Lecture Notes: Models & Sets

(MATH3120/5120M)

Asaf Karagila

Last Update:
December 19, 2025

Introduction

These notes are the lecture notes for the classes “Models and Sets” (3120) and “Advanced Models and Set” (5120M). Each chapter roughly corresponds to a week, with the exception of Chapters 5 & 6 which will be combined into a single week.

The notes were developed over three years of teaching this class. The preliminary requirements are minimal, but generally require some understanding of logic and some mathematical knowledge in order to understand the examples.

Please inform me of any mistakes, typos, and otherwise things that could improve the notes. Your help will be appreciated. On that note, I’d like to thank Aaron Katz who sent a list of corrections and Matthew Choy for his careful attention. As well as Aris Papadopoulus, Vincenzo Mantova, Andrew Brooke-Taylor, Connie Bromham, and Calliope Ryan-Smith for making helpful suggestions.

Contents

1 Sets, classes, and collections 1
- 1.1 What is a set? 1
- 1.2 Enter the Axiom: Zermelo's set theory 5
- 1.3 Sets as a universal interpreter 7

2 Order! Order! 12
- 2.1 Partial and total orders 12
- 2.2 Order-preserving functions 14
- 2.3 Products of orders 16

3 Well, well, well... foundedness 18
- 3.1 Well-founded relations 18
- 3.2 Finiteness done "wrong" 19
- 3.3 Induction and recursion in general 20
- 3.4 Some general comments on general recursion 22
- 3.5 Well-orders 23
- 3.6 Induction and recursion for well-ordered sets 24

4 Ordinal numbers 27
- 4.1 The von Neumann ordinals 27
- 4.2 The Natural Numbers 31
- 4.3 Rational approach to the real numbers that is too complex 33

5 Cardinal numbers 35
- 5.1 What do you mean "how many elements?" 35
- 5.2 Cardinalities of well-ordered sets 39
- 5.3 Cardinal numbers 41

6 The Axiom of Choice 43
- 6.1 "I choose to believe what I was programmed to believe!" 43
- 6.2 The Big Six 44
- 6.3 Applications 46

7 First-Order Logic 50
- 7.1 Syntax 50
- 7.2 Semantics 53
- 7.3 Structures, theories, and models 54
- 7.4 Complete theories 56

8 Elementary, my dear Moose. 58
8.1 Elementarity . . . 58
8.2 Łoś's Theorem . . . 60
8.3 Compactness . . . 63

9 Skolem in the House of Lions 65
9.1 Going down? . . . 65
9.2 Going up? . . . 67
9.3 Definability and lack thereof . . . 68
9.4 More uses for automorphisms . . . 70

10 Where the Moose and the Lions dread to step 72
10.1 Second-order logic . . . 72
10.2 Infinitary logic . . . 74
10.3 There exists a quantifier . . . 76

11 You and V 79
11.1 Axiom of Foundation . . . 79
11.2 The universe as we know it . . . 81
11.3 The L . . . 84

12 Models of fragments of ZFC 87
12.1 Setting the stage . . . 87
12.2 Powerless Models . . . 88
12.3 Not beyond infinity . . . 90
12.4 Replacement Killers . . . 90
12.5 Elementary submodels of fragments and their stories . . . 91

A Selected solutions 93

Chapter 1

Sets, classes, and collections

Chapter Goals

In this chapter we will learn about

- The naive concept of a set, as well as about some of the axioms which govern the behaviour of sets from the modern perspective.
- The difference between sets and proper classes.
- The concept of ordered pairs including how this concept can be modelled using sets.
- The basics of relations, equivalence relations, and functions.

1.1 What is a set?

Sets are mathematical objects which are collections of mathematical objects. In other words, sets are used to formalise the notion of a collection in mathematics. Since the need to collect objects is universal to all mathematical disciplines it is a good idea to have a naive understanding of what sets are. We can understand, for example, a line, or a circle, or a disk, as collection of points.

The notion of a collection, and therefore that of a set, was present from the very beginning of mathematics. It was not formalised and studied, however, until the 19th century, when *Georg Cantor* began the study of what is now known as **set theory**. Initially studying the concepts of well-orders and cardinality, Cantor's work became popular and his concept of sets became ever more important in the study of the foundations of mathematics.

We are not going to go into the philosophical question of what is a set, exactly, but the goal of set theory is to describe the properties which sets have. Much like that the axioms of a field guarantee the existence of a multiplicative inverse for non-zero elements, the axioms of set theory guarantee that if we are given some sets, then another set with some wanted property exists. Which are the correct properties of sets, then? This is left for the philosophers to debate, but we will remark that there is one major set theory, which we will see later. But others exist as well, and they are of interest in certain parts of the mathematical world.

1.1.1 Naive beginnings

Notation 1.1. We will write $\{\dots\}$ to denote a collection of mathematical objects.

For example, $\{0, 1, 2\}$ denotes the collection which includes **exactly** the numbers 0, 1, and 2. Nothing more, and nothing less than that.

$$\mathbb{N} = \{0, 1, 2, \dots\},$$

denotes the collection which contains exactly all the natural numbers. Similarly,

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

denotes the collection whose members are exactly the integers: the natural numbers and their additive inverses. We will use $\mathbb{Q}$ to denote the collection of all the rational numbers and $\mathbb{R}$ will denote the collection of all the real numbers.

Notation 1.2. If P is a 'property' (e.g., being a natural number, being a positive fraction, etc.), then

$$\{x \mid P(x)\}$$

is the collection of exactly those objects which have the property P.

For example, $\{x \mid x \text{ is an integer power of } 2\}$ is the collection $\{1, 2, 4, 8, 16, \dots\}$. We will see later what exactly constitutes as a property.

Notation 1.3. If A is a collection, we write "$a \in A$" to mean that a is one of the objects inside the collection A and we say that a is a *member* or an *element* of A. If a is not an element of A, we write "$a \notin A$".

So, for example, $0 \in \mathbb{N}$ and $0 \notin \{x \mid x \text{ is a positive integer}\}$.

Notation 1.4. We use $\varnothing$ to denote the collection with no elements. This is the collection which satisfies $x \notin \varnothing$ for any x.

Axiom: Extensionality

Two collections A and B are equal if and only if they have exactly the same elements.

