natural numbers is definable from ω.*

(b) If a is transitive, then every element in a is definable from a.

(c) If α is an ordinal, $\alpha \subset a$, and a is transitive, then α is definable from a.

(d) If $x \in a$ then $\{x\}$ is definable from a.

(e) If $x, y \in a$ then $\{x, y\}$ is definable from a.

(f) If $x \in a$ and a is transitive, then $\bigcup x$ is definable from a.

Proof. We give the formula Φ:[90]

For (a): $\exists x_0\ x = x_0 + x_0$

For (b): Let $x \in a$, a transitive, Φ is: $x_0 \in x$.

For (c): there are two cases: *Case 1.* $\alpha \in a$. This follows from (b) *Case 2.* $\alpha \notin a$. Then, since a is transitive, $\alpha = ON \cap a$, so Φ is: x_0 is an ordinal.

For (d): $x_0 = x$.

For (e): $x_0 = x$ or $x_0 = y$.

For (f): $\exists x_0 \in x\ x_1 \in x_0$. □

Def(a) is necessarily small.

Proposition 6.11. $|a| \leq |Def(a)| \leq \omega \cdot |a|$.

Proof. A formula Φ with parameters in a can be represented as a finite sequence whose elements are either elements in a (the parameters) or elements in the language of set theory (which is countable). Hence there are at most $\omega \cdot |a| = \sup\{|a|, \omega\}$ many formulas with parameters in a. There can't be more sets definable from a than formulas with parameters in a, so $|\text{Def}(a)| = \omega \cdot |a|$. □

[89] hence, replacing each "$\exists x_i$" by "$\exists x_i \in a$.

[90] and are careful about using $x_n, n < \omega$ to denote variables, reserving x, y, z, etc. for parameters, i.e., names of specific but unspecified sets.

Corollary 6.12. *If a is infinite, then Def$(a) \neq \mathcal{P}(a)$.*

In fact, if a is infinite, then $|\mathcal{P}(a) \setminus \mathrm{Def}(a)| = 2^{|a|} > |\mathrm{Def}(a)| = |a|$. We will construct L by using the hierarchy given by definability. By corollary 6.12 the levels will grow very slowly.

Definition 6.13. $L_0 = \emptyset$. If $\alpha = \beta + 1$ then $L_\alpha = \mathrm{Def}(L_\beta)$. If α is a limit, then $L_\alpha = \bigcup L_\beta$. $L = \bigcup_{\alpha \in ON} L_\alpha$.

Definition 6.13 has an absolute quality to it. All the transitive models in which x appears agree on Def(x). I.e., Def(x) does not vary from model to model the way $\mathcal{P}(x)$ might. This is the basic understanding behind the following theorem due to Gödel, whose proof we do not give:

Theorem 6.14. *Every transitive class containing all ordinals which is a model of ZF contains L as a subclass.*

For this reason, L is known as an inner model.

Let's explore some of the basic properties of L.

Proposition 6.15. *Each L_α is transitive.*

Proof. Suppose L_β is transitive for each $\beta < \alpha$. If α is a limit, we are done, so suppose $L_\alpha = \mathrm{Def}(L_\beta)$ for some β. If $y \in x \in L_\alpha$, $x \subseteq L_\beta$, so $y \in L_\beta$. By proposition 6.10(c), $y \in L_\alpha$. □

Proposition 6.16. *Every ordinal is an element of L.*

Proof. By proposition 6.10(b), it suffices to show that every ordinal is a subset of L. By induction, suppose every $\beta < \alpha$ is a subset of L. If α is a limit, we're done. If $\alpha = \beta + 1$ for some β, then $\beta \in L_\gamma$ for some γ, hence $\{\beta\} \in L_{\gamma+1}$ by proposition 6.10(d), so $\alpha \subset L_{\gamma+1} \subset L$. □

The next proposition says that L correctly identifies ordinals.

Proposition 6.17. *α is an ordinal iff $L \models \alpha$ is an ordinal.*

Proof. L is transitive and contains ON, so α is transitive and well-ordered by $\in$ iff $L \models \alpha$ is transitive and well-ordered by $\in$. □

As we'll see below, proposition 6.17 has no parallel for cardinals.

Gödel showed that $L \models \mathrm{ZFC} + \mathrm{GCH}$. We content ourselves with proving some simple fragments of this theorem.

Proposition 6.18. *$L \models$ extensionality, regularity, infinity, pairing, union, and power set.*

Proof. Regularity, extensionality and infinity are an exercise. Pairing and union follow from proposition 6.10(e) and (f). For power set, suppose $a \in L$. By the axiom of replacement there is some ordinal β so $\mathcal{P}(a) \cap L \subset L_\beta$. By the formula "$x \subseteq a$", $\mathcal{P}(a) \cap L \in L_{\beta+1}$. □

We sketch the proof that $L \models$ GCH:

First note that by proposition 6.11, each $|L_{\beta+1}| = |L_\beta|$. Hence, by induction, if $|\alpha| = |\kappa|$ then $|L_\alpha| = |L_\kappa|$, and $|L_{\kappa^+}| = |L_\kappa|^+$. Hence, by induction, each $|L_\kappa| = \kappa$. The final step is to prove that $\mathcal{P}(\kappa) \cap L \subseteq L_{\kappa^+}$.

Hence, by fact 6.3

Proposition 6.19. *If $L \models \kappa$ is weakly inaccessible, then $L \models \kappa$ is strongly inaccessible.*

A cardinal κ in the real universe V is always a cardinal in L: L can't have a function from a smaller ordinal α onto κ if there isn't one in V. Similarly, a weakly inaccessible cardinal in V is weakly inaccessible (hence strongly inaccessible) in L.

Hence, if you carefully inspect the proof (which we didn't give) that $L \models$ ZFC, you can prove

Theorem 6.20. *If κ is weakly inaccessible, then L_κ is a model of ZFC and GCH.*

By theorem 6.20, you cannot prove the existence of a weakly inaccessible cardinal in ZFC. (The details are an exercise.)

The main use of L is that it models not only GCH but many other useful combinatorial statements, where useful means: can be used to show a wide variety of mathematical statements consistent.[91] The question then becomes: does $V = L$?

The statement "$V = L$" is known as the axiom of constructibility. It is called an axiom not because there is general agreement it is true — I doubt that anyone would make a claim for its philosophically self-evident nature. Instead, the word "axiom" is used to indicate a statement which makes a fundamental claim about the nature of the mathematical universe, a claim which cannot be refuted within ZFC; whether we believe the claim to be true or not is irrelevant. CH is such an axiom. In the next chapter, we'll discuss Martin's axiom. And the assumption that there is, for example, a weakly inaccessible cardinal, or a measurable cardinal, or... is called a large cardinal axiom.

Since $L \models$ ZFC + $V = L$, the axiom of constructibility cannot be refuted in ZFC: just work inside the little piece of the unvierse we call L — how could you know anything else existed? But there are many models in which $V = L$ fails.[92]

There are two basic techniques for constructing models of $V \neq L$. One is to find models of statements which are false in L. For example, $L \models$ CH, so if you construct a model of $\neg$ CH (which Cohen did in 1963), you've constructed a model of $V \neq L$. Cohen's method, called forcing, has been explosively developed in the years since his work. It is firmly based in mathematical logic, and is beyond the scope of this book. The standard basic reference is Kunen.

The other technique is to show that ordinals which L thinks are cardinals actually aren't. How is this possible? Consider the α that L thinks is ω_1 (denoted ω_1^L). I.e., $L \models$ "$\forall \beta < \omega_1^L$ β is countable", and $L \models$ "ω_1^L isn't countable". But it might be that $V \models$ "ω_1^L is countable."

How can that happen? By proposition 6.17 we know that ω_1^L is an ordinal. $V \models$ "ω_1^L is countable" means that $\exists f : \omega \to \omega_1^L$, f onto. But no such f need exist in L, even if, $\forall \beta < \omega_1^L$ there is $f \in L$ $f : \omega \to \beta$, f onto.

[91] We will discuss one of these statements, $\diamondsuit$, in the next chapter.

[92] That is, if ZFC is consistent...

For example, there is a large cardinal hypothesis known as "$0^\#$ exists.". This is a highly model-theoretic statement: $0^\#$ is defined to be a particular set of formulas which completely describe a subclass of a certain kind of model of set theory.[93] While the existence of $0^\#$ is not known to be equivalent to saying that a large cardinal with simply stated specific combinatorial properties exists, it is implied by the existence of a measurable cardinal, and implies that there is a weakly inaccessible cardinal.[94] Silver proved

Theorem 6.21. *If $0^\#$ exists, then ω_1 is strongly inaccessible in L.*

Note that the ω_1 of theorem 6.21 is the real ω_1. If $0^\#$ exists, then not only is ω_1^L countable, but there are many ordinals that L thinks are cardinals between it and the real ω_1, in fact L thinks there are strongly inaccessibly many.

Another theorem about the extreme smallness of L under large cardinal hypotheses is Jensen's covering theorem, stated in the form: the failure of "$0^\#$ exists" is equivalent to L being reasonably large.

Theorem 6.22. *$0^\#$ does not exist iff every uncountable set of ordinals x is contained in a set of ordinals $y \in L$ where $|x| = |y|$.*[95]

Jensen proved his covering theorem because he was interested in the singular cardinals problem. By theorem 6.22, this is completely settled if $0^\#$ does not exist.

Corollary 6.23. *If $0^\#$ does not exist, then for every singular strong limit cardinal κ $2^\kappa = \kappa^+$.*

Proof. Let κ be a singular strong limit. By exercise 38 of chapter 5 it suffices to show that $\kappa^{\mathrm{cf}(\kappa)} = \kappa^+$. Suppose $x \subset \kappa, |x| = \mathrm{cf}(\kappa) = \lambda < \kappa$. By theorem 6.22, $x \subseteq y_x$ for some $y_x \in L$ where $|y_x| \leq \lambda^+ < \kappa$ and $y \subset \kappa$.[96] Since GCH holds in L, there are at most κ^+ many such y_x. By hypothesis, each $\mathcal{P}(y_x) < \kappa$. Whatever the real κ^+ is, it is $\geq (\kappa^+)^L$. ($(\kappa^+)^L$ is the way we denote the ordinal that L thinks is κ^+.) Hence $\kappa^\lambda \leq \kappa^+ \cdot 2^{\lambda^+} = \kappa^+$. And by theorem 5.55, $\kappa^\lambda \geq \kappa^+$. □

[93] and hence can be coded by a single real number. The set described is called the class of indiscernibles. Not every model contains such a class.

[94] There are better "lower bounds."

[95] $|x|$ and $|y|$ are the real cardinalities, not necessarily the cardinalities in L.

[96] The reason we write $|y_x| \leq \lambda^+$ is to cover the case $\lambda = \omega$.

6.3 Exercises

1. Prove fact 6.3.
2. Prove corollary 6.6 and apply it to $V_{\omega+\omega}$ and V_{ω_1}.
3. Show that $V_{\omega+\omega} \not\models$ replacement.
4. Show $\forall \alpha\ V_{\alpha+1} \not\models$ pairing.
5. Show that if κ is strongly inaccessible and $\lambda < \kappa$ then λ is regular iff $V_\kappa \models \lambda$ is regular.
6. Show that, if n is a finite cardinal, then $n \in \text{Def}(\omega)$.
7. Show that $\text{Def}(\emptyset) = \{\emptyset\}$.
8. (a) Show that if $a, b \in x$ and x is transitive, then $a \times b \in \text{Def}(\text{Def}(\text{Def}((x)))$.

   (b) Give an example where $a, b, \in x$ but $a \times b \notin \text{Def}(\text{Def}(\text{Def}((x)))$.
9. Show that if a is finite, then $\text{Def}(a) = \mathcal{P}(a)$.
10. Show that $L_\omega = V_\omega$ but $L_{\omega+1} \neq V_{\omega+1}$.
11. (a)Show that $L_\alpha \in L_\beta$ iff $\alpha < \beta$.

    (b) Show that each $L_\alpha \subseteq V_\alpha$.
12. Show that $L \models$ regularity, extensionality, and infinity.
13. Show that α is a limit ordinal iff $L \models \alpha$ is a limit ordinal.
14. (a) Show that if κ is regular then $L \models \kappa$ is regular.

    (b) Show that if κ is uncountable, then $L \models \kappa$ is uncountable.

    (c) Show that if κ is a limit cardinal then $L \models \kappa$ is a limit cardinal.

    (d) Show that if κ is a strong limit, then $L \models \kappa$ is a strong limit.
15. Show that you cannot prove the existence of a weakly inaccessible cardinal in ZFC.
16. Show that if $a \in L$ and $L \models |a| = \alpha$ then $\alpha \geq |a|$.

7 Semi-advanced set theory

In this chapter we use the techniques and ideas already developed to explore some areas of advanced set theory, mostly infinite combinatorics. The areas chosen are not exhaustive. In particular, we are leaving out major areas such as core models, descriptive set theory and determinacy, elementary submodels and Ω-logic (all of which are based in advanced mathematical logic); infinite games, pcf theory, and guessing principles other than $\diamondsuit$ (which are combinatorially complex).

The areas we will touch on (and that's all we're doing) are: partition calculus, trees, measurable cardinals, cardinal invariants of the reals , CH, Martin's axiom, stationary sets and $\diamondsuit$. The sections in this chapter are somewhat but not completely independent of each other. A reader who knows some finite combinatorics and graph theory should find some of the concepts in the first two sections familiar; a reader who knows some measure theory should find some of the concepts in the third section familiar.

There is a grand theme in this chapter, and that is the interrelatedness of the various topics. Thus, trees and partition calculus lead to the large cardinal property of strong compactness; trees are used to understand the real line; and the combinatorial principles CH, MA and $\diamondsuit$ shed different light on the same combinatorial issues.

7.1 Partition calculus

Here is a party trick: If you have at least six people together in a room, either three of them knew each other previously, or three of them were strangers to each other previously. Why is this? Suppose no three of them knew each other previously. Pick out a person, call him John, and the rest of the group divides into two sets: the people John knew previously, call it group I, and the people John didn't know before, call it group II. If two people in group II didn't know each other before, say Jane and Joan, then we are done, since John, Jane and Joan had all been strangers to each other previously. So we may assume that all the people in group II knew each other before. If there are three people in group II we're done, so we may assume that group II has at most two people in it, i.e., group I has at least three people in it. If two of the people in group I, say Sean and Ian, knew each other before, then Sean and Ian and John all knew each other before. So we may assume none of the people in group I knew each other before, but then we've got at least three people who didn't know each other before. Phew.

This party trick is the simplest example of a class of theorems known as Ramsey theorems, after the brilliant English mathematician who discovered them and died tragically young. To generalize this party trick it helps to develop some notation; this notation in turn inspires variations which give rise to some of the most difficult, interesting, and useful (as well as some of the most arcane) concepts in infinite combinatorics. Since all of these ideas have to do with partitions, their study is known as partition calculus.

We develop the notation by analyzing the party trick.

We have a property of (unordered) pairs of people: either they know each other or they do not. Thus we really have a partition of $[X]^2 = \{\{x, y\} : x, y \in X \text{ and } x \neq y\}$. This partition of $[X]^2$ has two pieces: "know each other," and "do not know each other." We are asking: can you find a three-element subset of X all of whose pairs lie in the same piece of the partition? Such a subset is known as a homogeneous set.

More generally:

Recall the definition of $[X]^\kappa$ in example 5.66: $[X]^\kappa = \{y \subseteq X : |y| = \kappa\}$. Similarly, $[X]^{<\kappa} = \{y \subseteq X : |y| < \kappa\}$; $[X]^{>\kappa} = \{y \subseteq X : |y| > \kappa\}$; $[X]^{\leq\kappa} = \{y \subseteq X : |y| \leq \kappa\}$; $[X]^{\geq\kappa} = \{y \subseteq X : |y| \geq \kappa\}$

Recall the definition of a partition in definition 1.22 and note that some elements of a partition may be empty.

Definition 7.1. If $[X]^\kappa$ is partitioned into sets $\{P_i : i \in I\}$, then $Y \subseteq X$ is homogeneous for the partition iff, for some i, $[Y]^\kappa \subseteq P_i$.

Thus the party trick can be rephrased as follows: if $|X| \geq 6$ and $[X]^2$ is partitioned into two sets, then there is some $Y \in [X]^3$, Y is homogeneous for the partition.

More compactly, we write this as $6 \to (3)^2_2$, meaning that if you partition the pairs (this is the upper 2) of a set of six elements into two pieces (this is the lower 2) you will have a homogeneous subset with at least three elements.

Here are some examples of homogeneous and non-homogeneous sets.

Example 7.2. Partition $[\omega]^2$ into two pieces: the pairs whose product is even, and the pairs whose product is odd. Any set consisting solely of even numbers is homogeneous. Any set consisting solely of odd numbers is homogeneous. A set with exactly one odd element and the rest even elements is homogeneous. Any set with at least one even and at least two odd numbers is non-homogeneous.

Example 7.3. Partition $[\mathbb{R}]^3$ into two pieces: $P_0 = \{\{x, y, z\} \in [\mathbb{R}]^3 : x, y, z \text{ are collinear}\}$. $P_1 = \{\{x, y, z\} \in [\mathbb{R}]^3 : x, y, z \text{ are not collinear}\}$. Straight lines are homogeneous fo P_0. Circles are homogeneous for P_1. Parabolas are homogeneous for P_1. Cubics are not homogenous.

Now we generalize the arrow notation.

Definition 7.4. Let $\kappa, \lambda, \rho, \sigma$ be cardinals. $\kappa \to (\lambda)^\rho_\sigma$ iff for every partition of $[\kappa]^\rho$ into σ pieces, there is a homogeneous subset of size λ.[97]

Ramsey's theorem for finite sets is that for all finite j, m, k with $j \geq m$ there is an n with $n \to (j)^m_k$. Finding the least possible n can be quite difficult, and if you fix m, k and vary j, n will grow very quickly.

The infinite version of Ramsey's theorem is

Theorem 7.5. *For all finite $n, m, \omega \to (\omega)^n_m$.*

Before proving Ramsey's theorem, we will give an application of it and a few easy facts involving arrow notation.

Theorem 7.6. *Every infinite partial order has either an infinite antichain or an infinite set of pairwise compatible elements.*

[97] One rich generalization of Ramsey theory (which we leave out) is using order types instead of cardinality for κ, λ.

Proof. Let A be an infinite countable subset of the given partial order, and let $[A]^2 = P_0 \cup P_1$ where $\{x, y\} \in P_0$ iff x, y are incompatible; otherwise $\{x, y\} \in P_1$. Let H be an infinite homogeneous set for this partition. If $[H]^2 \subseteq P_0$ then H is an infinite antichain; if $[H]^2 \subseteq P_1$ then H is an infinite set of pairwise compatible elements. □

Theorem 7.7. *An infinite partial order with no infinite antichain has either an infinite chain, or an infinite pairwise incomparable pairwise compatible set.*

Proof. Let A be an infinite countable subset of the given partial order, and let $[A]^2 = P_0 \cup P_1 \cup P_2$ where $\{x, y\} \in P_0$ iff x, y are incompatible; $\{x, y\} \in P_1$ iff x, y are comparable; $P_2 = [X]^2 \setminus (P_0 \cup P_1)$. Let H be an infinite homogeneous set for this partition. $H^2 \not\subseteq P_0$, so either $H^2 \subseteq P_1$ (and we have an infinite chain) or $[H]^2 \subseteq P_2$ and we have an infinite pairwise incomparable pairwise compatible set. □

The exercises give several more applications of Ramsey's theorem.

Proposition 7.8. *Let $\kappa, \lambda, \rho, \sigma, \tau$ be cardinals and suppose $\kappa \to (\lambda)^\rho_\sigma$. Then*

(a) If $\tau > \kappa$ then $\tau \to (\lambda)^\rho_\sigma$.

(b) If $\tau < \lambda$ then $\kappa \to (\tau)^\rho_\sigma$.

(c) If $\tau < \sigma$ then $\kappa \to (\lambda)^\rho_\tau$.

The proof is left to the exercises.

Note that for each infinite κ, $\kappa \to (\kappa)^1_2$ and, for any cardinal $\lambda \leq \kappa$, $\kappa \to (\kappa)^\lambda_1$. Thus the simplest nontrivial arrow relation on an infinite κ is $\kappa \to (\kappa)^2_2$. We prove some results about this relation.

Proposition 7.9. *If κ is infinite and $\kappa \to (\kappa)^2_2$ then $\kappa \to (\kappa)^2_m$ for all $m \in \omega$.*

Proof. The proof is by induction on m. If $\kappa \to (\kappa)^2_{m-1}$ then, given a partition $P_1, ...P_m$ of $[\kappa]^2$ into m pieces, let $P^*_1 = P_1 \cup P_2$, and for $i > 2$ let $P^*_i = P_{i+1}$. By induction hypothesis, there is a homogeneous set $Y \in [\kappa]^2$ and i with $[Y]^2 \subseteq P^*_i$. If $[Y]^2 \subseteq P^*_i$ for some $i > 1$, we're done. Otherwise we apply $\kappa \to [\kappa]^2_2$ to Y to get a set $Z \in [Y]^\kappa$ which is homogeneous either for P_1 or P_2. □

Which cardinals satisfy the hypothesis of proposition 7.9? Our first result in this direction is

Proposition 7.10. $\omega \to (\omega)^2_2$.

Proof. Suppose $[\omega]^2 = P_1 \cup P_2$ where $P_1 \cap P_2 = \emptyset$. We recursively build a sequence of natural numbers $\{k_j : j < \omega\}$ and a sequence of infinite sets $\{A_j : j < \omega\}$ so that each $A_{j+1} \subseteq A_j$, each $k_j \in A_j$, and $\forall j\ \exists i\ A_{j+1} \subseteq \{m > k_j : \{k_j, m\} \in P_i\}$.

How do we do this? Let $k_0 = 0$ and $A_0 = \omega$. If we have $A_j, k_j, j \leq m$ satisfying the above requirements, notice that there is some i so that $A = \{k \in K_m : k > k_m \text{ and } \{k_m, k\} \in P_i\}$ is infinite. Let $A_{m+1} = A, k_{m+1} \in A_{m+1}$ and continue.

Now that we have the k_j's, define $f(j) = i$ iff for all $m \in A_{j+1}, \{k_j, m\} \in P_i$. $f : \omega \to \{1, 2\}$ and, since its range is finite, f is constantly equal to some i on some infinite set B. But then if $j, m \in B, \{k_j, k_m\} \in P_i$, so $\{k_j : j \in B\}$ is homogeneous. □

Hence, by proposition 7.9, $\omega \to (\omega)^2_m$ for all $n < \omega$.

The proof of Ramsey's theorem will be completed by proving

Theorem 7.11. *If $\omega \to (\omega)^2_2$ then $\omega \to (\omega)^n_m$ for all $n, m < \omega$.*

Proof. By proposition 7.9 it suffices to fix m and work by induction on n. So suppose $\omega \to (\omega)^n_m$ and let $\mathcal{P} = \{P_1, ..P_m\}$ be a partition of $[\omega]^{n+1}$. For each $k \in \omega$ define $\mathcal{P}_k$ a partition of $[\omega \setminus \{k\}]^n$ as follows:

$\mathcal{P}_k = \{P_{k,1}...P_{k,m}\}$ where $P_{k,i} = \{s \in [\omega \setminus \{k\}]^n : s \cup \{k\} \in P_i\}$. Note that $\mathcal{P}_k$ is a partition of $[\omega \setminus \{k\}]^n$ as promised.

Again we have sequences $\{k_j : j < \omega\}, \{A_j : j < \omega\}$ with the following slightly more demanding requirements: each $A_{j+1} \subseteq A_j$ infinite; each $k_j \in A_j$, each A_j is homogeneous for each $\mathcal{P}_{k_r}$, $r \leq j$.

How is this done? Given $A_j, k_j \in A_j$, let $B_{j,0} \subseteq A_j$ be infinite homogeneous for $\mathcal{P}_{k_0}$, $B_{j,1} \subseteq B_{j,0}$ be infinite homogeneous for $\mathcal{P}_{k_1}$, etc. Then $B_{j,j}$ is infinite homogeneous for all $\mathcal{P}_{k_r}$ with $r \leq j$, so define $A_{j+1} = B_{j,j}$.

Define $f(j) = i$ iff $A_{j+1} \subseteq P_{k_j,i}$ and note that there is some i so $B = \{j : f(j) = i\}$ is infinite. Again, $\{k_j : j \in B\}$ is homogeneous for the original partition $\mathcal{P}$. □

Proposition 7.10 showed that $\omega \to (\omega)^2_2$. Are there any other cardinals with this property?

Definition 7.12. An uncountable cardinal κ is a weakly compact cardinal iff $\kappa \to (\kappa)^2_2$.

It turns out that if κ is weakly compact then κ is a regular strong limit cardinal,[98] hence, if uncountable, strongly inaccessible. Our immediate goal is to prove this. On the way, we'll prove some negative partition properties.

Theorem 7.13. *If κ is weakly compact then κ is regular.*

Proof. Suppose κ is weakly compact. Let $\{\kappa_\alpha : \alpha < \mathrm{cf}(\kappa)\}$ be an increasing sequence of cardinals cofinal in κ. For each ordinal $\beta < \kappa$ we write $f(\beta) =$ the least α with $\beta < \kappa_\alpha$ and let $P_1 = \{\{\gamma, \beta\} \in [\kappa]^2 : f(\gamma) = f(\beta)\}; P_2 = \{\{\gamma, \beta\} \in [\kappa]^2 : f(\gamma) \neq f(\beta)\}$. Let $Y \in [\kappa]^\kappa, Y$ homogeneous. Y cannot be homogeneous for P_1, so it must be homogeneous for P_2, i.e., $\mathrm{cf}(\kappa) = \kappa$. □

The next theorem says that the cardinality of a power set fails to be weakly compact.

Theorem 7.14. $\forall \lambda\ 2^\lambda \not\to (\lambda^+)^2_2$.

Proof. First we prove a subclaim of independent interest:

Subclaim 7.14.1. There are no increasing or decreasing chains of size λ^+ in 2^λ under $\leq_L$

[98] like ω

Proof. Suppose we have $G = \{g_\alpha : \alpha < \lambda^+\} \subseteq 2^\lambda$ where if $\alpha < \beta$ then $g_\alpha <_L g(\beta)$. For each α define γ_α to be the least $\gamma < \lambda$ so that, for some β, $g_\alpha|_\gamma = g_\beta|_\gamma$ and $0 = g_\alpha(\gamma) < g_\beta(\gamma) = 1$. Let β_α be the least such β.

Let $H_\gamma = \{\alpha : \gamma_\alpha = \gamma\}$, and suppose $\alpha, \alpha^* \in H_\gamma$. By definition, $g_\alpha|_\gamma = g_{\beta_\alpha}|_\gamma = g_{\alpha^*}|_\gamma$. By the preceding paragraph $g_\alpha(\gamma) = g_{\alpha^*}(\gamma) = 0 < g_{\beta_\alpha}(\gamma) = 1$. So $g_{\alpha^*} < g_{\beta_\alpha}$.

Hence $\forall \alpha, \alpha^* \in H_\gamma$ $\beta_\alpha = \beta_{\alpha^*}$. Let δ_γ be this β_α.

Hence each $H_\gamma \subseteq \delta_\gamma$. So each $|H_\gamma| \leq \lambda$. But $\lambda^+ = \bigcup_{\gamma<\lambda} H_\gamma$ a contradiction.

A similar proof shows that there is no $\{g_\alpha : \alpha < \lambda^+\} \subseteq 2^\lambda$ where if $\alpha < \beta$ then $f_\alpha >_L f_\beta$. □

Now suppose $\{f_\alpha : \alpha < 2^\lambda\}$ lists the elements of 2^λ. We partition $[2^\lambda]^2$ as follows: $P_1 = \{\{f_\alpha, f_\beta\} : \alpha < \beta \text{ and } f_\alpha <_L f_\beta\}$; $P_2 = \{\{f_\alpha, f_\beta\} : \alpha < \beta \text{ and } f_\alpha >_L f_\beta\}$. By the subclaim, neither P_1 nor P_2 has a homogeneous set of size λ^+. □

By theorem 7.13 and theorem 7.14

Corollary 7.15. *A weakly compact cardinal is a regular strong limit cardinal.*

Proof. We already know that weakly compact cardinals are regular. Suppose κ is not a strong limit. Then there is $\lambda < \kappa$ with $2^\lambda \geq \kappa$. If κ were weakly compact, then by proposition 7.8 $\kappa \to (\lambda^+)^2_2$ hence $2^\lambda \to (\lambda^+)^2_2$, contradicting theorem 7.14. □

Theorem 7.14 is especially interesting in light of the Erdös-Rado theorem, which we will not prove here.

Theorem 7.16. *(Erdös-Rado)* $(2^\lambda)^+ \to (\lambda^+)^2_2$.

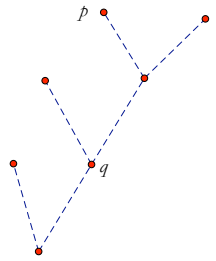
Thus, we know exactly which cardinals κ satisfy $\kappa \to (\lambda^+)^2_2$.

In the next section we will see that weakly compact cardinals have other interesting combinatorial properties.

7.2 Trees

Definition 7.17. A tree is a partially ordered set in which the predecessors of each element are well-ordered, i.e., each $t\downarrow = \{s : s < t\}$ is well-ordered.

Here's a picture of a piece of a tree. The points are in the tree. The lines are not in the tree, but show the order relation, i.e., if a line connects q to p and q is below p then $q < p$.



We introduce more terminology.

Definition 7.18. Let T be a tree, $t \in T$. The height of t, ht t, is the order type of $t\downarrow$. The height of T, ht T, is the smallest α so $\forall t \in T$ ht $t < \alpha$. $T(\alpha)$, called the α^{th} level of T, is the set of elements of T with height α. $T_\alpha = \bigcup_{\beta<\alpha} T(\beta)$. Each T_α is called an initial segment of T.

For example, if the above diagram were a complete tree, then ht $q = 1$ and height $p = 3$.

A useful observation is:

Fact 7.19. *If $t \in T(\alpha)$, $t < s$, $s \in T(\beta)$, and $\alpha < \gamma < \beta$ then $\exists r \in T(\gamma)$ $t < r < s$*

More terminology:

Definition 7.20. Let T be a tree, $p, q \in T$. p is a successor of q iff $p > q$ iff q is a predecessor of p; p is an immediate successor of q iff $p > q$ and ht $p =$ ht $q + 1$ iff q is an immediate predecessor of p.

I.e., in the diagram above, p is a successor but not an immediate successor of q and q is a predecessor but not an immediate predecessor of p.

Example 7.21. For any x, κ define $T = \bigcup_{\beta<\kappa} x^\beta$ under the ordering of extension: $t \leq s$ iff $t \subseteq s$. Each $T(\alpha) = x^\alpha$ and each $T_\alpha = \bigcup_{\beta<\alpha} x^\beta$. If $x = 2$ we call this the binary tree of height κ; if $\lambda = \kappa$, we call this the
lambda-branching tree of height κ.