The role of equality in modern mathematics is simple: two objects are equal if and only if they are the same object. The Axiom of Extensionality, therefore, says that the description of the property, while important, is not as important as what are the actual objects satisfying it.

So, for example, $\mathbb{N} = \{x \mid x \in \mathbb{Z} \text{ and } x \text{ is non-negative}\}$. On the other hand, $\mathbb{N} \neq \varnothing$, since $0 \in \mathbb{N}$ but $0 \notin \varnothing$.

Exercise 1.1. Show that if A, B, and C are collections, then

- $A = A$;
- if $A = B$, then $B = A$; and
- if $A = B$ and $B = C$, then $A = C$.

Exercise 1.2. Show that $\{0, 1, 2, 3\} = \{x \mid x \in \mathbb{N} \text{ and } x < 4\}$.

Notation 1.5. If A and B are collections, we write $A \subseteq B$ if every element of A is also an element of B. Otherwise, we write $A \nsubseteq B$. Not to be confused, we will use $A \subsetneq B$ to mean "$A \subseteq B$ and $A \neq B$".

Using the $\subseteq$ notation we can rewrite the Axiom of Extensionality as "$A = B$ if and only if $A \subseteq B$ and $B \subseteq A$".

Proposition 1.6. *If A is a collection, then $\varnothing \subseteq A$.*

Proof. Suppose that $\varnothing \nsubseteq A$. Then there is some x such that $x \in \varnothing$ and $x \notin A$. However $\varnothing$ is the empty collection, and therefore there cannot be such x. Therefore $\varnothing \subseteq A$. $\square$

Remark

An argument is *vacuous* if it holds simply due to the lack of possible counterexamples. The claim that $\varnothing \subseteq A$ is *vacuously true*, simply since there cannot be an example to the contrary.

Note that as an immediate corollary, if $A \subseteq \varnothing$, then $A = \varnothing$. Therefore, by the Axiom of Extensionality, $\varnothing$ is unique, and using "the" is now justified.

1.1.2 Boolean operations

Definition 1.7. Let A and B be two collections:

- The *union* (of A and B) is $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$.
- The *intersection* (of A and B) is $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$.
- The *difference* (of A and B) is $A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}$.

Exercise 1.3. Check that for any two collections A and B, $A \cup B = B \cup A$ and $A \cap B = B \cap A$.

Exercise 1.4. Find two collections A and B such that $A \setminus B = B \setminus A$, and two collections C and D such that $C \setminus D \neq D \setminus C$.

Exercise 1.5. Show that for any two collections A and B, $A \cap B \subseteq A \subseteq A \cup B$.

Exercise 1.6. Show that for any two collections A and B, $A \subseteq B$ if and only if $A \setminus B = \varnothing$ and $A \cap B = \varnothing$ if and only if $A \setminus B = A$. (Visit solution)

Exercise 1.7. Show that for any collection A, $A \cup \varnothing = A$ and $A \cap \varnothing = \varnothing$.

Exercise 1.8. Show that for any collections A, B, and C, $A \cap (B \cap C) = (A \cap B) \cap C$ and $A \cup (B \cup C) = (A \cup B) \cup C$.

The above exercise allows us to omit the parentheses and simply write $A \cup B \cup C$ when we need to. On the other hand the following exercise shows that we cannot omit the parentheses when we are mixing unions and intersections.

Exercise 1.9. Show that for any collections A, B, and C, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Exercise 1.10. Let $A = \{1, 2, 3\}$ and $B = \{2, 1, 4\}$. Calculate $A \cap B$, $A \cup B$, $A \setminus B$, and $B \setminus A$.

Definition 1.8. The *symmetric difference* of two collections A and B is $A \triangle B = (A \cup B) \setminus (A \cap B)$.

Exercise 1.11. Prove that for any two collections A and B, $A \triangle B = B \triangle A$.

Proposition 1.9. $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Proof. To show the equality, we need to show that $A \triangle B \subseteq (A \setminus B) \cup (B \setminus A)$ and vice versa. We begin by showing that $A \triangle B \subseteq (A \setminus B) \cup (B \setminus A)$.

Let $x \in A \triangle B$, then by definition $x \in A \cup B$ and $x \notin A \cap B$. Since $x \in A \cup B$, either $x \in A$ or $x \in B$. If $x \in A$, since $x \notin A \cap B$, it must be that $x \notin B$, and therefore $x \in A \setminus B$, and so $x \in (A \setminus B) \cup (B \setminus A)$. Similarly, if $x \in B$, then $x \notin A \cap B$, and therefore $x \notin A$, and the conclusion follows.

In the other direction, let $x \in (A \setminus B) \cup (B \setminus A)$. Suppose that $x \in A \setminus B$, then $x \in A$ and so $x \in A \cup B$. On the other hand, $x \notin B$, so $x \notin A \cap B$. Therefore, $x \in (A \cup B) \setminus (A \cap B) = A \triangle B$. The argument is similar for the case that $x \in B \setminus A$. □

Exercise 1.12. Show that for any collection A, $A \triangle \varnothing = A$, $A \triangle A = \varnothing$. (Visit solution)

Exercise 1.13. Show that for any three collections, A, B, and C, $A \triangle (B \triangle C) = (A \triangle B) \triangle C$.

1.1.3 Frege's failed attempt

Friedrich Ludwig *Gottlob Frege* had made an attempt to base all of mathematics on the foundation of logic. His naive approach to set theory can be understood from a modern perspective as a set theory with two "axioms". The first is the Axiom of Extensionality which we saw earlier. The second is Comprehension.

Axiom: Comprehension

Suppose that P is a property, then $\{x \mid P(x)\}$ is a set.

In other words, every collection, in the eyes of Frege, is a set. More accurately, any collection that "exists in the scope of the mathematical universe" is a set. And in this setting we can think of all mathematical objects as sets as well.

Theorem 1.10. *Given a property P, the set $\{x \mid P(x)\}$ is unique.*

Proof. Suppose that A and B are instances of the Axiom of Comprehension for the same property, it follows that if $a \in A$, then $P(a)$ holds, and therefore $a \in B$. Similarly, if $b \in B$, then $P(b)$ holds, and so $b \in A$. Therefore the by the Axiom of Extensionality, $A = B$. □

Now that we have the word "set" in our language, we can use it. If $B \subseteq A$, we say that B is a **subset** of A.