Example 7.22. Let $f : \omega \to \omega$, let $x = \{g : \omega \to \omega : g =^* f\}$. Let $T = x \cup \bigcup_{n<\omega} \omega^n$, ordered by extension. Then T has height $\omega + 1$, T is countable, and $x = T(\omega)$.

Fact 7.23. *Every subset of a tree is a tree under the induced order.*

Recall the definitions of chain and antichain in definition 1.11. Thus every $T(\alpha)$ is an antichain.

Definition 7.24. A maximal chain in a tree is called a branch.

In the tree of example 7.21, each branch has the form $\{f|_\alpha : \alpha < \kappa\}$ for some $f : \kappa \to x$; and in this same tree, if, for all $y \in x$, we choose some t_y with $t_y(17) = y$, then $\{t_y : y \in x\}$ is an antichain.

The following definition is an abuse of notation that makes sense if you think of a tree as hanging upside down.

Definition 7.25. Two elements of a tree are incompatible iff they are incompatible under $\leq^{\leftarrow}$, where $p \leq^{\leftarrow} q$ iff $p \geq q$; otherwise they are compatible.

Theorem 7.26. *Two elements in a tree are comparable iff they are compatible.*

Proof. Note that p, q are comparable under $\leq$ iff they are comparable under $\leq^{\leftarrow}$. Note that in a partial order comparability implies compatibility. So assume p, q are compatible. There is r with $p \leq r, q \leq r$. If either $p = r$ or $q = r$ then p, q are comparable, so assume $p < r, q < r$. Then $p, q \in r^{\downarrow}$ which is well-ordered, hence linear ordered, and p, q are comparable. □

The main focus of this section is the relation between the length of a tree's branches and the size of its levels or, more generally, of its antichains. These rather innocent-looking questions will quickly involve us in consistency results and large cardinals.

The first property we look at is

Definition 7.27. A cardinal κ has the tree property iff for every tree T whose levels and branches each have size $< \kappa$, $|T| < \kappa$.

Equivalently

Fact 7.28. *κ has the tree property iff every tree of cardinality κ whose levels each have cardinality $< \kappa$ has a branch of cardinality κ.*

Proposition 7.29. *Singular cardinals do not have the tree property.*

Proof. Let κ be singular, $\lambda = \mathrm{cf}(\kappa) < \kappa$, and let $\{\kappa_\alpha : \alpha < \lambda\}$ be an increasing sequence whose union is κ. For $\alpha < \lambda$ define $f_\alpha : \kappa_\alpha \to \{\alpha\}$. Let $T = \bigcup_{\alpha<\lambda}\{f_\alpha|_\beta : \beta < \kappa_\alpha\}$ ordered by extension. T has height κ, each $|T(\alpha)| = \lambda < \kappa$, so $|T| = \lambda \cdot \kappa = \kappa$. The branches of T are all $b_\alpha = \{f_\alpha|_\beta : \beta < \kappa_\alpha\}$, so no branch of T has cardinality κ. □

On the other hand

Theorem 7.30. *(König) ω has the tree property.*

Proof. Suppose T is an infinite tree with every level finite. By induction we construct an increasing sequence $\{t_n : n < \omega\}$ where each $t_n \in T(n)$ and $\forall n\ t_n$ has infinitely many successors.

Suppose we have $t_0, ...t_n$. Since t_n has infinitely many successors, it has at least one successor in $T(n+1)$. Since it has only finitely many successors in $T(n+1)$, at least one of them has infinitely many successors. So we can find $t_{n+1} \in T(n+1)$, $t_{n+1} > t_n$, t_{n+1} has infinitely many successors. □

Theorem 7.30 is known as König's lemma.[99]

[99]The König of König's theorem was Gyula König; the König of König's lemma was his son, Dénes König.

Theorem 7.30 is handy when trees must be constructed ad hoc in the course of a proof. For example, if you have a tree of sets with all levels finite, ordered so that $s < t$ implies $t \in s$ (hence, by regularity, no infinite branches) then you know your tree is finite.

At this point we know that ω has the tree property and that no singular cardinal does. What about ω_1? The young Aronszajn, taking a quick side-trip from his usual preoccupation of functional analysis, proved

Theorem 7.31. *(Aronszajn) ω_1 does not have the tree property.*

The proof of theorem 7.31 is to construct a counterexample. Not surprisingly, these counterexamples are known as Aronszajn trees.

Definition 7.32. A tree is Aronszajn iff it has height ω_1 but all levels and branches are countable.

We can restate theorem 7.31 as

Theorem 7.33. *There is an Aronszajn tree.*

Proof. There are, as it turns out, many ways to construct Aronszajn trees. Here are two of them.

Construction I. (Aronszajn) Our Aronszajn tree consists of well-ordered (under the usual ordering) subsets of $\mathbb{Q}$, with the order as end-extension: $\sigma < \tau$ iff σ is an initial segment of τ. Thus there are no uncountable branches. We have to keep the tree pruned in order to ensure each level is countable, but if we prune it too much, its height will be countable.

Before stating the requirements, we need some notation. Given σ a well-ordered subset of $\mathbb{Q}$, we write $q_\sigma = \sup \sigma$, and given $q \in \mathbb{Q}$ with $q > r$ for all $r \in \sigma$, we write $\sigma^\frown q$ for σ followed by q.

The requirement that helps prune the tree is: (*) if $\sigma \in T$ then $q_\sigma \in \mathbb{Q}$.

The requirement that allows us to continue is: (**) if $\gamma < \beta, \sigma \in T_\gamma, q \in \mathbb{Q}, q > 0$ then $\exists \tau \in T(\beta)\ \sigma < \tau, q_\tau = q + q_\sigma$.

Suppose we have T_α satisfying (*), T_α satisfies (**) for $\gamma < \beta < \alpha$, and T_α satisfies $\forall \beta < \alpha\ T(\beta)$ is countable.

If, for some β, $\alpha = \beta + 1$, then define $T(\alpha) = \{\sigma^\frown q + q_\sigma : \sigma \in T(\beta),\ q \in \mathbb{Q}, q > 0.\}$. $T(\alpha)$ is countable, (*) holds, and (**) holds below $\alpha + 1$.

Suppose α is a limit. Then there is an increasing sequence $\{\alpha_n : n < \omega\}$ with $\alpha = \sup_{n<\omega} \alpha_n$. Similarly, for each $q \in \mathbb{Q}$ with $q > 0$ there is an increasing sequence of positive rationals $\{p_{n,q} : n < \omega\}$ converging to q.

Let $T_\alpha = \{\sigma_k : k < \omega\}$. Given σ_k, q, by induction hypothesis we can fix a sequence $\vec{\tau}_{k,q} = \{\tau_{k,n,q} : n < \omega\} \subseteq T_\alpha$ so $\sigma_k \leq \tau_{k,0,q}$, each $\tau_{k,n,q} < \tau_{k,n_1,q}$, each $q_{\tau_{k,n,q}} = p_{n,q} + q_{\sigma_k}$, and $\tau_{k,n,q} \notin T_{\alpha_n}$. Define $T(\alpha) = \{\bigcup \vec{\tau}_{k,q} : k \in \omega, q \in \mathbb{Q}, q > 0\}$. Again, $T(\alpha)$ is countable, (*) holds, and (**) holds below $\alpha + 1$.

Finally, let $T = \bigcup T(\alpha)$. T is an Aronszajn tree.

Construction II (Shelah) Now our Aronszajn tree consists of infinite subsets of ω, ordered by inverse inclusion: $A < B$ iff $A \supset B$.

There are two requirements:

(*) If $A, B \in T(\alpha)$ then $A \cap B$ is finite.

(**) If $A \in T_\beta$ and $a \in [A]^{<\omega}$ then $\exists B \in T(\beta)\ a \subset B \subset A$.

Again the tree is constructed by induction. So suppose we have T_α meeting all relevant requirements.

Subclaim 7.33.1. If $A \in T_\alpha, a \in [A]^{<\omega}$ and $\alpha > \beta >$ ht A, then there are infinitely many $B \in T(\beta)$ with $a \subset B \subset A$.

Proof. Given A, a, β, let $\mathcal{A} = \{B \in T(\beta) : a \subset B \subset A\}$. If $\mathcal{A}$ is finite, for each $B \in \mathcal{B}$ let $k_B \in B \setminus (a \cup \bigcup (A \setminus \{B\}))$. There is $C \in T(\beta)$ with $a \cup \{c_B : B \in \mathcal{A}\} \subset C \subset A$. $C \notin \mathcal{A}$. □

If, for some β, $\alpha = \beta + 1$, then for every $A \in T(\beta)$ let $\{B_{n,A} : n < \omega\}$ be a partition of A into infinite sets. Let $[A]^{<\omega} = \{a_n : n < \omega\}$, and let $T(\beta + 1) = \{B_{n,A} \cup a_n : A \in T(\beta)\}$.

If α is a limit, let $\{\alpha_n : n < \omega\}$ be an increasing sequence converging to α. Enumerate $\{(A, a) : A \in T_\alpha, a \in [A]^{<\omega}\}$ as $\{(A_n, a_n) : n < \omega\}$. Suppose, for each $m < n$ we have $b_{m,\alpha}$ a branch in T_α where $a_m \subset \bigcap b_{m,\alpha} \subset A_m$, each $\bigcap b_{m,\alpha}$ is infinite, and if $m \neq m'$ then either $b_{m,\alpha} = b_{m',\alpha}$ or $\bigcap b_{m,\alpha} \cap \bigcap b_{m',\alpha}$ is finite.

Consider A_n, a_n. If, for some $m < n$, $a_n \subset \bigcap b_{m,\alpha} \subset A_n$, let $b_{m,\alpha} = b_{n,\alpha}$. Otherwise, let k be least with ht $A_n \leq \alpha_k$. Let $B_{0,n,\alpha} \in T(\alpha_k \setminus \bigcup_{m<n} b_{m,\alpha}$ with $a_n \subset B_{0,n,\alpha} \subset A_n$ and choose $r_0 \in B_{0,n,\alpha} \setminus (a \cup \bigcup_{m<n} \bigcap b_{m,\alpha})$. Recursively contstruct $B_{j+1,n,\alpha} \in T(\alpha_{k+j+1} \setminus \bigcup_{m<n} b_{m,\alpha}$ with $a_n \subset B_{j+1,n,\alpha} \subset B_{j,n,\alpha}$. Let $r_{j+1} \in B_{j+1,n,\alpha} \setminus (a \cup \{r_s : s < j+1\} \cup \bigcup_{m<n} \bigcap b_{m,\alpha})$.

By construction, $b_{n,\alpha} = \{C \in T_\alpha : \exists j < \omega\ C \supset B_{j,n,\alpha}\}$ is a branch in T_α, $\bigcap b_{n,\alpha}$ is infinite, $\bigcap b_{m,\alpha} \cap \bigcap b_{n,\alpha}$ is finite for all $m < n$, and $a_n \subset \bigcap b_{n,\alpha} \subset A_n$. Let $B_{n,\alpha} = \bigcap_{j<\omega} B_{j,n,\alpha}$ and define $T(\alpha) = \{B_{n,\alpha} : n < \omega\}$.

$T = \bigcup_{\alpha<\omega_1} T(\alpha)$.

□

There are many kinds of Aronszajn trees. Here is one kind.

Definition 7.34. An uncountable tree is said to be special iff it is the countable union of antichains.

The tree we constructed in the first proof of theorem 7.33 is easily seen to be special: $A_q = \{\sigma \in T : q_\sigma = q\}$ is an antichain for each $q \in \mathbb{Q}$, and $T = \bigcup_{q\in\mathbb{Q}} A_q$. EATS is the statement: "every Aronszajn tree is special." Does EATS hold?

Definition 7.35. A Suslin tree is an uncountable tree with no uncountable branches and no uncountable antichains.

The Suslin hypothesis, SH, is: "there are no Suslin trees."

Suslin trees, if they exist, are definitely not special. So EATS + ¬SH is a contradiction.

A useful fact about Suslin trees is

Fact 7.36. *Every uncountable subset of a Suslin tree for is, under the induced order, a Suslin tree.*

The proof is an exercise.

So are there Suslin trees? It depends on the combinatorics of the particular model of set theory. In particular, we've already met CH and $V = L$. Later we will meet Martin's Axiom, MA. Here's a brief survey of what is known about SH and EATS:

Theorem 7.37. *(a) If $V = L$ then $\neg$ SH.*[100]

(b) If MA + $\neg$ CH then EATS.[101]

We will prove theorem 7.37 in the sections on $\diamondsuit$ and MA.

What about the relationship, if any, between SH and the size of the continuum?

Theorem 7.38. *The following are consistent:*

(a) CH + SH + $\neg$ EATS.[102]

(b) SH + $2^\omega = \aleph_{\omega_1}$.[103]

None of this could remotely have been imagined when Suslin, back in the 1920's, over ten years before Aronszajn constructed an Aronszajn tree, asked about what we have come to call Suslin trees.

Actually, he didn't ask about trees, he asked about lines, and we now turn our attention to the intimate relationship between trees and lines.

Suslin was thinking about the following properties of the real line:

(i) No collection of pairwise disjoint intervals is uncountable. (This is called the ccc property.)

(ii) There is a countable dense subset, i.e., a countable set D so every non-empty open interval intersects D.

Clearly (ii) implies (i). Do either of these properties alone characterize subsets of the real line? We will show below that

Theorem 7.39. *A linear order satisfies (ii) iff it is order-isomorphic to a subset of the reals.*

But does (i) suffice? Suslin's hypothesis, the way Suslin stated it, was that (i) does suffice, i.e., (i) implies (ii). $\neg$ SH, then, says that there is a linear order satisfying (i) and not (ii). Such an order is known as a Suslin line, and it clearly cannot embed as a suborder of the reals, since every subset of the reals has a countable dense subset.

The connection between lines and trees showing that both versions of SH are equivalent was worked out independently by a number of people with much duplication of effort over the next 20 years, and it is to this that we now turn our attention.

[100] This was proved by Jensen; Jech and Tennenbaum had earlier proved the consistency of $\neg$ SH.

[101] This was proved by Baumgartner, Reinhardt, and Malitz. Earlier, Solovay, Martin and Tennenbaum proved that under MA + $\neg$ CH, SH holds; MA itself is derived from an earlier proof of the consistency of SH by Solovay and Tennenbaum.

[102] Jensen's proof that CH + SH is consistent is seminal; Shelah proved SH + $\neg$ EATS consistent, and about ten years later Schlindwein proved theorem 7.38(a).

[103] This is due to Laver.

Definition 7.40. Let X be a linear order, $A \subseteq X, x, y \in X, x < y$.

(a) $(x, y) = \{z : x < z < y\}$; its endpoints are x, y and if $z \in (x, y)$ then we say (x, y) is an interval around z.

(b) $x \in \text{cl } A$ iff every interval around x contains an element of A. (Here, cl A is called the closure of A.)

(c) A is dense in X iff $X = \text{cl } A$.

Recall from chapter 1 that X is a dense linear order iff it is a linear order and if $x < y$ then $(x, y) \neq \emptyset$.

Note the distinct meanings of "dense."

Before proving theorem 7.39 we prove

Theorem 7.41. *(Cantor) Every nonempty countable dense linear order without endpoints is order-isomorphic to* $\mathbb{Q}$.

Proof. Let X be a countable linear order, $X = \{x_n : n < \omega\}$. Let $\mathbb{Q} = \{q_n : n < \omega\}$. We construct an order isomorphism $\varphi : X \to \mathbb{Q}$ as follows:

Suppose we have $\varphi : X_n \to Q_n$ an order-isomorphism, where $X_n \in [X]^{\leq 2n}, Q_n \in [Q]^{\leq 2n}$, and if $i < n$ then $x_i \in X_n$ and $q_i \in Q_n$.

Consider x_n. If $x_n \notin X_n$, let $S_n = \{x \in X_n : x < x_n\}$ and $T_n = \{x \in X_n : x > x_n\}$. There is $q \in \mathbb{Q} =$ with $\varphi(\sup S_n) < q < \varphi(\inf T_n)$.[104] Assign $\varphi(x_n) = q$ and put $x_n \in X_{n+1}, q \in Q_{n+1}$.

Consider q_n. If $q_n \notin Q_n, q_n \neq \varphi(x_n)$, let $R_n = \{q \in Q_n \cup \{\varphi(x_n)\} : q < q_n\}$ and $W_n = \{q \in Q_n \cup \{\varphi(x_n)\} : q > q_n\}$. There is $x \in X$ with $\varphi(\sup R_n) < x < \varphi(\inf W_n)$.[105] Assign $\varphi(x) = q_n$ and put $x \in X_{n+1}, q_n \in X_{n+1}$. □

The technique of this proof is known as the back-and-forth argument, and is used extensively in model theory. It is the third major technique of Cantor we have seen.

Corollary 7.42. *Every countable dense linear ordering X is isomorphic to a subset of* $\mathbb{Q}$.

Proof. Given a countable dense linear ordering, extend it (by countable recursion; see exercise 24) to a countable dense linear order without endpoints. Apply theorem 7.41. □

A careful analysis shows the following:

Proposition 7.43. *Consider the proof of theorem 7.41. If $\{x_n : n < \omega\}$ is an increasing sequence in X, $\{y_n : n < \omega\}$ is a decreasing sequence in X, each $x_n < y_m$, and $\bigcap_{n<\omega}(x_n, y_n) = \emptyset$ then $\bigcap_{n<\omega}(\varphi(x_n), \varphi(y_n)) = \emptyset$.*

Now we are ready to prove theorem 7.39, that a linear order has a countable dense set iff it order-isomorphic to a subset of $\mathbb{R}$.

[104] hence $q \notin Q_n$.
[105] hence $x \notin X_n \cup \{x_n\}$.

Proof. First note that every subset of $\mathbb{R}$ has a countable dense set.

Now let X be a linear order with a countable dense subset D. Without loss of generality, if X has a minimum d,then $d \in D$. Every point in X is the supremum of an initial segment of D. By corollary 7.42, D is order-isomorphic to $Y \subseteq \mathbb{Q}$, where $\varphi : Y \to D$ is the order-isomorphism. By proposition 7.43, if $x = \sup A = \inf B$ where A, B are infinite subsets of D, then, in $\mathbb{R}$, $\sup \varphi[A] = \inf \varphi[B]$. So X is order-isomorphic to a subset of cl $Y \subseteq \mathbb{R}$. □

Theorem 7.39 tells us that SH is reasonable.

The next task is to relate trees and lines, in particular to show that there's a Suslin tree iff there's a Suslin line.

Definition 7.44. For T a tree, $t \in T, \alpha <$ ht t, we define $t(\alpha)$ to be the unique element of $T(\alpha)$ below t. If b is a branch of T, $b(\alpha) = b \cap T(\alpha)$.

Proposition 7.45. *Let T be a tree, B the set of branches of T. For each $\alpha <$ ht T let $\leq_\alpha$ be a linear order on $T(\alpha)$. For $b, c \in B$ we define $b \leq_* c$ iff $b = c$ or, for α the least ordinal with $b(\alpha) \neq c(\alpha)$, $b(\alpha) <_\alpha c(\alpha)$. Then $\leq_*$ is a linear order.*

The proof is left as an exercise.

A linear order as in proposition 7.45 will be called a canonical linear order on the branches..

We want to show that a canonical linear order on the branches of a Suslin tree is Suslin. To make the proof work, the tree needs an additional property:

Theorem 7.46. *A canonical linear order on the branches of a Suslin tree is a Suslin line.*

Proof. Let T be a Suslin tree under $\leq_T$, and let $\leq_*$ be a canonical linear order on the set of branches B.

Subclaim 7.46.1. No countable set is dense in the linear order $\leq_*$.

Proof. Since every branch is countable, if $A \subseteq B$ is countable, then $\bigcup A$ is contained in some T_α. There are at most countably many $t \in (T \setminus T_\alpha)$ with only countably many branches through t (because every branch is countable). If $t \in (T \setminus T_\alpha)$ and there are uncountably many branches through t there are $s, r \in T$ with $t \leq_T s, r$ and s incompatible with r. Let b be a branch through s, c be a branch through r and, without loss of generality, assume $b < c$ Then $(b, c) \neq \emptyset$ but $(b, c) \cap A = \emptyset$. Hence A is not dense. □

Subclaim 7.46.2. In the linear order $\leq_*$, there is no uncountable pairwise disjoint set of intervals.

Proof. Let $\mathcal{I}$ be pairwise disjoint collection of intervals.

For $I = (b_I, c_I) \in \mathcal{I}$ let α_I be least so $b_I(\alpha_I) \neq c_I(\alpha_I)$. For all $d \in I, b_I(\alpha_I) \leq d(\alpha_I) \leq c_I(\alpha_I)$. Pick $d_I \in (b_I, c_I)$.

Case 1. Suppose $d_I(\alpha_I) = b_I(\alpha_I)$. Define β_I to be least so $b_I(\beta_I) \neq d_I(\beta_I)$ and set $t_I = d_I(\beta_I)$.

Case 2 Suppose $d_I(\alpha_I) = c_I(\alpha_I)$. Define β_I to be least so $d_I(\beta_I) \neq c_I(\beta_I)$ and set $t_I = d_I(\beta_I)$.

Case 3 Suppose $b_I(\alpha_I) < d_I(\alpha_I) < c_I(\alpha_I)$. Define $\beta_I = \alpha_I$ and set $t_I = d_I(\beta_I)$.

In all three cases, if e is a branch through t_I then $e \in I$.

We show that $\{t_I : I \in \mathcal{I}\}$ is an antichain in T, hence countable.

If $t_I <_T t_J$ then $t_J(\beta_I) = t_I$, hence, for any branch e through t_J, $e \in I$. But there is some branch e through t_J with $e \in J$, a contradiction.

□

By the subclaims, B is a Suslin line under $\leq_*$.

□

Definition 7.47. A tree is a splitting tree iff above every element there are two incomparable elements.

Proposition 7.48. *If there is a Suslin tree, then there is a Suslin splitting tree.*

Proof. Let T be a Suslin tree. Let $S = \{t \in T : \text{if } t \uparrow \text{ is countable}\}$, where $t \uparrow = \{s \in T : s > t\}$. Let A be the set of minimal elements of S. A is an antichain, so it is countable. So $T \setminus \bigcup_{t \in A} t \uparrow$ is an uncountable subtree of T, hence Suslin. And $T \setminus \bigcup_{s \in A} s \uparrow$ is a splitting tree. □

Definition 7.49. Let X be linearly ordered. A tree of intervals on X is a tree whose elements are intervals in X with the order $I \leq J$ iff $I \supseteq J$ so that incomparable intervals in the tree are disjoint.

I.e., as we move up a branch, the intervals grow smaller.

Proposition 7.50. *A dense linear order has a countable splitting tree of intervals.*

Proof. Let X be a dense linear order. Let $x_0 < x_1 < x_2 \in X$, $T(0) = \{(x_0, x_1), (x_1, x_2)\}$.

Given $T(n)$ a collection of 2^{n+1} pairwise disjoint intervals, where $T(n)$ refines $T(n-1)$, and given $I \in T(n), I = (x_I, y_I)$, let $z_I \in (x_I, y_I)$ and define $T(n+1) = \{(x_I, z_I) : I \in T(n)\} \cup \{(z_I, y_I) : I \in T(n)\}$. $T = \bigcup_{n<\omega} T(n)$ is a splitting tree of intervals. □

We want to construct a Suslin tree of intervals from a Suslin line. But not just any Suslin line will do.

Proposition 7.51. *If there is a Suslin line there is a dense Suslin line.*

Proof. Let X be a Suslin line under $\leq$.

Subclaim 7.51.1. A strictly increasing well-ordered sequence in X must be countable. Similarly, a strictly decreasing well-ordered sequence in X must be countable.

Proof. We prove the decreasing case: If $x_\alpha > x_\beta$ for $\alpha < \beta < \delta$ then $\{(x_{\alpha+2}, x_\alpha) : \alpha < \delta, \alpha \text{ even}\}$ is a pairwise disjoint collection of non-empty intervals, so δ is countable. □

As an immediate corollary:

Subclaim 7.51.2. If $Y \subseteq X$ and Y is infinite then there are $\{(x_n, y_n) : n < \omega\}$ with each $x_n, y_n \in Y$, $x_{n+1} < x_n < y_m < y_{m+1}$ for all $n, m < \omega$ and $Y \subseteq \bigcup_{n<\omega} (x_n, y_n)$.

For $x < y \in X$ we define $x \equiv y$ iff (x, y) is countable (possibly empty). $\equiv$ is easily seen to be an equivalence relation, and

$$(\dagger) \text{ if } x < z < y \text{ and } x \equiv y \text{ then } z \equiv y.$$

Subclaim 7.51.3. Each $[x]$ is countable.

Proof. Let $\{(x_n, y_n) : n < \omega\}$ be as in subclaim 7.51.2 where $[x] \subseteq \bigcup_{n<\omega}(x_n, y_n)$ and each $x_n, y_n \in [x]$. Since each (x_n, y_n) is countable, so is $[x]$. □

Define $X^\dagger = \{[x]_\equiv : x \in X\}$ with $[x] \leq^\dagger [y]$ iff $x \leq y$. By subclaim 7.51.3, $X^\dagger$ is uncountable. By $\dagger$, $\leq^\dagger$ is linear. Note that if $([x], [y])$ is countable, then (x, y) is countable, so $x \equiv y$. Hence if $[x] <^\dagger [y]$, then $([x], [y])$ is uncountable, so certainly nonempty. I.e., $X^\dagger$ is a dense linear order.

We show $X^\dagger$ has no countable dense subset: Let $A^\dagger \in [X^\dagger]^\omega$, and define $A = \{x : [x] \in A^\dagger\}$. A is countable, so there are $y < z \in X$ with $(y, z) \neq \emptyset$ and $(y, z) \cap A = \emptyset$. But then $([y], [z]) \cap A^\dagger = \emptyset$.

We show $X^\dagger$ has no uncountable pairwise disjoint collection of intervals: If $([x], [y]) \cap ([z], [w]) = \emptyset$ where $[x] <^\dagger [y] <^\dagger [z] <^\dagger [w]$ then $(x, y) \cap (z, w) = \emptyset$, so an uncountable pairwise disjoint collection of intervals in $X^\dagger$ gives rise to an uncountable pairwise disjoint collection of intervals in X. □

Theorem 7.52. *If there is a Suslin line there is a Suslin tree.*

Proof. We may assume that our Suslin line X is dense.

Subclaim 7.52.1. Every countable splitting tree of intervals T in X has an extension $S \neq T$ which is also a countable splitting tree of intervals.

Proof. Let T be a countable splitting tree of intervals. Let $A = \{x \in X : x \text{ is an endpoint of some interval } I \in T\}$. A is countable and not dense in X, so there are $z < w \in X \setminus A$ with $(z, w) \cap A = \emptyset$. Then for all $I \in T$ either $(z, w) \subset I$ or $(z, w) \cap I = \emptyset$. $S = T \cup \{(z, w)\}$ is the desired extension. □

By proposition 7.50, a recursive construction of length ω_1 constructs an uncountable tree of intervals T.

If T had an uncountable antichain, X would have an uncountable pairwise disjoint family. If T had an uncountable chain, the left endpoints of the intervals in this chain would be a set $\{x_\alpha : \alpha < \omega_1\} \subset X$ so that $x_\alpha < x_\beta$ for $\alpha < \beta$. But then $\{(x_\alpha, x_{\alpha+1}) : \alpha < \omega_1\}$ would be an uncountable pairwise disjoint family. So T is Suslin. □

By theorem 7.46 and theorem 7.52

Corollary 7.53. *There is a Suslin tree iff there is a Suslin line.*

Thereom 7.53 has first proved by Kurepa in 1935 and published in a Belgrade journal where it languished unnoticed by the rest of the world. It was rediscovered independently in the 1940's by Sierpinki in Poland and by Miller in the United States.

We now leave small cardinals and connect trees to large cardinals.

Theorem 7.54. *$\kappa \to (\kappa)^2_2$ iff κ is strongly inaccessible and has the tree property.*

In particular, an uncountable cardinal is weakly compact iff it is strongly inaccessible and has the tree property.

Proof. Suppose $\kappa \to (\kappa)^2_2$. By corollary 7.15 we only need to prove that κ has the tree property.

Let T be a tree of size κ under the order $\leq_T$ in which each $|T(\alpha)| < \kappa$. Then $|T| = \kappa$ so we may list its elements as $T = \{t_\alpha : \alpha < \kappa\}$.

We extend the tree order $\leq$ on T to a linear order, as follows:

For all $\alpha < \kappa$ let $\leq_\alpha$ be a linear order on $T(\alpha)$. For $s \neq t \in T$ define $\alpha_{s,t}$ to be the least α with $s(\alpha) \neq t(\alpha)$. Finally, for $s, t \in T$, define $s \leq_* t$ iff $s \leq_T t$ or $s(\alpha) \leq t(\alpha)$.

Define $P_0 = \{\{\alpha, \beta\} : \text{if } \alpha < \beta \text{ then } t_\alpha <_* t_\beta\}$, $P_1 = [T]^2 \setminus P_0$. Let H be homogeneous for this partition, $|H| = \kappa$, and let $A = \{t \in T : \exists \kappa \text{ many } \alpha \in H \text{ with } t <_T t_\alpha\}$. A is a subtree of T of size κ, so $\forall \delta\ \exists \alpha \in H$ ht $t_\alpha > \delta$. Hence $A \cap T(\gamma) \neq \emptyset$ for all $\gamma < \kappa$.

We show that A is a branch.

Suppose not. Then there are two incompatible (in the sense of $\leq_T$) elements $t, s \in A$ with $t <_* s$. If $[H]^2 \subseteq P_0$, there are $\alpha < \beta < \gamma \in H$ with $t < t_\alpha, t_\gamma$ and $s < t_\beta$. So $t_\alpha, t_\gamma < t_\beta$, which contradicts the homogeneity of H. If $[H]^2 \subseteq P_1$, there are $x, t \in A$ with $t >^* s$, and $\alpha < \beta < \gamma \in H$ with $t < t_\alpha$ and $s < t_\beta, t_\gamma$. So $t_\alpha > t_\gamma, t_\beta$, which contradicts the homogeneity of H.

For the other direction, suppose κ is strongly inaccessible and has the tree property. Let $[\kappa]^2 = P_0 \cup P_1$ where $P_0 \cap P_1 = \emptyset$. We construct a tree $T = \{t_\alpha : \alpha < \kappa\}$ of distinct functions where each dom t_α is some ordinal $\leq \alpha$ ordered by extension as follows: $t_0 = \emptyset$. Given $\{t_\beta : \beta < \alpha\}$, we define t_α inductively as follows: $t_\alpha(0) = i$ iff $\{\alpha, 0\} \in P_i$; at stage γ, if $\forall \beta < \alpha\ t_\alpha|_\gamma \neq t_\beta$, then dom $t_\alpha = \gamma$. Otherwise there is some $\beta < \alpha$ with $t_\alpha|_\gamma = t_\beta$. There is exactly one such β, so we define $t_\alpha(\gamma) = i$ iff $\{\alpha, \beta\} \in P_i$.