Definition 1.11. Suppose that A is a set, the *power set* of A is $\mathcal{P}(A) = \{B \mid B \subseteq A\}$.

The power set, therefore, is the set of all the subsets of A.

Exercise 1.14. Compute $\mathcal{P}(\varnothing)$, $\mathcal{P}(\{\varnothing\})$, and $\mathcal{P}(\{\varnothing, \{\varnothing\}\})$. (Visit solution)

Exercise 1.15. Show that for any set A, $\mathcal{P}(A) \neq \varnothing$.

In Frege's naive set theory, every set has a power set, since the collection is a well-defined property. Similarly, the unions and intersections of sets are well-defined properties, so the collections they define are also sets.

Unfortunately, Frege's naive set theory did not last very long. It did not take a very long time for many paradoxical observations to appear, and as mathematics became more and more formal and reliant on logic for its rigour, the most famous of them was formalised by *Bertrand Russell*.

Theorem 1.12 (Russell's paradox). *Let $R = \{x \mid x \notin x\}$, then R cannot be a set.*

Proof. Suppose that R is a set, then $R \in R$ if and only if $R \notin R$. This is of course impossible. □

This means that some collections are not (and cannot be) sets. Frege's Axiom of Comprehension is too strong of an axiom, and we must find a way to "fix" the definition of a set, or at least the list of properties that govern the behaviour of sets.

Remark

We use the term "**class**" to refer to a collection (defined by a property) in a more formal setting. All sets are classes, and some classes are sets. We use the term "**proper class**" to mean that a particular class is not a set. For example, R from Russell's paradox is a proper class.

1.2 Enter the Axiom: Zermelo's set theory

The first notable attempt to fix Frege's set theory came from *Ernst Zermelo*. He proposed axioms which define the properties of a set, and while he did not intend to fix Frege's naive set theory, but instead to prove the Well-Ordering Theorem, his work turned out to be seminal to modern set theory and mathematics. These axioms were augmented by other mathematicians later, and most notably *Abraham Fraenkel*, to what is now known as the **Zermelo–Fraenkel axioms of set theory**. We will slowly incorporate these into our discourse. We begin with a few basic axioms.

Remark

One striking property of modern set theory is that **every mathematical object** is a set. This include 0 and $\frac{1}{4}$ and $\sqrt{38}$. We can, and should, wonder what are the elements of e or π, and we will see later why this question, while interesting on its surface, is meaningless to an extent. But for now, we will subtly put this fact in the background.

The problem with Frege's naive set theory was never the Axiom of Extensionality. Indeed, there is a good philosophical argument to be made that this axiom is the bare minimum wanted from anything which is supposed to be a set, or even a collection. And so we will keep this axiom for the rest of the course.

Axiom: Empty Set

$\varnothing$ is a set.

This axiom is useless. It simply states that some sets exist. We will see that it is a consequence of other axioms. Nevertheless, we do want to claim that some sets do exist.

Axiom: Pairing

Let x, y be two mathematical objects, then the **unordered pair** $\{x, y\}$ is a set.

Remark

Note that the role of x and y is symmetrical, giving credence to the term "unordered". So in particular, as a consequence of the Axiom of Extensionality, $\{x, y\} = \{y, x\}$.

Exercise 1.16. Suppose that a is a mathematical object, then $\{a\}$ is a set.

Exercise 1.17. Show that $\varnothing \neq \{\varnothing\}$ and $\{\varnothing\} \neq \{\{\varnothing\}\}$.

Notation 1.13. Suppose that A is a set, $\bigcup A$ denotes $\{x \mid \text{There exists } a \in A \text{ such that } x \in a\}$.

As we said, in the context of modern set theory, all objects are sets, which means that the question of whether or not A contains "only sets" is moot. However, even if we do want to allow non-set objects to exist, we still can notice that by requiring $x \in a$ we invariably restrict ourselves to the elements of A which are already sets.

Axiom: Union

If A is a set, then $\bigcup A$ is a set.

Theorem 1.14. *If A and B are sets, then $A \cup B$ is a set.*

Proof. By the Axiom of Pairing, $\{A, B\}$ is a set, and by the Axiom of Union $\bigcup\{A, B\}$ is a set. It is enough to check that $A \cup B = \bigcup\{A, B\}$. □

Exercise 1.18. Complete the proof of the above theorem.

One of the important properties we want to have about sets is to be able to define and carve them out. For example, we want to be able to carve out the set of all even integers out of $\mathbb{N}$ and the set of all rational numbers between 0 and 1 out of $\mathbb{Q}$. For these sets to exist we need the following axiom.

Axiom: Separation

For any property P and set A, $\{a \mid a \in A \text{ and } P(a)\}$ is a set.

Notation 1.15. To emphasise the importance of A, we write $\{a \in A \mid P(a)\}$.

Remark

The property P in the Axiom of Separation is allowed to have parameters, which are additional predefined objects that we can use when we express the property. We will usually omit them from the discussion for the sake of readability.

Theorem 1.16. *If A is a set, then there is a set $B \subseteq A$ such that $B \notin A$.*

Proof. Let $B = \{a \in A \mid a \notin a\}$, by the Axiom of Separation B is a set. We now repeat Russell's paradox: If $B \in A$, then $B \in B$ if and only if $B \notin B$. Therefore $B \notin A$, as wanted. □

Theorem 1.17. *The class of all sets is a proper class.*

Proof. Let V denotes the class of all sets. If V is a set, then by the previous theorem there is some $B \subseteq V$ such that $B \notin V$. But since B is a set, B must be an element of V. □

Exercise 1.19. If A and B are sets, then $A \cap B$, $A \setminus B$, and $A \triangle B$ are also sets.

Axiom: Power Set

If A is a set, then $\mathcal{P}(A) = \{B \mid B \subseteq A\}$ is a set.

Remark

Perhaps the most important thing to understand about the Axiom of Power Set, and in general about set theoretic foundations of mathematics, is that we are given a universe of sets. What the axioms do is describe to us which properties this universe and its objects satisfy. The Axiom of Power Set does not tell us which are the subsets of a set, nor how many of them exist. All that it is telling us is that we can collect *all* of the subsets into a single set.

Exercise 1.20. For any set A, $\mathcal{P}(A) \nsubseteq A$. Find an example of a set A such that $A \subseteq \mathcal{P}(A)$. (Visit solution)

Exercise 1.21. Prove Theorem 1.17 using the Power Set axiom and the previous exercise.