Elements of T with the same height have the same domain, so by strong inaccessiblity each $|T(\alpha)| < \kappa$.

By the tree property, T has a branch b of size κ. Define $c_i = \{\alpha : t_\alpha, t_\alpha^\frown i \in b\}$. For some i, $|c_i| = \kappa$. If $\alpha < \beta \in c_i$ then $t_\beta \supseteq t_\alpha^\frown i$. I.e., $t_\beta(\alpha) = i$, so $\{\alpha, \beta\} \in P_i$. I.e., c_i is homogeneous. □

Is every strongly inaccessible cardinal weakly compact? A theorem which we not prove here says no:

Theorem 7.55. *If κ is weakly compact, then there are κ many strongly inaccessible cardinals below κ.*

Weakly compact cardinals have a rich combinatorial structure — they can be characterized in terms of infinitary languages, elementary structures, descriptive set theory, and so on.

7.3 Measurable cardinals

Measurable cardinals are the last large cardinals we shall talk about. They are also the smallest of the really big cardinals, in a sense we will make precise later. They are defined in terms of filters, so to define them, recall the definitions of section 1.6: a nonprincipal proper ultrafilter on a set X is a family $\mathcal{F}$ of infinite subsets of X which is closed under supersets and finite intersections, so that if $x \subseteq X$ then either $x \in \mathcal{F}$ or $X \setminus x \notin \mathcal{F}$.

Definition 7.56. An filter $\mathcal{F}$ is said to be κ-closed iff $\forall \mathcal{A} \in [\mathcal{F}]^{<\kappa} \bigcap \mathcal{A} \in \mathcal{F}$.

I.e., every filter is ω-closed, and a principal filter is κ-closed for all κ.

Fact 7.57. *A κ-closed nonprincipal ultrafilter contains no set of size smaller than κ.*

We are ready to define measurable cardinals.

Definition 7.58. An uncountable cardinal κ is measurable iff there is a κ-closed nonprincipal ultrafilter on κ.

If you know any measure theory, you will understand the following explanation of why these cardinals are called measurable: Given a κ-closed nonprincipal ultrafilter $\mathcal{F}$ on κ, define a measure $\mu : \mathcal{P}(\kappa) \to 2$ by: $\mu(A) = 1$ iff $A \in \mathcal{F}$. This is a measure, in fact a κ-additive measure, i.e., if $\mathcal{A} \subseteq [\kappa]^{\leq \kappa}, |\mathcal{A}| < \kappa$ and $\mathcal{A}$ is pairwise disjoint, then $\mu(\bigcup \mathcal{A}) = \Sigma_{A \in \mathcal{A}} \mu(A)$ — since no two sets in $\mathcal{F}$ are disjoint, at most one element of $\mathcal{A}$ is an element of $\mathcal{F}$, so κ-additivity is trivial.

How large are measurable cardinals?

Theorem 7.59. *Every measurable cardinal is weakly compact.*

Proof. Let κ be measurable, $\mathcal{F}$ a κ-closed ultrafilter on κ, and suppose $[\kappa]^2 = P_0 \cup P_1$ where P_0, P_1 are disjoint. We imitate the proof of theorem 7.10:

Let $\alpha_0 = 0$. There is a unique i_0 with $A_0 = \{\beta > 0 : \{0, \beta\} \in P_{i_0}\} \in \mathcal{F}$.

Now suppose we have, for some $\delta < \kappa$, $\{\alpha_\gamma : \gamma < \delta\}, \{A_\gamma : \gamma < \delta\}, \{i_\gamma : \gamma < \delta\}$ where, if $\gamma < \gamma'$ then $\alpha_{\gamma'} \in A_\gamma$, $A_\gamma \supset A_{\gamma'}$, and $A_\gamma \subset \{\alpha > \alpha_\gamma : \{\alpha_\gamma, \alpha\} \in P_{i_\gamma}\} \in \mathcal{F}$.

$\mathcal{F}$ is κ-closed, so $\bigcap_{\gamma<\delta} A_\gamma \in \mathcal{F}$. Let $\alpha_\delta \in \bigcap_{\gamma<\delta} A_\gamma$. There is i_δ so $A_\delta = \{\alpha \in \bigcap_{\gamma<\delta} A_\gamma : \alpha > \alpha_\delta, \{\alpha_\delta, \alpha\} \in P_{i_\delta}\} \in \mathcal{F}$.

There is $i \in 2$ so $y = \{\delta : i_\delta = i\}$ has cardinality κ. Finally, let $H = \{\alpha_\delta : \delta \in y\}$. H is homogeneous for P_i. □

Just as strongly inaccessible cardinals need not be weakly compact, so too weakly compact cardinals need not be measurable. A theorem we will not prove is

Theorem 7.60. *If κ is measurable, then there are κ many weakly compact cardinals below κ.*

The set-theoretic power of measurable cardinals comes from the following

Theorem 7.61. *κ is measurable iff there is a 1-1 map j from V to a transitive subclass $M \supset ON$ where*

(1) If Φ is a formula then $\forall a_1 ... a_n$ $V \models \Phi(a_1, ...a_n)$ iff $M \models \Phi(j(a_1), ...j(a_n))$.

(2) $j(x) = x$ for all $x \in V_\kappa$.

(3) $j(\kappa) \neq \kappa$.

κ is called the critical point of j.

In the taxonomy of large cardinals, it is exactly the property of an elementary embedding from V to a proper subclass that makes a large cardinal really big. By theorem 7.61, measurability is the weakest such property.

To give some idea of how j works, note that, since j is 1-1 and $j(\alpha) = \alpha$ for each ordinal $\alpha \in V_\kappa$, by (3), $j(\kappa) > \kappa$. But, since $M \supset$ ON, $\kappa \in M$, so j is not onto.

To explain (1) a little bit, if α is an ordinal, let Φ be the formula "x is transitive and well-ordered by $\in$." $V \models \Phi(x)$ iff, $\forall N$ transitive with $x \in N, N \models \Phi(x)$, so, since $M \models \Phi(j(\alpha))$, $j(\alpha) \in$ ON.

On the other hand, this if λ is a cardinal, even though $M \models j(\lambda)$ is a cardinal, there is no guarantee that $j(\lambda)$ is really a cardinal: there might be some $g \in V, \alpha < j(\lambda), g : \alpha \to j(\lambda), g$ is onto; but there might be no such g in M. In fact, $j(\kappa)$ will be such a non-cardinal image of a cardinal.

We will prove sufficiency in theorem 7.61 below — necessity needs logical apparatus that we don't have, but first we need

Proposition 7.62. *Suppose $j : V \to M$ is a 1-1 map, where M is a transitive subclass of V, $M \supset ON$, and property (1) of theorem 7.61 holds. Then*

(a) $\forall x, y, j(x \setminus y) = j(x) \setminus j(y)$.

(b) $\forall x, y$ if $x \subseteq y$ then $j(x) \subseteq j(y)$.

(c) If $x = \{y_\alpha : \alpha < \lambda\}$ then $j(x) = \{z_\alpha : \alpha < j(\lambda)\}$ where if $\alpha < \kappa$ then $z_\alpha = j(y_\alpha)$, and if $\alpha < \kappa$ and $y_\alpha \in V_\kappa$ then $z_\alpha = y_\alpha$.

(d) $\forall x \neq \emptyset$ if $|x| < \kappa$ then $j(\bigcap x) = \bigcap j(x)$.

Proof. We proof (a) and (c), leaving the rest to the reader.

For (a): Let $z = x \setminus y$. Let Φ be: $\forall w\ w \in z \leftrightarrow w \in x \wedge w \notin y$. Then $M \models \forall w\ w \in j(z)$ iff $w \in j(x) \setminus j(y)$.

For (c) Let $x = \{y_\alpha : \alpha < \lambda\}$ and let $f : \lambda \to x$ with $f(\alpha) = y_\alpha$ for all $\alpha < \lambda$. Then $M \models j(f) : j(\lambda) \to j(x)$ and $j(f)$ is onto, i.e., $j(x) = \{j(f)(\alpha) : \alpha < j(\lambda)\}$. For each $\alpha < j(\lambda)$ we write $z_\alpha = j(f)(\alpha)$. Suppose $y = f(\alpha)$. Then $j(y) = j(f)(j(\alpha)) = z_{j(\alpha)}$. If $\alpha < \kappa$ then $j(y) = j(f)(\alpha) = z_\alpha$. If $y \in V_\kappa$ and $\alpha < \kappa$ then $y = j(y) = z_\alpha$. □

Proposition 7.62(c) is a bit counterintuitive. It says that if you start with a short sequence in V, its image in M has the same length, but if you start with a long sequence in V, its image in

M is even longer. In particular, if you start with a κ-sequence $\vec{s}$, where κ is the critical point of j, then $j(\vec{s})$ is a $j(\kappa)$ sequence. I.e., $j(\{y_\alpha : \alpha < \kappa\}) \neq \{j(y_\alpha) : \alpha < \kappa\}$.

We prove half of theorem 7.61.

Theorem 7.63. *Suppose there is a 1-1 map $j : V \to M$ as in theorem 7.61 where κ is the critical point (i.e., κ satisfies properties (2) and (3)). Then κ is measurable.*

Proof. Define $\mathcal{F} = \{x \subseteq \kappa : \kappa \in j(x)\}$.

By proposition 7.62,

(i) $\mathcal{F}$ is closed under finite intersection and superset.

(ii) $\forall x \subseteq \kappa$, either $x \in \mathcal{F}$ or $\kappa \setminus x \in \mathcal{F}$.

(iii) $\emptyset \notin \mathcal{F}$.

(iv) If $x \in \mathcal{F}$ then $|x| = \kappa$.

So $\mathcal{F}$ is a nonprincipal ultrafilter on κ. Why is it κ-closed?

Suppose $\lambda < \kappa$ and $X = \{x_\alpha : \alpha < \lambda\} \subset \mathcal{F}$. Then $j(X) = \{j(x_\alpha) : \alpha < \lambda\}$ and $\forall \alpha < \lambda \ \kappa \in j(x_\alpha)$. So $\kappa \in \bigcap_{\alpha<\lambda} j(x_\alpha)$, i.e., by proposition 7.62(d), $\bigcap X \in \mathcal{F}$. □

Another measure of how large measurable cardinals are is that they don't exist in L.

Theorem 7.64. *(Scott) If $V = L$ then there are no measurable cardinals.*

Proof. . If there is a measurable cardinal then there is a smallest one, κ. Let j, M be as in theorem 7.61. Then $M \models j(\kappa)$ is the smallest measurable cardinal. If $V = L$ then $V \subseteq M \subseteq V$, so $M = V$. Hence $V \models$ "$\kappa < j(\kappa), \kappa$ is the smallest measurable cardinal ,and $j(\kappa)$ is the smallest measurable cardinal." This is a contradiction. □

With inaccessible, weakly compact, and measurable cardinals we have barely touched on the rich subject of large cardinals. It has close connections to descriptive set theory (the classification of sets of reals begun by Borel and continued by the Polish school of Kuratowski and Sierpinski), and to the axiom of determinacy and its variations (i.e., the investigation of which infinite games have guaranteed winning strategies), as well as applications in many fields of mathematics (theorems of the form "if there is this kind of large cardinal then a statement of mathematical interest is true or, more usually, consistent").

7.4 Cardinal invariants of the reals

To a set theorist, "the reals" does not necessarily mean $\mathbb{R}$. Sometimes it means 2^ω, sometimes $\mathcal{P}(\omega)$, sometimes ω^ω. These are closely related.

There is a strong connection between $\mathcal{P}(\omega)$ and 2^ω via characteristic functions ($\chi_x(n) = 0$ iff $n \in x$). Similarly, every function $f : \omega \to \omega$ is a set of ordered pairs, i.e., a subset of $\mathcal{P}(\mathbb{N} \times \mathbb{N})$ hence, since $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$, f can be coded as a subset of $\mathbb{N}$, thus identifying ω^ω with $\mathcal{P}(\omega)$. If you know a little topology, you can easily prove that ω^ω under the product topology is homeomorphic to the irrationals. It's even easier to see that 2^ω is homeomorphic to the Cantor set, another subset of $\mathbb{R}$. For these and other reasons, set theorists move among these structures easily, considering all of them "the reals."

A little notation: we often use $\mathfrak{c}$ to denote the cardinal number 2^ω, as opposed to 2^ω, the set of all functions from ω into 2.

We introduce a number of cardinal invariants. To define them we will refer back to concepts first introduced in chapter 1. We will also need

Definition 7.65. (a) A family $A \subset [\omega]^\omega$ is called almost disjoint iff $\forall a \neq b \in A\ a \cap b$ is finite.[106] A is maximal almost disjoint iff it is almost disjoint and $\forall a \in [\omega]^\omega\ \exists b \in A\ a \cap b$ is infinite.

(b) For a, b infinite, $a \subseteq^* b$ iff $a \setminus b$ is finite. If b is infinite and $b \subseteq^* a$ for all $a \in A \subset [\omega]^\omega$ then b is called a pseudo-intersection of A.

(c) Let $f, g \in \omega^\omega$. $f =^* g$ iff $\{n : f(n) \neq g(n)\}$ is finite; $f \leq^* g$ iff $\{n : f(n) > g(n)\}$ is finite.

It's easy to see that $=^*$ is an equivalence relation and $\leq^*$ is a pre-order.

Definition 7.66. (a) $\mathfrak{p} = \inf\{|\mathcal{A}| : \mathcal{A}$ is a non-principal filterbase on ω with no pseudo-intersection$\}$. (Recall the definition of filterbase from definition 1.45; a filterbase is non-principal iff every intersection of finitely many of its sets is infinite.)

(b) $\mathfrak{a} = \inf\{|\mathcal{A}| \subset [\omega]^\omega : \mathcal{A}$ is maximal almost disjoint$\}$.

(c) $\mathfrak{b} = \inf\{|F| : F$ is unbounded in ω^ω under $\leq^*\}$. (The definition of unbounded is in definition 1.16.)

(d) $\mathfrak{d} = \inf\{|F| : F$ is dominating in ω^ω under $\leq^*\}$. (The definition of dominating is definition 1.17.)

There are many more cardinal invariants, but these are the ones we will focus on.

Clearly

Fact 7.67. $\mathfrak{p}, \mathfrak{a}, \mathfrak{b}, \mathfrak{d} \leq \mathfrak{c}$.

In this and the next chapter we will look at cardinal invariants for their own sake, but they are also useful in proving consistency results. Equalities such as $\mathfrak{b} = \mathfrak{d}, \mathfrak{d} = \omega_1$, etc. have consequences in many fields of mathematics. One way to prove consistency results is to show that a statement follows from a consistent statement about cardinal invariants of the reals.

Our first theorem says that none of these cardinal invariants is countable.

[106] I.e., the levels of the second Aronszajn tree we constructed were almost disjoint.

Theorem 7.68. $\omega < \mathfrak{p}, \mathfrak{a}, \mathfrak{b}, \mathfrak{d}$

Proof. For $\omega < \mathfrak{p}$: This follows from theorem 1.43.

For $\omega < \mathfrak{a}$: Let $A \subset [\omega]^\omega$ be a countable almost disjoint family, $A = \{a_n : n < \omega\}$. Let $b_0 = a_0, b_{n+1} = a_{n+1} \setminus \bigcup_{i \leq n} a_n$. $\{b_n : n < \omega\}$ is a pairwise disjoint family of infinite sets. Let $k_n \in b_n$ for all $n < \omega$, $a = \{k_n : n < \omega\}$. Then $a \cap a_n$ is finite for all n.

For $\omega < \mathfrak{b}, \mathfrak{d}$: This follows from exercise 24 in chapter 1.