1.3 Sets as a universal interpreter

1.3.1 Pairs and products

Definition 1.18. An *ordered pair* is a mathematical object which has two elements, "left" and "right". We denote the ordered pair whose left element is x and whose right element is y by $\langle x, y\rangle$. The defining property of an ordered pair is

$$\langle x, y\rangle = \langle a, b\rangle \text{ if and only if } x = a \text{ and } y = b.$$

Theorem 1.19. *We can interpret the concept of an ordered pair using sets.*

Proof. We define $\langle x, y\rangle$ to denote the set $\{\{x\}, \{x, y\}\}$. Firstly, this is a set, since $\{x\}$ and $\{x, y\}$ are both sets, and therefore $\{\{x\}, \{x, y\}\}$ is also a set.

Next we need to check that this definition satisfies the defining property of ordered pairs. Note that if $x = a$ and $y = b$, then by the Axiom of Extensionality, $\langle x, y\rangle = \langle a, b\rangle$. We only have to check the other implication. Suppose that $\{\{x\}, \{x, y\}\} = \{\{a\}, \{a, b\}\}$. By the Axiom of Extensionality we know that either $\{x\} = \{a\}$ or $\{x\} = \{a, b\}$.

Case I: $\{x\} = \{a\}$. In that case, again by the Axiom of Extensionality, $x = a$. Therefore it must be that $\{x, y\} = \{a, b\}$ and since $x = a$, it must be that $y = b$, as wanted.

Case II: $\{x\} = \{a, b\}$. In that case, $x = a = b$. So, $\{a\} = \{a, b\} = \{x\}$. It now follows that $\{\{x\}, \{x, y\}\} = \{\{x\}\}$. Therefore $\{x, y\} = \{x\}$ and therefore $x = y$. So we have that $y = x = a = b$, as wanted again. □

Remark

The above definition of an ordered pair is known as the **Kuratowski ordered pair**. It is not the only way of defining an ordered pair using sets, but it is an incredibly convenient one, and we will use this one going forward. However, the vast majority of the proof can be thought of as "templates" where we can plug in the different definitions of various objects.

Exercise 1.22. Check whether or not $\{\{\varnothing, \{x\}\}, \{\{y\}\}\}$ defines an ordered pair.

Exercise 1.23. Check whether or not $\{x, \{\varnothing, y\}\}$ defines an ordered pair. (Visit solution)

Exercise 1.24. We can "iterate" the concept of an order pair to define an ordered triplet: $\langle x, y, z\rangle = \langle x, \langle y, z\rangle\rangle$, for concreteness sake. Show that this definition works. Namely, $\langle x, y, z\rangle = \langle a, b, c\rangle$ if and only if $x = a$, $y = b$, and $z = c$. On the other hand, show that Kuratowski's definition does not extend to triplets. In other words, using the definition $\langle x, y, z\rangle = \{\{x\}, \{x, y\}, \{x, y, z\}\}$ does not define an ordered triplet.

Definition 1.20. Given sets, A and B, their *Cartesian product* is

$$A \times B = \{\langle a, b\rangle \mid a \in A \text{ and } b \in B\}.$$

Theorem 1.21. *For any two sets, A and B, $A \times B$ is a set.*

Proof. $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B))$. □

Exercise 1.25. Complete the above proof.

Exercise 1.26. Compute $\{1, 2\} \times \{3, 4\}$.

Exercise 1.27. Show that for any set A, $A \times \varnothing = \varnothing \times A = \varnothing$. Moreover, if $A \times B = \varnothing$, then either $A = \varnothing$ or $B = \varnothing$.

Exercise 1.28. Prove or disprove: $A \cap (B \times C) = (A \cap B) \times (A \cap C)$. (Visit solution)

1.3.2 Relations

Definition 1.22. We say that R is a *relation* if R is a set of ordered pairs.

Definition 1.23. Let R be a relation, the *domain* of R, denoted by $\operatorname{dom}(R)$, is

$$\operatorname{dom}(R) = \{a \mid \text{There is some } b \text{ such that } \langle a, b\rangle \in R\}.$$

The *range* of R, denoted by $\operatorname{rng}(R)$, is

$$\operatorname{rng}(R) = \{b \mid \text{There is some } a \text{ such that } \langle a, b\rangle \in R\}.$$

We say that R is a *relation on A* if $\operatorname{dom}(R) \cup \operatorname{rng}(R) \subseteq A$.

We always have that $R \subseteq \operatorname{dom}(R) \times \operatorname{rng}(R)$, but these need not be equal. For example, $R = \{\langle 1, 2\rangle, \langle 2, 1\rangle\}$ has $\operatorname{dom}(R) = \operatorname{rng}(R) = \{1, 2\}$. But $\{1, 2\} \times \{1, 2\} = \{\langle 1, 1\rangle, \langle 1, 2\rangle, \langle 2, 1\rangle, \langle 2, 2\rangle\}$ so they are not equal.

Exercise 1.29. Find all relations R such that $\operatorname{dom}(R) \subseteq \{0, 1\}$ and $\operatorname{rng}(R) \subseteq \{\varnothing, \mathbb{N}\}$.

Notation 1.24. If R is a relation we write $a \mathrel{R} b$ to mean $\langle a, b\rangle \in R$, and $a \not\mathrel{R} b$ otherwise.

Definition 1.25. Let R be a relation on a set A.

1. R is *reflexive (on A)* if for every $a \in A$, $a\ R\ a$.
2. R is *symmetric* if whenever $a\ R\ b$, then $b\ R\ a$.
3. R is *transitive* if whenever $a\ R\ b$ and $b\ R\ c$, then $a\ R\ c$.

Remark

Here we see an interesting distinction of properties: being "reflexive" is an *extrinsic* property of a relation. Namely, it only makes sense when additional information is added in the form of the set A on which the relation is defined. So "R is a reflexive relation" is not a meaningful statement on its own. In contrast, symmetric and transitive are *intrinsic* properties which only depend on the relation itself.

Definition 1.26. If R is a relation on A which is reflexive, symmetric, and transitive we say that R is an *equivalence relation (on A)*.

Exercise 1.30. Show that the relation E defined on $\mathbb{Z}$ by $a\ E\ b$ if and only if $a - b$ is even is an equivalence relation. (Visit solution)

Exercise 1.31. For any two distinct properties {reflexive, symmetric, transitive}, find a relation on a set A satisfying exactly those two and not the third.

Definition 1.27. Suppose that E is an equivalence relation on a set A and $a \in A$. The *equivalence class of a* is $a/E = \{b \in A \mid a\ E\ b\}$. The *quotient set* is $A/E = \{a/E \mid a \in A\}$.

Theorem 1.28. *Let E be an equivalence relation on A. The following are equivalent:*

1. $a\ E\ b$.
2. $a/E = b/E$.

Proof. Assume that $a\ E\ b$, we will show that $a/E = b/E$. For this we will show that $a/E \subseteq b/E$ and that $b/E \subseteq a/E$. Let $c \in a/E$, then $a\ E\ c$ by definition. Moreover, since $a\ E\ b$, by symmetry, $b\ E\ a$. Therefore, by transitivity we have $b\ E\ c$, and so $c \in b/E$. In the other direction, if $c \in b/E$, we have that $a\ E\ b$ and $b\ E\ c$, and therefore $a\ E\ c$, so $a\ E\ c$ as wanted.

In the other direction, assume now that $a/E = b/E$. Since E is reflexive on A, $b\ E\ b$, so $b \in b/E = a/E$. Therefore $a\ E\ b$ as wanted. □

Exercise 1.32. Let E be the equivalence relation from Exercise 1.30. Compute $6/E$ and $\mathbb{Z}/E$.

Definition 1.29. If R is a relation, the inverse relation, R^{-1} is $\{\langle b, a\rangle \mid a\ R\ b\}$.

Exercise 1.33. Show that R is symmetric if and only if $R = R^{-1}$.

1.3.3 Functions

Definition 1.30. We say that a relation R is a *function* if it satisfies the following property: if $\langle a, b\rangle \in R$ and $\langle a, c\rangle \in R$, then $b = c$.

Notation 1.31. If f is a function, we write $f\colon A \to B$ to denote that $\operatorname{dom}(f) = A$ and $\operatorname{rng}(f) \subseteq B$. For $a \in \operatorname{dom}(f)$, we write $f(a)$ to denote the unique b such that $\langle a, b\rangle \in f$.

For example $\{\langle n, n+1\rangle \mid n \in \mathbb{N}\}$ is a function, it is "the successor function" mapping a natural number n to its successor, $n+1$.

Remark

We read the notation above as "f is a function from A to B" and we say that B is "the codomain". Note this is an extrinsic property of the function, in particular, we may enlarge B as well, so the notion of a codomain is not unique and the article "the" is formally incorrect, however in the context of the notation it is understood as the codomain of interest. This is not the only way to encode a function, and in other field of mathematics, the sets A and B are intrinsic to f itself.

Definition 1.32. Let $f\colon A \to B$ be a function.

1. f is *injective* (or 1-1) if whenever $a \neq a'$ we have that $f(a) \neq f(a')$.
2. f is *surjective* (or onto) if whenever $b \in B$, there is some $a \in A$ such that $f(a) = b$.
3. f is *bijective* if it is injective and surjective.

Note that the definition of injectivity can be recast as "if $f(a) = f(a')$, then $a = a'$".

Remark

Note that while injectivity is an intrinsic property, surjectivity is extrinsic and depends very much on B. However, since we use the notation $f\colon A \to B$, the context is clear that we mean "onto B".

Notation 1.33. *The identity function* (on a set A), $\operatorname{id}\colon A \to A$, is the function $\operatorname{id}(a) = a$.

Exercise 1.34. Show that if F is both a function and an equivalence relation, then $F = \operatorname{id}$.

Notation 1.34. Suppose that $f\colon A \to B$ and $X \subseteq A$. We write $f[X] = \{f(x) \mid x \in X\}$ (the *direct image*) and we write $f \restriction X = \{\langle x, f(x)\rangle \in f \mid x \in X\}$ (the *restriction*).

Exercise 1.35. Prove or disprove for $f\colon A \to B$. $f[X] \cup f[Y] = f[X \cup Y]$; $f[X] \cap f[Y] = f[X \cap Y]$; $f[X] \setminus f[Y] = f[X \setminus Y]$.

Proposition 1.35. *Let f be a function, then f is injective if and only if f^{-1} is a function.*

Proof. Let A denote $\operatorname{dom}(f)$ and let B denote $\operatorname{rng}(f)$. Suppose that f is injective, and suppose that $\langle b, a\rangle$ and $\langle b, a'\rangle$ both belong to f^{-1}. By the definition of f^{-1} this means that $\langle a, b\rangle$ and $\langle a', b\rangle$ are both in f, or in the simplified notation, $f(a) = b$ and $f(a') = b$. Since f is injective, it means that $a = a'$, which means that f^{-1} is a functional relation and is therefore a function.

In the other direction, suppose that f^{-1} is a function, and let $a, a' \in A$ be such that $f(a) = f(a') = b$. Then $\langle b, a\rangle, \langle b, a'\rangle \in f^{-1}$. Since f^{-1} is a function it follows that $a = a'$, and therefore f is injective. □

Exercise 1.36. Suppose that f is a function with $\operatorname{dom}(f) = A$. Let K_f denote the relation $\{\langle a, b\rangle \in A \times A \mid f(a) = f(b)\}$. Show that K_f is an equivalence relation on A. (Visit solution)

Exercise 1.37. Show that if E is an equivalence relation on A, then there is a set B and a function $f\colon A \to B$ such that $E = K_f$. (Visit solution)

Exercise 1.38. Show that $\{f \mid f \text{ is a function}\}$ is a proper class.

Notation 1.36. Let A and B be two sets, then $A^B = \{f \mid f : B \to A\}$.

Exercise 1.39. Show that if A and B are two sets, then A^B is a set.

Definition 1.37. Suppose that f and g are functions, we write $g \circ f$ to denote the composition of the functions given by $g(f(x))$.

Exercise 1.40. Let $f : A \to B$ and $g : C \to D$ be two functions. Show that $g \circ f$ is a function and compute its domain and range.

Exercise 1.41. If $f : A \to B$ and $g : B \to C$ are injective (surjective), then $g \circ f : A \to C$ is injective (surjective).