□

As an immediate corollary

Corollary 7.69. *Assume CH.* $\mathfrak{p} = \mathfrak{a} = \mathfrak{b} = \mathfrak{d} = \omega_1 = \mathfrak{c}$.

Our next proposition says what can be said about relations among these invariants.

Theorem 7.70. *(a)* $\mathfrak{b} \leq \mathfrak{d}$

(b) $\mathfrak{p} \leq \mathfrak{a}$

(c) $\mathfrak{p} \leq \mathfrak{b}$.

Note that theorem 7.70 greatly simplifies the proof of theorem 7.68.

Proof. For (a): A dominating family is unbounded.

For (b): If A is infinite almost disjoint, then $B = \{\omega \setminus a : a \in A\}$ is a non-principal filterbase. B has no pseudo-intersection iff A is maximal almost disjoint. So there is always a filterbase B with no pseudo-intersection and $|B| = \mathfrak{a}$.

For (c): Let F be an unbounded family in ω^ω, and for $f \in F$ define $g_f(n) = \sup\{f(i) : i \leq m\}$. Each g_f is non-decreasing, and $\{g_f : f \in F\}$ is also unbounded. Also, if F is unbounded we may assume

$$(\dagger) \quad \forall m \; \exists f \in F \; \forall n f(n) > m.$$

So when we consider a potential unbounded family, we may assume its functions are non-decreasing and that it has property (†).

Now let $F \subseteq \omega^\omega$ be a collection of non-decreasing functions so that $\forall m \; \exists f \in F \; \forall n f(n) > m$. For $f \in F$ define $a_f = \{(n, m) : m > f(n)\}$. Suppose b is a pseudo-intersection of $A_F = \{a_f : f \in F\}$. If some $(n, m) \in b$ we define $g(n) = \inf\{m : (n, m) \in b\}$. g is a partial function from ω to ω with infinite domain, and $\forall f \in F \; \{n : g(n) \leq f(n)\}$ is finite.

We define $h \in \omega^\omega$ as follows: Let $\{k_i : i < \omega\} = \text{dom } g$ so that each $k_i < k_{i+1}$. For $k \in [k_i, k_{i+1})$, define $h(k) = g(k_{i+1})$. We show that $\forall f \in F \; h \geq^* f$.

Fix $f \in F$. There is i so $\forall j \geq i \; f(k_j) < g(k_j)$. Since f is non-decreasing, for $j \geq i$ and $k \in [k_j, k_{j+1}), f(k) \leq f(k_{j+1}) < g(k_{j+1}) = h(k)$.

I.e., if F is unbounded then A_F has no pseudo-intersection. So there is always a filterbase A with no pseudo-intersection, $|A| = \mathfrak{b}$. □

Theorem 7.70 is nearly all that can be said about relative sizes in ZFC. For example, $\omega_1 < \mathfrak{b} < \mathfrak{d} < \mathfrak{c}$ is consistent, as is $\omega_1 = \mathfrak{a} < \mathfrak{d}$, as is $\mathfrak{d} < \mathfrak{a}$[107] and so on.

[107]This was quite difficult, and is due to Shelah.

7.5 CH and Martin's axiom

In the previous section we saw CH's effect on cardinal invariants of the reals. CH has many other consequences, throughout mathematics. We give one example, from partition calculus.

Theorem 7.71. *(Erdös-Rado). Assume CH. Then $\omega \times \omega_1 = H \cup K$ where $H \cap K = \emptyset$ and there is no homogeneous subset of the form $A \times B$ where $A \in [\omega]^\omega$ and $B \in [\omega_1]^{\omega_1}$.*[108]

Proof. By CH, let $\{a_\alpha : \alpha < \omega_1\}$ enumerate $[\omega]^\omega$. Using a recursive construction of length ω_1, at stage β we decide which $(k, \beta) \in H$ and which $(k, \beta) \in K$.

At stage β, let $\{c_{\beta,n} : n < \omega\}$ enumerate $\{a_\alpha : \alpha < \beta\}$. Since each $c_{\beta,n}$ is infinite, we pick disjoint infinite sets $s_{\beta,0}, s_{\beta,1} \subset \omega$ with $c_{\beta,n} \cap s_{\beta,n,i} \neq \emptyset$ for $n < \omega, i \in \{0, 1\}$. If $k \in s_{\beta,0}$ put $(k, \beta) \in H$; if $k \in s_{\beta,1}$ put $(k, \beta) \in K$.

Suppose B is uncountable and A is infinite. Then A is some a_α and there is some $\beta \in B$ with $\beta > \alpha$. Let $k \in s_{\beta,0} \cap A, m \in s_{\beta,1} \cap A$. Then $(k, \beta) \in H$ and $(m, \beta) \in K$ so $A \times B$ is not homogeneous. □

The proof of theorem 7.71 is typical of many CH proofs: line things up in order type ω_1, create ω_1 many requirements you need to meet, and then systematically meet all of them by a recursive construction.

Martin's axiom is stated quite different from CH, but as we will learn, in fact it is a consequence of CH. Generally, we are interested in the axiom MA + ¬CH, which was proven consistent by Solovay and Martin.

There are several equivalent ways to state MA. We will use the version using partial orders. In doing so, we need to introduce yet a third meaning for the word "dense" (for the others, see the theory DLO in chapter 2, and example 5.61.)

Definition 7.72. Let P be a partial order under $\leq$. D is dense in P iff $\forall p \in P\ \exists d \in D\ d \leq p$.

MA is not just about any partial orders, but ccc partial orders.

Definition 7.73. A partial order is ccc iff it has no uncountable antichains.[109]

I.e., a ccc linear order is one in which the set of intervals under reverse inclusion is ccc.

Finally, MA is about filters on partial orders.

Definition 7.74. Let P be a partial order under $\leq$. G is a filter in P iff $\forall p, q \in G$ there is $r \in G$ with $r \leq p, q$, and if $\in G$ and $q > p$ then $q \in G$.

I.e., a filter on a set X is a filter on the partial order $\mathcal{P}(X)$ where $x \leq y$ iff $x \subseteq y$.

We are ready to define MA.

Definition 7.75. MA is the following statement: If P is a ccc partial order and $\mathcal{D}$ is a family of dense subsets of $P, |\mathcal{D}| < \mathfrak{c}$, then there is a filter G in P with $G \cap D \neq \emptyset$ for all $D \in \mathcal{D}$ (G is called $\mathcal{D}$-generic).

[108] Another way to write this would be: $\omega \times \omega_1 \not\to (\omega \times \omega_1)^1_2$.

[109] "ccc" abbreviates "countable chain condition" even though it is about antichains. Sorry.

The following weak version of MA is simply true in ZFC:

Theorem 7.76. *(Rasiowa-Sikorski) If P is a partial order and $\mathcal{D}$ is a countable family of dense subsets of P then there is a $\mathcal{D}$-generic filter G in P.*

The proof is left as an exercise. Theorem 7.76 is known as the Rasiowa-Sikorski lemma.

Corollary 7.77. *If CH, then MA.*

Proof. By CH, "$< \mathfrak{c}$" means "countable", so we are done (in fact have a slightly stronger statement) by theorem 7.76. □

MA + ¬CH has a number of weaker variations which are independent of ZFC, and it is often useful in using MA to state the weakest variation needed for the proof. For example, $\mathrm{MA}_{\mathrm{countable}}$ substitutes "countable partial order" for "ccc partial order"; MA_{ω_1} substitutes "$|\mathcal{D}| = \omega_1$" for "$|\mathcal{D}| < \mathfrak{c}$", etc.

The reader may recall from section 7.2 that CH is independent of SH. However

Theorem 7.78. *Assume MA_{ω_1}. Then SH holds.*

Proof. First we need to show that if there's a Suslin tree, there's one in which every element has sucessors of arbitrarily high height.

Subclaim 7.78.1. If T is a Suslin tree, there is a Suslin tree $T^* \subseteq T$ so that

$$(*) \quad \forall t \in T^* \; \forall \alpha < \omega_1 \; \exists s \in T^*(\alpha) \; s, t \text{ are comparable.}$$

Proof. If T does not satisfy $(*)$, let $A = \{t \in T : \; \exists \alpha \; \forall \beta > \alpha \; t \text{ has no successor in } T(\beta)\}$. Let B be the set of minimal elements in A. B is an antichain, hence countable. If $t \in B$ then $t \uparrow$ is countable, so A is countable. $T^* = T \setminus A$ is the tree required. □

Assume MA_{ω_1} and suppose T is a Suslin tree under $\leq_T, T$ satisfies $(*)$. Define $t \leq_* s$ iff $t \geq_T s$. By definition, T is a ccc partial order under $\leq_*$. Let $D_\alpha = \{t \in T : t \in T(\alpha)\}$. By property $(*)$, each D_α is dense. By MA_{ω_1}, there is a filter G which is $\{D_\alpha : \alpha < \omega_1\}$-generic. Since G is a filter, it's a branch. Hence an uncountable branch. So T was not Suslin. □

In some sense, MA is derived from the proof of theorem 7.78. That is, originally Solovay and Tennenbaum constructed a proof of the consistency of SH via forcing. Forcing works very much like MA, that is, it is about creating generic filters.[110] Contemplating what makes the the forcing proof work, Solovay and Martin came up with MA, and proved

Theorem 7.79. *MA + ¬CH is consistent.*

The technique used to prove theorem 7.79 is known as iterated forcing, and again has its origins in the Solovay and Tennenbaum construction.

The reader recalling the independence of SH from CH will note that, by theorem 7.78, MA_{ω_1} does not imply CH. In fact

[110] The filters in forcing meet so many dense sets that they can't exist in the original model, so when you force you enlarge the original model.

Theorem 7.80. *Assume MA_{ω_1}. Then CH fails.*

Proof. Let $F \in [2^\omega]^{\omega_1}$. We show that $F \neq 2^\omega$.

Define $P = \bigcup_{n<\omega} 2^n$ where $p \leq q$ iff $p \supseteq q$.[111] P is countable, so ccc.

For $f \in F$, define $D_f = \{p \in P : \exists k \in \text{dom } p\ p(k) \neq f(k)\}$. We show each D_f is dense.

Fix $f \in F$. Given $q \in P$, let $n = \text{dom } q$. Let $p \supset q$ with $p(n) = 1 - f(n)$. Then $p \leq q$ and $p \in D_f$. So D_f is dense.

By MA_{ω_1}, let G be a $\{D_f : f \in F\}$-generic filter in P. Then $\bigcup G$ is a (possibly partial) function from ω to 2. There is $g \in 2^\omega$ with $g \supseteq \bigcup G$. By genericity, $\forall f \in F\ g \neq f$.

Hence CH fails. □

The partial order P in the proof of theorem 7.80 is known as Cohen forcing (actually, as forcing with one Cohen real). If G had been generic not just for all D_f with $f \in F$, but for all D_f with $f \in 2^\omega$ the function g would not have been in our original model; it would have been a new function. Roughly speaking, that is how forcing works.[112] Cohen's original forcing to show the consistency of ¬CH was a variation in which he simultaneously forced with ω_2 versions of P.

CH implies that the cardinal invariants of section 3.4 are all equal to $\mathfrak{c}$. So does MA, although its version of $\mathfrak{c}$ can be larger than CH's version.

Theorem 7.81. *Assume MA. Then* $\mathfrak{p} = \mathfrak{a} = \mathfrak{b} = \mathfrak{d}$.

Proof. Assume MA. By theorem 7.70 and fact 7.67 it suffices to show that $\mathfrak{p} \geq \mathfrak{c}$.

So let $A \subset [\omega]^\omega$, A is a non-principal filterbase, $|A| < \mathfrak{c}$. Without loss of generality we may assume that A is closed under finite intersections.

Let P be the following partial order: an element of P has the form $p = (\sigma_p, a_p)$ where $a_p \in A$ and $\sigma_p \in \bigcup_{n<\omega} 2^n$. We also write $c_p = \{k : \sigma_p(k) = 0\}$ (i.e., σ_p is the characteristic function of c_p). The order on P is: $p \leq q$ iff $\sigma_p \supseteq \sigma_q$, $a_p \subseteq a_q$, and $c_p \setminus c_q \subseteq a_q$.

I.e., we are trying to construct a set c out of a union of some of the c_p's; and each p insists that any new element of c must already be in a_p.

Note that if $\sigma_p = \sigma_q$ then $(\sigma_p, a_p \cap a_q) \in P$. Hence if p, q are incompatible, $\sigma_p \neq \sigma_q$. $\bigcup_{n<\omega} 2^n$ is countable. So if $Q \in [P]^{>\omega}$ there is an uncountable $R \subseteq Q$ and σ with $\sigma_p = \sigma$ for all $p \in R$. R is pairwise compatible. Hence P is ccc.

Now we have to find the right dense sets. We want to construct a pseudo-intersection of A, so, for $a \in A$, let $D_a = \{p : a_p \subseteq a\}$. We show each D_a is dense:

Let $q \in P$. Let $b = a \cap a_p$. $b \in A$, so define $p = (\sigma_q, b)$. $p \in D_a$ and $p \leq q$.

Finally, let G be a $\{D_a : a \in A\}$-generic filter in P. Let $c = \bigcup_{p \in G} c_p$. We show that c is a pseudo-intersection of A:

[111] This reversal of the usual order — the more you know, the smaller you are — is common in most discussions of forcing (although Shelah reverses it); and MA is very close to an axiom about forcing, whatever that is.

[112] And the logical hoops you need to jump through to show that it makes sense to do this are quite complicated.

Let $a \in A$. Let $p \in G \cap D_a$. If $k \in c \setminus c_p$ then $k \in a_p \subseteq a$. Since c_p is finite, we're done. □

The partial order used in the proof of theorem 7.81 is, in fact, a countable union of filterbases. Such partial orders are called σ-centered, so the proof of theorem 7.81 actually uses only a piece of MA, $\mathrm{MA}_{\sigma-\mathrm{centered}}$, i.e., MA when "ccc" is replaced by "σ-centered."

The reader may wonder whether all of these versions of MA + ¬CH are in fact equivalent. We state without proof

Theorem 7.82. *(a) MA_{ω_1} is strictly weaker than MA + ¬CH.*

(b) $MA_{\mathrm{countable}}$ + ¬CH is strictly weaker than $MA_{\sigma-\mathrm{centered}}$ + ¬CH which is strictly weaker then MA + ¬CH.

Under GCH, $2^\omega < 2^{\omega_1}$. But not under MA + ¬CH:

Theorem 7.83. *Assume $MA_{\sigma-\mathrm{centered}}$. If $\kappa < \mathfrak{c}$ then $2^\kappa = 2^\omega$.*

Proof. First, note that if $T = \bigcup_{n<\omega} 2^n$ is the binary tree of height ω, where the order is end-extension, then the set of branches of T is an almost disjoint family on T. Since T is countable, we can identify it with ω, so there is always an almost disjoint family on ω of size $\mathfrak{c}$, hence of any smaller cardinality.

So suppose $\kappa < \mathfrak{c}$, and let A be an almost disjoint family on ω of size κ. For $C \subseteq A$ we construct a set $a_C \subseteq \omega$ so that if $C \neq E$ then $a_C \neq a_E$. I.e., each subset of κ is associated with a subset of ω in 1-1 fashion, so $2^\kappa \leq 2^\omega \leq 2^\kappa$.

We define $p \in P_C$ iff $p = (\sigma_p, E_p)$ where $\sigma_p \in \bigcup_{n<\omega} 2^n$, $E_p \in [C]^{<\omega}$, and define $a_p = \{k : \sigma_p(k) = 0\}$. $p \leq q$ iff $\sigma_p \supseteq \sigma_q, E_p \supseteq E_q$ and $(a_p \setminus a_q) \cap \bigcup E_q = \emptyset$. I.e., once $c \in E_q$, nothing in c can be added to a_C.