Theorem 1.38 (Cantor's Theorem). *Let A be any set, then there is no surjective function $f : A \to \mathcal{P}(A)$.*

Proof. Suppose that $f : A \to \mathcal{P}(A)$ and let $A_f = \{a \in A \mid a \notin f(a)\}$, this is a set by the Axiom of Separation. We claim that $A_f \notin \operatorname{rng}(f)$ and therefore f is not surjective. If $f(a) = A_f$ for some a, then $a \in A_f$ if and only if $a \notin f(a) = A_f$, which would be a contradiction, so not such a can exist. □

Exercise 1.42. For any set A, define $f : A \to \mathcal{P}(A)$ by $f(a) = \{a\}$. Show that f is well-defined and injective. (Visit solution)

Exercise 1.43. Calculate A_f for the function from the previous exercise.

Remark

We can now understand what is a mathematical structure as a set with some specified relations and functions. For example, a group is a set G with a function $\cdot : G \times G \to G$ satisfying certain properties. We can understand a field as a set F with two functions, $+ : F \times F \to F$ and $\times : F^\times \times F^\times \to F$, where $F^\times$ is the set of non-zero elements of F (and we can understand the zero element as a distinguished element of F as well).

Chapter 2

Order! Order!

Chapter Goals

In this chapter we will learn about

- Partially and totally ordered sets.
- Strictly ordered sets.
- Minimal and maximal elements, as well as minimum and maximum elements.
- Chains and antichains.
- Order-preserving functions, embeddings, and isomorphisms (of orders).
- Combining orders to produce new ones: pointwise products and lexicographic products.

2.1 Partial and total orders

Definition 2.1. Let R be a relation on a set A.

1. R is *irreflexive (on A)* if for all $a \in A$, $a \not R\, a$.
2. R is *antisymmetric* if whenever $a\, R\, b$ and $b\, R\, a$, then $a = b$.

Exercise 2.1. Show that if R is symmetric and antisymmetric, then $R = \mathrm{id}$ on its domain.

Definition 2.2. We say that a relation R on a set A is a *partial order* if it is reflexive, antisymmetric, and transitive. We say that R is a *strict* partial order if it is irreflexive and transitive. We say that $\langle A, R\rangle$ is a *partially ordered set* if R is a partial order on A.

Exercise 2.2. Suppose that R is a reflexive relation on a set A. Show that R is a partial order if and only if $R \setminus \mathrm{id}$ is a strict partial order.

Remark

We will often (but not always) use $\leq$ to denote a partial order and $<$ to denote its strict counterpart. These by no means hints that we are ordering numbers of any kind or even that the orders are somehow "nice".

Exercise 2.3. Given a set A, show that $\langle \mathcal{P}(A), \subseteq \rangle$ is a partially ordered set.

Definition 2.3. Let $\langle A, \leq \rangle$ be a partially ordered set and let $a, b \in A$. We say that a and b are *comparable* if $a \leq b$ or $b \leq a$ (or $a = b$ in the case of a strict order). Otherwise they are *incomparable*.

Definition 2.4. Let $\langle A, \leq \rangle$ be a partially ordered set.

1. $a \in A$ is a *maximal* element if whenever $a \leq b$, then $a = b$.
2. $a \in A$ is a *maximum* element if for any $b \in A$, $b \leq a$.
3. $a \in A$ is a *minimal* element if whenever $b \leq a$, then $a = b$.
4. $a \in A$ is a *minimum* element if for any $b \in A$, $a \leq b$.

Note that minimum and maximum elements must be comparable with any other element, whereas minimal and maximal elements are not necessary comparable with other elements.

Proposition 2.5. *Let $\langle A, \leq \rangle$ be a partially ordered set. If $a \in A$ is a maximum element, then a is the only maximal element.*

Proof. We first show that a is a maximal element. Let $b \in A$ such that $a \leq b$, we will show that $a = b$. Since a is a maximum element, it follows that $b \leq a$, and by antisymmetry, $a = b$. Suppose now that $b \in A$ is another maximal element, since a is a maximum, $b \leq a$, and since b is maximal, it must be that $a = b$. □

Exercise 2.4. Write the definitions of maximum, maximal, minimal, and minimum in a strict partial order.

Exercise 2.5. Find a partially ordered set without maximal or minimal elements. (Visit solution)

Exercise 2.6. Find a partially ordered set which has a unique maximal element but no maximum.

Definition 2.6. We say that a partially ordered set $\langle A, \leq \rangle$ is a *totally (or linearly) ordered set* if whenever $a, b \in A$ either $a \leq b$ or $b \leq a$.

Definition 2.7. If $\langle A, \leq \rangle$ is a totally ordered set, the *endpoints* of A are the minimum and maximum elements, if they exist.

For example, the standard orderings on $\mathbb{N}$ or $\mathbb{R}$ are total orders, and $\mathbb{R}$ does not have endpoints. Therefore, $\langle \mathbb{N}, \leq \rangle$ is a linearly ordered set.

Exercise 2.7. Find all the sets, A, such that $\langle \mathcal{P}(A), \subseteq \rangle$ is a totally ordered set. Is this collection a set? (Visit solution)

Definition 2.8. Let $\langle A, \leq \rangle$ be a partially ordered set and let $B \subseteq A$.

1. B is a *chain* if any two elements of B are comparable.
2. B is an *antichain* if any two distinct elements of B are incomparable.

We say that B is a *maximal* chain (antichain) if it is maximal with respect to $\subseteq$ in the set of all chains (antichains).

Remark

We can understand the definitions as follows: B is a *chain* if $B \times B \cap \leq$ is a total order on B, and B is an *antichain* if $B \times B \cap \leq = \mathrm{id}$ on B.

Exercise 2.8. Write the definitions for a chain and antichain for strict partial orders.

Proposition 2.9. *Let $\langle A, \leq \rangle$ be a partially ordered set and let $C \subseteq A$ be a chain. The following are equivalent:*

1. *C is a maximal chain.*
2. *Whenever $a \in A \setminus C$, there is some $c \in C$ that is incomparable with a.*

Proof. Suppose that C is a chain and let $a \in A \setminus C$. Then $C \cup \{a\}$ is a chain if and only if a is comparable with all the elements of C, therefore C is maximal if and only if any $a \in A \setminus C$ is incomparable with some $c \in C$. □

Exercise 2.9. Find and prove a condition similar to the above proposition for antichains.