Again, if $\sigma_p = \sigma_q$, p, q are compatible, so P_C is σ-centered.

For $c \in C$ we define $D_c = \{p : c \in E_p\}$. D_c is dense: given $q \in P_C$ let $p = (\sigma_q, E_q \cup \{c\})$. Then $p \in D_c, p \leq q$.

For $c \notin C, n \in \omega$ we define $D_{c,n} = \{p \in P_C : |a_c \cap c| \geq n\}$. We show $D_{c,n}$ is dense: given $q \in P_C$, let $e \in [c]^n, e \cap (\operatorname{dom} \sigma_p \cup \bigcup E_q) = \emptyset$ — we can do this because A is almost disjoint. Then let $p = (\sigma_p, E_q)$ where $\sigma_p : 1 + \sup e \to 2, \sigma_p \supseteq \sigma_q$, and if $k \in e$ then $\sigma_p(k) = 0$. $p \in D_{c,n}$ and $p \leq q$.

Finally, let G be generic for $\{D_c : c \in C\} \cup \{D_{c,n} : c \notin C, n < \omega\}$. Let $a_C = \bigcup_{p \in G} a_p$. $a_C \cap c$ is finite for all $c \in C$, since there is $p \in G \cap D_c$ and $a_C \cap c \subset a_c$. Similarly, $a_C \cap c$ is infinite for all $c \notin C$, since for all n there is $p \in G \cap D_{c,n}$, hence $\forall n \ |a_C \cap c| > n$.

□

The partial order in the proof of theorem 7.83 is known as almost disjoint forcing, and is due to Solovay.

You may recall Easton's theorem (theorem 5.56) which said, in part, that it is consistent for $2^\omega = \aleph_{\omega_1}$. But not under MA.

Corollary 7.84. *Assume MA. Then 2^ω is regular.*

Proof. If $\kappa < 2^\omega$ then $2^\kappa = 2^\omega$, so cf $2^\omega =$ cf $2^\kappa > \kappa$. □

The reader may wonder if there are any other restrictions on the cardinality of 2^ω under MA, and the answer is no.

So far, except for the result on Suslin trees, MA + ¬CH looks a lot like CH, except things are moved up from ω_1 to $\mathfrak{c}$. Here's a theorem in which it is clear that they are different.

Theorem 7.85. *(Baumgartner-Hajnal) Assume* $\mathfrak{p} = \mathfrak{c}$.[113] *Then* $\omega \times \omega_1 \to (\omega \times \omega_1)^1_2$, *i.e., if* $\omega \times \omega_1 = H \cup K$ *where* $H \cap K = \emptyset$ *then there are* $A \in [\omega]^\omega, B \in [\omega_1]^{\omega_1}$ *so either* $A \times B \subseteq H$ *or* $A \times B \subseteq K$.

I.e., under $\mathrm{MA}_{\sigma-\text{centered}}$ + ¬CH, the conclusion of theorem 7.71 fails.

Proof. Fix some $\mathcal{F}$ a non-principal ultrafilter on ω.

For each $\alpha < \omega_1$ let $H_\alpha = \{n : (n, \alpha) \in H\}$ and $K_\alpha = \{n : (n, \alpha) \in K\}$. $\omega = H_\alpha \cup K_\alpha$, so for each α either $H_\alpha \in \mathcal{F}$ or $K_\alpha \in \mathcal{F}$. Either uncountable many H_α's are in $\mathcal{F}$, or uncountably many K_α's are in $\mathcal{F}$. Without loss of generality assume the former, and let $E = \{\alpha : H_\alpha \in \mathcal{F}\}$.

$\{H_\alpha : \alpha \in E\}$ is a non-principal filter of size $< \mathfrak{c}$, so by theorem 7.81, it has a pseudo-intersection, A. For $\alpha \in E$ let n_α be the least n so $A \setminus H_\alpha \subseteq n$. There is an uncountable $B \subseteq E$ with $n_\alpha = n$ for all $\alpha \in B$. Hence $(A \setminus n) \times B \subseteq H$, as desired. □

There are other axioms similar to MA which guarantee the existence of sets which are "enough" generic. The two most widely used are the proper forcing axiom, PFA, and Martin's Maximum. Both of these axioms need large cardinals to be proven consistent, and both of them imply that $\mathfrak{c} = \omega_2$.

[113] In particular, $\mathrm{MA}_{\sigma-\text{centered}}$ + ¬CH will do.

7.6 Stationary sets and $\diamondsuit$

In this final section we discuss a combinatorial principle which strengthens CH. This principle, known as $\diamondsuit$ (diamond), holds in L and was formulated by Jensen as a crystalization of his proof that Suslin trees exist in L (just as MA was formulated as a crystalization of the proof that Suslin trees need not exist). L is a rich source of powerful combinatorial principles, with fanciful ad hoc names such as $\clubsuit$ (club, a weakening of $\diamondsuit$ — $\diamondsuit = \clubsuit + CH$), $\square$ (box), morasses, and so on. Most of these principles involve some kind of guessing — essentially we are given a fixed sequence (the guessing sequence) that simultaneously manages to approximate (or guess) a large collection of sets.

Before stating $\diamondsuit$ we need to characterize some interesting sets of ordinals.

Definition 7.86. Let α be an ordinal

(a) $C \subseteq \alpha$ is said to be closed iff, for all $A \subseteq C$ if A is not cofinal in α then $\sup A \in C$.

(b) A closed unbounded (a.k.a. club) subset of α is a closed subset cofinal in α.

For example, all subsets of ω are closed, so every unbounded subset of ω is a club. If κ is an uncountable cardinal, then $\{\alpha < \kappa : \alpha \text{ is a limit ordinal}\}$ is a club, since the limit ordinals are cofinal in κ and every supremum of a subset of limit ordinals is itself a limit ordinal.

Fact 7.87. *If C is a club in α and $cf(\alpha) = \lambda$, then C must contain elements of every cofinality below λ.*

Proof. Every set cofinal in α has order type $\geq \lambda$, so if ρ is regular and $\rho < \lambda$ then each club C has a subset A of order type ρ, hence $\sup A$ has cofinality ρ, and $\sup A \in C$. $\square$

Fact 7.88. *If $cf(\kappa) = \lambda$ then the intersection of fewer than λ many club sets in κ is a club.*

The proof is left to the reader.

We generalize theorem 5.39.

Theorem 7.89. *Let α be an orrdinal with uncountable cofinality. A continuous strictly increasing function from a club in α to α has a club of fixed points.*

Proof. Let α have uncounable cofinality, $f : C \to \alpha$ where C is a club in α and f is continuous and strictly increasing.

If A is a set of fixed points in C and $\sup A < \alpha$ then by continuity $\sup A$ is a fixed point, hence in C.

If $\gamma < \alpha$ there is $\gamma_0 \in C$ with $\gamma_0 \geq \gamma$. By induction, for all n there is $\gamma_{n+1} \in C$ with $f(\gamma_n) \leq \gamma_{n+1}$. Let $\beta = \sup\{\gamma_n : n < \omega\}$. By continuity $f(\beta) = \beta$, so the set of fixed points is unbounded. $\square$

Definition 7.90. Let κ be a cardinal. Then $S \subseteq \kappa$ is said to be stationary iff $S \cap C \neq \emptyset$ for all C club in κ.

For example, by fact 7.87, if $\rho < \mathrm{cf}(\kappa), \{\alpha : \mathrm{cf}(\alpha) = \rho\}$ is stationary in κ. By fact 7.88, every club is stationary. And if $\mathrm{cf}(\kappa) = \omega$ then S is stationary in κ iff $\kappa \setminus S$ is bounded below κ (the proof is an exercise).

The division of sets into stationary and nonstationary sets is an important one. To give an idea of the power of stationary sets, we state three theorems about them, proving two.

Theorem 7.91. *(Fodor) Let κ be a regular uncountable cardinal. If $f : S \to \kappa$, S is stationary, and $f(\alpha) < \alpha$ for all nonzero $\alpha \in S$ (such a function f is called regressive), then there is a stationary $R \subseteq S$ with f constant on R.*

Proof. Given f, S as in the hypothesis, $\alpha < \kappa$, define $S_\alpha = \{\beta : f(\beta) = \alpha\}$. If S_α is not stationary, then there is a club $C_\alpha \subseteq \kappa$ with $C_\alpha \cap S_\alpha = \emptyset$. Let $C = \{\beta : \forall \beta < \alpha \ \beta \in C_\alpha\}$. By exercise 7.40, C is club. So there is some $\beta \in S \cap C$. Then $f(\beta) \geq \beta$, a contradiction.

□

Our second important theorem about stationary sets is

Theorem 7.92. *(Solovay) Every regular uncountable cardinal κ is the union of κ many disjoint stationary subsets of κ.*

Proof. Let $S = \{\alpha < \kappa : \mathrm{cf}(\alpha) = \omega\}$, and for $\alpha \in S$ let $\{\beta_{n,\alpha} : n < \omega\}$ be an increasing sequence converging to α.[114]

For $n < \omega, \eta < \kappa$ let $S_{n,\eta} = \{\alpha \in S : \beta_{n,\alpha} \geq \eta\}$.

Subclaim 7.92.1. $\exists n \ \forall \eta \ S_{n,\eta}$ is stationary.

Proof. If not, $\forall n \ \exists \mu_n \ S_{n,\mu_n}$ is not stationary, so let C_n be club with $C_n \cap S_{n,\mu_n} = \emptyset$, and let C be club with $C \subseteq \bigcap_{n<\omega} C_n$. Let $\mu = \sup\{\mu_n : n < \omega\}$. C is a club and $\forall n \ \forall \alpha \in C \cap S \ \beta_{n,\alpha} < \mu_n < \mu$. I.e., $\mu = \sup(C \cap S)$, a contradiction. □

Fix n as in subclaim 7.92.1. For $\alpha \in S$ define $f(\alpha) = \beta_{n,\alpha}$. f is regressive. The proof of theorem 7.92 will be completed by constructing two κ sequence of stationary sets $\{S_\eta : \eta < \kappa\}$, $\{R_\eta : \eta < \kappa\}$ so that

(i) each $R_\eta \subset S_\eta$

(ii) $\{R_\eta : \eta < \kappa\}$ is pairwise disjoint

(iii) $\nu < \eta$ then $S_\nu \supset S_\eta$.

(iv) $\forall \eta \ \exists \ \gamma_\eta \ \forall \alpha \in R_\eta \ f(\alpha) = \gamma_\eta$

Let $S_0 = S$. By Fodor's theorem (theorem 7.91) $\exists \gamma_0$ with $R_0 = \{\alpha \in S_0 : f(\alpha) = \gamma_0\}$ stationary. Now suppose we have S_ν, R_ν, γ_ν for $\nu < \eta$. Define $S_\eta = S \setminus \bigcup_{\nu<\eta} R_\eta$. Note that, by subclaim 7.92.1 $S_{n,\eta} \subseteq S_\eta$. By Fodor's theorem $\exists \gamma_\eta$ with $R_\eta = \{\alpha \in S_\eta : f(\alpha) = \gamma_\eta\}$ stationary. □

And the third theorem we present (without proof) on stationary sets is a more general version of Silver's singular cardinals theorem (theorem 5.57).

[114] This is sometimes called a ladder system.

Theorem 7.93. *Suppose $\aleph_\alpha$ is a singular cardinal of uncountable cofinality and, for some fixed $\gamma < cf\,\alpha$ there is a set S stationary in α so that, for all $\beta \in S$ $2^{\aleph_\beta} = \aleph_{\beta+\gamma}$. Then $2^{\aleph_\alpha} = \aleph_{\alpha+\gamma}$.*

Stationary sets permeate modern set theory. For example, they are crucial in forcing and in large cardinal arguments. But we give only one further example of their use, the principle $\diamondsuit$.

Definition 7.94. $\diamondsuit$ is the following statement: there is a sequence $\{a_\alpha : \alpha < \omega_1\}$ so that each $a_\alpha \subseteq \alpha$ and, for each set $a \subseteq \omega_1$, $\{\alpha : a \cap \alpha = a_\alpha\}$ is stationary.

We call the sequence $\{a_\alpha : \alpha < \omega_1\}$ a $\diamondsuit$-sequence, and say that it captures each $a \subseteq \omega_1$ on a stationary set.

Another way of looking at $\diamondsuit$ is to say that the a_α's are guessing initial segments of a, and that stationarily many of them guess correctly.

Fact 7.95. *If $\{a_\alpha : \alpha < \omega_1\}$ is a $\diamondsuit$ sequence, then $\forall a \in [\omega_1]^\omega$ $\{\alpha : a_\alpha = a\}$ is stationary.*

By fact 7.95

Fact 7.96. *If $\diamondsuit$, then CH.*

An important theorem about $\diamondsuit$ is

Theorem 7.97. *$L \models \diamondsuit$.*

The proof of theorem 7.97 requires careful inspection of how L is constructed, and is beyond the scope of this book.

$\diamondsuit$ is useful for many constructions. We content ourselves with

Theorem 7.98. *Assume $\diamondsuit$. Then there is a Suslin tree.*

Proof. Let $\{a_\alpha : \alpha < \omega_1\}$ be a $\diamondsuit$ sequence. We use this sequence to construct a Suslin tree T whose elements will be countable ordinals.

Let Λ be the set of all countable limit ordinals, $\Lambda = \{\delta_\alpha : \alpha < \omega_1\}$ where if $\alpha < \beta$ then $\delta_\alpha < \delta_\beta$. Λ is a club in ω_1. We will require that

$$(\dagger\dagger) \quad \text{each } T_\alpha \subseteq \delta_\alpha, \text{ and each } T_\alpha \text{ is a splitting tree.}$$

At stage α we are given T_α. We check a_α. Is it a maximal antichain in the tree T_α? If not, we do whatever we like to extend T_α to $T_{\alpha+1}$ so $(\dagger\dagger)$ is not violated. If a_α is a maximal antichain in T_α then, for each element β of a_α we let b_β be a branch of T_α with $\beta \in b_\beta$. We have chosen a countably infinite set of such b_β's and $\delta_{\alpha+1} \setminus \delta_\alpha$ is countably infinite, so we extend T_α to $T_{\alpha+1}$ so each $\gamma \in \delta_{\alpha+1} \setminus \delta_\alpha$ is an upper bound for some b_β, and each b_β is bounded by some $\gamma \in \delta_{\alpha+1} \setminus \delta_\alpha$. Thus every element in $T_{\alpha+1} \setminus T_\alpha$ is comparable to some $\beta \in a_\alpha$, hence a_α will remain a maximal antichain in T.

Since T is a splitting tree, it suffices to show that it has no uncountable antichain. Suppose c is an antichain. By exercise 19 in chapter 4, c extends to a maximal antichain a. Let $C = \{\alpha : a \cap T_\alpha$ is a maximal antichain in $T_\alpha\}$.

Subclaim 7.98.1. C is a club in ω_1.

Proof. Since for $\delta \in \Lambda$ $a \cap T_\delta = a \cap \bigcup_{\alpha<\delta} T_\alpha$, C is closed. We show C is unbounded in ω_1.