Exercise 2.10. Find all the maximal antichains in $\langle \mathcal{P}(\{0, 1, 2\}), \subseteq \rangle$.

Exercise 2.11. Find a partial order with a minimum element, exactly two maximal chains but no maximal elements. (Visit solution)

Definition 2.10. Let $\langle A, \leq \rangle$ be a partially ordered set and $a \in A$. We say that $b \in A$ is *a successor of a* if $a < b$ and there is no c such that $a < c < b$. We will say that $b \in A$ is *a successor* if it is a successor of some $a \in A$.

Exercise 2.12. Show that in $\mathbb{Z}$ with the standard order, every element is a successor. On the other hand, show that in $\mathbb{Q}$ with the standard order no element is a successor.

Definition 2.11. Suppose that $\langle A, \leq \rangle$ is a linearly ordered set. We say that A is *densely ordered* if whenever $a, b \in A$ are such that $a < b$, then there is some $c \in A$ such that $a < c$ and $c < b$.

Exercise 2.13. Show that $\mathbb{R}$ with its standard order is dense.

Exercise 2.14. A linear order $\langle A, \leq \rangle$ is dense if and only if no element has a successor. (Visit solution)

2.2 Order-preserving functions

Definition 2.12. Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be two partially ordered sets and let $F \colon A \to B$. We say that F is an *embedding (of orders)* if for all $a, a' \in A$

$$a \leq_A a' \text{ if and only if } F(a) \leq_B F(a').$$

We will sometimes write $F \colon \langle A, \leq_A \rangle \to \langle B, \leq_B \rangle$ to mean that F is an embedding of these orders.

Theorem 2.13. *Every partially ordered set embeds into its power set.*

Proof. Let $\langle A, \leq \rangle$ be a partially ordered set and let F be the function $F(x) = \{y \in A \mid y \leq x\}$. Note that by transitivity, $a \leq b$ implies that for any $c \leq a$, $c \leq b$, and therefore $F(a) \subseteq F(b)$. On the other hand, $a \in F(a)$, so if $F(a) \subseteq F(b)$, then $a \in F(b)$ and therefore $a \leq b$, so F is indeed an embedding. □

Proposition 2.14. *If $\langle A, \leq_A\rangle$ and $\langle B, \leq_B\rangle$ are partially ordered sets and $F\colon A \to B$ is an embedding, then F is injective.*

Proof. Suppose that $F(a) = F(a')$, then $F(a) \leq_B F(a')$ and so $a \leq_A a'$, similarly we get that $a' \leq_A a$. Therefore, by antisymmetry, $a = a'$. □

Exercise 2.15. Write a definition for an embedding for strict orders. (Visit solution)

Proposition 2.15. *The composition of embeddings is an embedding.*

Proof. Suppose that $F\colon \langle A, \leq_A\rangle \to \langle B, \leq_B\rangle$ and $G\colon \langle B, \leq_B\rangle \to \langle C, \leq_C\rangle$ are embeddings, then $a \leq_A a'$ if and only if $F(a) \leq_B F(a')$ if and only if $G(F(a)) \leq_C G(F(a'))$. □

Definition 2.16. Let $\langle A, \leq_A\rangle$ and $\langle B, \leq_B\rangle$ be two partially ordered sets. If $F\colon A \to B$ is an embedding such that F is surjective, then we say that F is an *isomorphism* and we say that the two partially ordered sets are *isomorphic*.

Note that if $F\colon \langle A, \leq_A\rangle \to \langle B, \leq_B\rangle$ is an isomorphism, then $F^{-1}\colon \langle B, \leq_B\rangle \to \langle A, \leq_A\rangle$ is an isomorphism.

Exercise 2.16. Show that $\langle \mathcal{P}(A), \subseteq\rangle$ and $\langle \mathcal{P}(A), \supseteq\rangle$ are isomorphic for any set A.

Exercise 2.17. Suppose that $F\colon \langle A, \leq_A\rangle \to \langle B, \leq_B\rangle$ is an embedding. If $C \subseteq A$ is a chain (antichain), $F[C]$ is a chain (antichain) in B. If F is an isomorphism, then minimal, maximal, minimum, maximum, and successor elements preserve their property as such. (Visit solution)

Exercise 2.18. Find two partial orders on $\mathbb{N}$ which are not isomorphic. (Visit solution)

Theorem 2.17. *Suppose that $F\colon \langle \mathcal{P}(A), \subseteq\rangle \to \langle \mathcal{P}(B), \subseteq\rangle$ is an isomorphism. Then there is a bijection $f\colon A \to B$ such that $F(X) = f[X]$ for all $X \subseteq A$.*

Proof. Since F is an isomorphism and $\varnothing$ is the minimum in both power sets, $F(\varnothing) = \varnothing$. It follows that if $a \in A$ is any point, $F(\{a\}) = \{b\}$ for some $b \in B$, since the singletons are exactly the successors of $\varnothing$, moreover, since F is surjective every $\{b\}$ is $F(\{a\})$ for some $a \in A$. We let $f = \{\langle a, b\rangle \in A \times B \mid F(\{a\}) = \{b\}\}$.

Firstly, we claim that f is a function. Namely, if $\langle a, b\rangle$ and $\langle a, c\rangle$ are both in f, then by the definition $F(\{a\}) = \{b\}$ and $F(\{a\}) = \{c\}$, but since F is a function, $b = c$.

Next, by the fact that F is an isomorphism, and in particular injective, f must be injective: if $a \neq a'$, then $\{a\} \neq \{a'\}$, then $\{f(a)\} = F(\{a\}) \neq F(\{a'\}) = \{f(a')\}$, then $f(a) \neq f(a')$. And by the observation that if $b \in B$, then there is some $a \in A$ such that $F(\{a\}) = \{b\}$, we have that f is onto B as well.

Finally, we need to show that if $X \subseteq A$, then $F(X) = f[X]$. We claim that $F(X) = \bigcup\{F(\{x\}) \mid x \in X\}$, and easily $\bigcup\{F(\{x\}) \mid x \in X\} = f[X]$.