For each $t \in T$ pick $s_t \in a$ with t, s_t comparable. Let $\alpha < \omega_1$. Let $\gamma_0 = \alpha$. Given γ_n, let $\gamma_{n+1} = \inf\{\gamma : T_{\gamma_n} \cup \{s_t : t \in T_{\gamma_n}\} \subset T_\gamma\}$. Each $\gamma_{n+1} \geq \gamma_n$. By induction, each γ_n is countable. Let $\gamma = \sup\{\gamma_n : n < \omega\}$. Then $\forall t \in T_\gamma$ $s_t \in T_\gamma$, so $a \cap T_\gamma$ is a maximal antichain in T_γ, i.e., $\gamma \geq \alpha$ and $\gamma \in C$. Thus C is unbounded.[115] □

Since $\{a_\alpha : \alpha < \omega_1\}$ is a $\diamondsuit$-sequence there is some $\delta \in C \cap \Lambda$ with $a_\delta = a \cap \delta$. a_δ is maximal in T_δ. Thus, by construction, $a = a_\delta$, i.e., a is countable. □

This is a classic use of $\diamondsuit$, killing off every possible counterexample by killing off some approximation in the $\diamondsuit$-sequence.

Another way of looking at this is that whatever happens in an uncountable set is reflected in some (in fact many) countable subsets. This technique of reflection permeates not only infinite combinatorics on small cardinals via guessing principles, but also is a key element of certain large cardinal principles (e.g., supercompact cardinals) and forcing techniques (e.g., proper forcing).

[115] This kind of construction is called a Lowenheim-Skolem closure argument.

7.7 Exercises

1. Let $[\mathbb{Z}]^2$ be partitioned as follows: $P_0 = \{\{x, y\} : x + y \text{ is even}\}$; $P_1 = \{\{x, y\} : x + y \text{ is odd}\}$. Characterize the sets homogeneous for P_0; for P_1.

2. For any finite n, let $[\mathbb{R}^2]^n$ be partitioned as follows: $P_0 = \{\{x_1, ...x_n\} : |x_1| = ... = |x_n|\}$, where $|x|$ is the distance from the point x to the origin; $P_1 = [\mathbb{R}^2]^n \setminus P_0$. Characterize the sets homogeneous for P_0; for P_1.

3. Given a partial order $\leq$ on a set X, let $[X]^3$ be partitioned as follows: $P_0 = \{\{x, y, z\} : x, y, z$ are linearly ordered by $\leq\}$; $P_1 = \{\{x, y, z\} : x, y, z$ are mutually incompatible$\}$; $P_2 = [X]^3 \setminus (P_0 \cup P_1)$. Characterize the sets homogeneous for P_0; for P_1; for P_2.

4. Use Ramsey's theorem to prove that every infinite partial order has either an infinite chain or an infinite set of pairwise incomparable elements.

5. Use Ramsey's theorem to snow that every infinite set of natural numbers has an infinite subset A so that either $\forall$ distinct $n, m, k \in A$ $n+m+k$ is prime, or $\forall$ distinct $n, m, k \in A$ $n+m+k$ is not prime.

6. Use Ramsey's theorem to show that every infinite linearly ordered set has either an infinite increasing sequence or an infinite decreasing sequence.

7. Prove proposition 7.8.

8. Prove the rest of subclaim 7.14.1.

9. Use the Erdös-Rado theorem (theorem 7.16) to show that if a partial order has size $> 2^\omega$ then it either has an uncountable chain or an uncountable pairwise incomparable subset.

10. Prove the claims of example 7.22.

11. Prove fact 7.23.

12. Prove that if T is a finite partial order, then T is a tree iff $\forall p, q \in T$ p, q are comparable iff p, q are compatible.

13. Consider T the set of increasing well-ordered subsets of $\mathbb{R}$, ordered by end-extension ($\vec{\sigma} \leq \vec{\tau}$ iff $\vec{\sigma}$ is an initial segment of $\vec{\tau}$).

(a) Show that T is a tree.

(b) What is its height?

(c) What is each $T(\alpha)$?

(d) What is each $|T(\alpha)|$?

14. Prove fact 7.19.

15. T is a tree of subsets of x iff $T \subseteq \mathcal{P}(x)$; $t \leq s$ iff $t \supseteq s$; if s, t are incompatible then $s \cap t = \emptyset$; and T is a tree. Let T be a tree of subsets of x.

(a) Show that every branch of T has size at most $|x|$

(b) Show that each $|T_\alpha| \leq |x|$

(c) Show that $C = \{b : b \text{ a branch and } \cap b \neq \emptyset\}$ has size at most $|x|$.

(d) Define $t \in T$ to be decisive if there is $y_t \in x \cap t$ so that $y_t \notin \bigcup T(\alpha)$ where $\alpha = \mathrm{ht}(t) + 1$. Show that there are at most $|x|$ many decisive nodes t.

(e) Show that $|\{t \in T : t \notin \bigcup C \text{ and } t \text{ not decisive}\}| \leq |x|$.

(f) Finally, show that $|T| = |x|$.

16. Using König's lemma, show that if the human race is to survive forever, then some woman must have a female descendant in each subsequence generation.[116]

17. Let T be a tree so each level of T is finite, and if $s < t$ then $t \in s$. Show that T is finite.

18. Prove fact 7.28.

19. Show that a Suslin tree is an Aronszajn tree but not a special Aronszajn tree.

20. (a) Prove fact 7.36.

(b) Show that an uncountable subset of an Aronszajn tree need not be, under the induced order, Aronszajn.

21. Prove proposition 7.43.

22. Prove proposition 7.45.

23. Prove fact 7.57.

24. Show that every countable linear ordering extends to a countable dense linear ordering without endpoints.

25. (a) Show that an uncountable subset of a Suslin line need not be Suslin.

(b) Show that if there is a Suslin line then there is one for which every uncountable subset is a Suslin line.

26. Consider a canonical linear ordering $\leq$ on $B = \{\text{branches of an Aronszajn tree}\}$. Show

(a)B has uncountably many pairwise disjoint intervals.

(b) B does not embed in an order-preserving fashion into $\mathbb{R}$.

27. Prove proposition 7.62(b) and (c).

28. Show directly, without using theorem 7.59, that every measurable cardinal is regular.

29. Let j be as in theorem 7.61.

(a) Show that $M \models j(\kappa)$ is measurable.

(b) Show that if κ is the least measurable cardinal, then $M \models \kappa$ is not measurable.

30. Prove proposition 7.62 (b) and (c).

31. ω^ω can be identified with the branches of the tree $T = \bigcup_{n<\omega} \omega^n$ and 2^ω can be identified with the branches of the tree $S = \bigcup_{n<\omega} 2^n$. S is clearly a subset of T. Show how to embed T as a subset of S so order is preserved, i.e., construct $\varphi : T \to S$ so that $t < t'$ iff $\varphi(t) < \varphi(t')$.

[116]This is assuming that we define "human" genetically, and do not classify, e.g., androids as human.

32. Show that $\mathfrak{b}, \mathfrak{p}$ are regular.

33. A filterbase $\mathcal{B}$ is a base for a filter $\mathcal{F}$ iff $\mathcal{F}$ is the closure of $\mathcal{B}$ under $\supseteq$, i.e., $\forall F \in \mathcal{F}\ \exists B \in \mathcal{B}\ F \supseteq B$. Define $\mathfrak{u} = \inf\{|\mathcal{B}| : \mathcal{B}$ is a base for a non-principal ultrafilter on $\omega\}$. Show that $\mathfrak{u}$ is uncountable.

34. Prove the Rasiowa-Sikorski lemma (theorem 7.76).

35. Use the Rasiowa-Sikorski lemma to prove that $\mathfrak{p}$ is uncountable, i.e., recast the proof of theorem 1.43 in terms of the Rasiowa-Sikorski lemma, using a partial order and dense sets.

36. A scale is a dominating family in ω^ω well-ordered by $\leq^*$. Show that if $\mathfrak{b} = \mathfrak{d}$ there is a scale of order-type $\mathfrak{d}$.

37. Prove directly, without using our knowledge of how cardinal invariants are related to each other, that under $\mathrm{MA}_{\sigma-\text{centered}}$

(a) $\mathfrak{b} = \mathfrak{c}$

(b) $\mathfrak{a} = \mathfrak{c}$

38. Assume $\mathfrak{b} = \mathfrak{d}$. Show there is a family $\{a_\alpha : \alpha < \mathfrak{b}\}$ where each $a_\alpha \subset \omega$ and

(a) if $\alpha < \beta$ then $a_\alpha \supseteq^* a_\beta$.

(b) $\{a_\alpha : \alpha < \mathfrak{b}\}$ has no pseudo-intersection.[117]

39. A Δ-system is a family of sets $\mathcal{A}$ so that for some set a, $a \subseteq \bigcap \mathcal{A}$ and if $A, B \in \mathcal{A}$, $A \neq B$, then $A \cap B = a$. (a is called the root of the Δ-system $\mathcal{A}$.) The purpose of this exercise is to prove the Δ-system lemma: if $\mathcal{A}$ is an uncountable collection of finite subsets then there is an uncountable Δ-system $\mathcal{B} \subseteq \mathcal{A}$.[118]

So let $\mathcal{A}$ be an uncountable collection of finite sets. Without loss of generality $\bigcup \mathcal{A} = \omega_1$.

(a) Show there is $n < \omega$ and an uncountable set $\mathcal{A}_0 \subseteq \mathcal{A}$ so $\forall A \in \mathcal{A}_0\ |A| = n$.

(b) For each $A \in \mathcal{A}_0$ let $A = \{\alpha_{i,A} : i < n\}$. Show that there is $k < n$ so $\{\alpha_{k,A} : A \in \mathcal{A}_0\}$ is uncountable.

(c) Let k be least so $\{\alpha_{k,A} : A \in \mathcal{A}_0\}$ is uncountable. Show that there is an uncountable set $\mathcal{A}_1 \subseteq \mathcal{A}_0$ so that $\{\{\alpha_{j,A} : j \geq k\} : A \in \mathcal{A}_1\}$ is pairwise disjoint.

(d) Let k be as in (c). Show there is an uncountable set $\mathcal{A}_2 \subseteq \mathcal{A}_1$ and a finite set a (possibly empty) with $\{\alpha_{i,A} : i < k\} = a$ for all $A \in \mathcal{A}_2$. (I.e., $\mathcal{A}_2$ is the uncountable Δ-system contained in $\mathcal{A}$.)

40. As a counterbalance to the Δ-system lemma, show that there is a family $\mathcal{A} \subset [\aleph_{\omega_1}]^2$ with $|\mathcal{A}| = \aleph_{\omega_1}$ and if $\mathcal{B} \subseteq A$ and $\mathcal{B}$ is a Δ-system, then $|\mathcal{B}| < \aleph_{\omega_1}$.

41. (a) Use the Δ-system lemma to show that the following partical order is ccc: $\mathcal{P} = \{\sigma : \mathrm{dom}\ \sigma$ is a finite subset of ω_1, range $\sigma \subset \omega\}$.

(b) Assume MA + $\neg$CH. Using the partial order of (a), show that if $\mathcal{F} \in [\omega^{\omega_1}]^{<\mathfrak{c}}$ then there is

[117] $\mathfrak{t}$ is the smallest cardinal for which there is such a descending sequence with no pseudo-intersection.This exercise shows that $\mathfrak{t} \leq \mathfrak{b}$.

[118] There is also a more general Δ-system lemma.

a function $g \in \omega^{\omega+1}$ so that $\forall f \in F\ |\{\alpha : g(\alpha) = f(\alpha)\}| = \omega_1 = |\{\alpha : g(\alpha \neq f(\alpha\}|$.

42. An ultrafilter U on ω is Ramsey iff for all P a partition of ω into infinite sets, either $P \cap U \neq \emptyset$ or $\exists b \in U\ \forall \alpha \in P\ |a \cap b| \leq 1$. Show that under CH there is a Ramsey ultrafilter.

43. A Canadian tree is a tree of size ω_1 and height ω_1 with at least ω_2 many branches. Prove that if CH holds then there is a Canadian tree.[119]

44. Prove that if κ is a cardinal and $\mathrm{cf}(\kappa) = \lambda$ then the intersection of fewer than λ many clubs in κ is a club.

45. Prove that if κ is a regular cardinal and $\{c_\alpha : \alpha < \kappa\}$ is a collection of club sets in κ then $c = \{\beta : \forall \alpha < \beta\ \beta \in c_\alpha\}$ is club.[120]

46. Show that if κ has countable cofinality then $S \subseteq \kappa$ is a stationary subset of κ iff $\kappa \setminus S$ is bounded below κ.

47. Show that if S is stationary in κ, where κ has uncountable cofinality, then $S \cap C$ is stationary for all C club in κ.

48. Prove fact 7.95.

49. Prove fact 7.96.

50. Let $\{a_\alpha : \alpha < \omega_1\}$ be a $\diamondsuit$-sequence. Show that for no set a is $\{\alpha : a = a_\alpha\}$ club.

51. (a) Show that the non-stationary subsets of κ form an ideal (see definition 1.50).

(b) If κ is regular, show that the union of fewer than κ non-stationary sets is non-stationary.[121]

52. S is stationary co-stationary in κ iff both S and $\kappa \setminus S$ are stationary. Let κ be regular. How many stationary co-stationary subsets of κ are there?

53. In the spirit of using a nuclear weapon to kill a mouse, use $\diamondsuit$ to get a short proof of Solovay's theorem for ω_1, i.e., if $\diamondsuit$ holds, then there is a pairwise disjoint family $\mathcal{F}$ of stationary subsets of ω_1 where $|\mathcal{F}| = \omega_1$.[122]

[119] "ZFC + there are no Canadian trees" is equiconsistent with "ZFC + there is an inaccessible cardinal."
[120] c is called the diagonal intersection of $\{c_\alpha : \alpha < \kappa\}$.
[121] We say that the non-stationary ideal is κ-additive.
[122] Of course the conclusion holds without any extra axioms...

8 Bibliography

Necessarily, a book at this level leaves a lot out. A selective bibliography is included here for the reader interested in learning more. Some acquaintance with the model-theory portions of the logic books is essential for understanding nearly all of the set-theory books mentioned. All books are listed in their most current edition. Some are out of print. That is why we have libraries.

Mathematical Logic

Enderton, H., *A Mathematical Introduction to Logic*, Academic Press, 2001

Hodel, R., *An Introduction to Mathematical Logic*, PWS, 1995

Shoenfield, J., *Mathematical Logic*, A.K. Peters, 2001

Set theory

Just, W. and Weese, M., *Discovering Modern Set theory I: The Basics*, American Mathematical Society, 1996

Just, W. and Weese, M., *Discovering Modern Set theory II: Set-Theoretic Tools for Every Mathematician*, American Mathematical Society, 1997

Jech, T., *Set Theory*, Springer, 2006

Forcing

Kunen, K., *Set Theory: An Introduction to Forcing and Independence Proofs*, North-Holland, 1983

Combinatorics

Erdös, P., Máté. A., Hajnal, A., Rado, R., *Combinatorial Set Theory: Partition Relations for Cardinals*, North Holland, 1984

Nash-Williams, C., *Combinatorial Set Theory*, North-Holland, 1977

***L* and related models**

Devlin, K., *Constructibility*, Springer-VErlag, 1974

Dodd, A.J., *The Core Model*, Cambridge University Press, 1982

Large Cardinals

Drake, F.R., *Set Theory: An Introduction to Large Cardinals*, North-Holland, 1974

Kanamori, A., *The Higher Infinite: Large Cardinals in Set Theory from their Beginnings*, Springer-Verlag, 2008

History

Dauben, J.W., *Georg Cantor: His Mathematics and Philosophy of the Infinite*, Harvard University Press, 1990

Hallett, M., *Cantorian Set Theory and Limitation of Size*, Oxford University Press, 1986

van Heijenoort, J., ed., *From Frege to Födel: A Source Book in Mathematical Logic, 1879 - 1931*, Harvard University Press, 2002

Moore, Gergeory H., *Zermelo's Axiom of Choice: Its Origins, Development and Influence*, Springer-VErlag, 1982