Indeed, $\{x\} \subseteq X$ for all $x \in X$, so easily $f[X] = \bigcup\{F(\{x\}) \mid x \in X\} \subseteq F(X)$. In the other direction, if $b \in F(X)$, we have that $\{b\} \subseteq F(X)$, and so there is some a such that $f(a) = b$ and $F(\{a\}) = \{b\}$, and since F is an isomorphism, $\{a\} \subseteq X$, so $a \in X$. Therefore $b \in f[X]$. □

Exercise 2.19. Suppose that $f\colon A \to B$ is an injective function. Show that $F\colon \mathcal{P}(A) \to \mathcal{P}(B)$ given by $F(X) = f[X]$ is an embedding. Moreover, if f is a bijection show that F is an isomorphism.

Exercise 2.20. Find an example of sets A and B and embedding $F\colon \mathcal{P}(A) \to \mathcal{P}(B)$ such that there is no $f\colon A \to B$ for which $F(X) = f[X]$.

Exercise 2.21. Let $\mathrm{PO}(A)$ be the set of all partial orders on A. The relation $E \prec E'$ given by "There exists an embedding $F\colon \langle A, E\rangle \to \langle A, E'\rangle$" is reflexive and transitive. Moreover, show that the relationship $E \cong E'$ given by "There exists an isomorphism $F\colon \langle A, E\rangle \to \langle A, E'\rangle$ is an equivalence relation on $\mathrm{PO}(A)$.

2.3 Products of orders

Given two partial orders we sometimes wish to combine them into one. Let us examine two ways of doing that.

Definition 2.18. Suppose that $\langle A, \leq_A\rangle$ and $\langle B, \leq_B\rangle$ are partial orders. The *pointwise product* is the relation on $A \times B$ given by $\langle a, b\rangle \leq_{\mathrm{pw}} \langle c, d\rangle$ if and only if $a \leq_A c$ and $b \leq_B d$.

Exercise 2.22. Verify that the pointwise product of two partially ordered sets is a partially ordered set.

Exercise 2.23. Suppose that $A \cap B = \varnothing$, then $\langle \mathcal{P}(A \cup B), \subseteq\rangle$ is isomorphic to the pointwise product of $\langle \mathcal{P}(A), \subseteq\rangle$ and $\langle \mathcal{P}(B), \subseteq\rangle$. (Visit solution)

Proposition 2.19. *Suppose that A and B have more than one element each. Then the pointwise product of any total orders $\leq_A$ and $\leq_B$ on A and B respectively is not a total order.*

Proof. Let $a_0 <_A a_1$ and $b_0 <_B b_1$ be two pairs of points in A and B respectively. Then $\langle a_0, b_1\rangle$ and $\langle a_1, b_0\rangle$ are incomparable, so $\leq_{\mathrm{pw}}$ is not a total order. □

This makes the pointwise product a tool that is somewhat too unwieldy for many of the purposes that we want to use products for. Instead, a more convenient definition is the lexicographic product.

Definition 2.20. Suppose that $\langle A, \leq_A\rangle$ and $\langle B, \leq_B\rangle$ are partially ordered sets. The *lexicographic order (or product)* is the relation $\leq_{\mathrm{Lex}}$ on $A \times B$ given by

$$\langle a_0, b_0\rangle \leq_{\mathrm{Lex}} \langle a_1, b_1\rangle \text{ if and only if } a_0 <_A a_1 \text{ or } (a_0 = a_1 \text{ and } b_0 \leq_B b_1).$$

Remark

The intuition of the lexicographic product is that we take a copy of A and we replace each point by a copy of B.

Proposition 2.21. *$\langle A \times B, \leq_{\mathrm{Lex}}\rangle$ is a partially ordered set.*

Proof. We need to verify the three properties. Reflexivity holds since both $\leq_A$ and $\leq_B$ are reflexive, it follows that for any $\langle a, b\rangle \in A \times B$, $a = a$ and $b \leq_B b$, and so $\langle a, b\rangle \leq_{\mathrm{Lex}} \langle a, b\rangle$.

For antisymmetry suppose that $\langle a, b\rangle \leq_{\mathrm{Lex}} \langle c, d\rangle$ and $\langle c, d\rangle \leq_{\mathrm{Lex}} \langle a, b\rangle$. If $a = c$, then we get that $b \leq_B d$ and $d \leq_B b$, so $b = d$ as well. If $a \neq c$, then $a <_A c$ and $c <_A a$ holds which is impossible.

Finally, suppose that $\langle a, b\rangle \leq_{\mathrm{Lex}} \langle c, d\rangle$ and $\langle c, d\rangle \leq_{\mathrm{Lex}} \langle e, f\rangle$. As $a \leq_A c \leq_A e$ and $\leq_A$ is transitive, we have either $a <_A e$ in which case we are done, or else $a = c = e$, in which case $b \leq_B d \leq_B f$ and by the transitivity of $\leq_B$ we are done. □

Exercise 2.24. Show that if $\langle A, \leq_A\rangle$ and $\langle B, \leq_B\rangle$ are linear orders, then their lexicographic product is a linear order. (Visit solution)

Exercise 2.25. Suppose that $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ are non-empty partially ordered sets. Show that both embed into $\langle A \times B, \leq_{\mathrm{pw}} \rangle$ and into $\langle A \times B, \leq_{\mathrm{Lex}} \rangle$.

Exercise 2.26. Show that with their standard orders, $\langle \mathbb{Z} \times \mathbb{Q}, \leq_{\mathrm{Lex}} \rangle$ and $\langle \mathbb{Q} \times \mathbb{Z}, \leq_{\mathrm{Lex}} \rangle$ are not isomorphic. (Visit solution)

Chapter 3

Well, well, well... foundedness

Chapter Goals

In this chapter we will learn about

- Well-foundedness.
- An unusual definition for the concept of "finite" and how to use it.
- Induction on well-founded relations and finite induction.
- Well-ordered sets, along with induction and recursion for that specific concept.
- Using these concepts to prove the Comparability Theorem and Hartogs' Theorem.

3.1 Well-founded relations

Definition 3.1. We say that a relation R on a set A is *well-founded* if for any $B \subseteq A$ which is non-empty, then some $b \in B$ is R-minimal in B. Namely, for all $c \in B$, if $c\,R\,b$, then $c = b$.

The idea is the generalisation of "if something happens, then it happens for a first time" to arbitrary relations. The case where the relation is a linear order is a particularly interesting one, and we will study those later in much closer details.

Still, even in the case where the relation is not linear, it is still useful to be able to pick a minimal "occurrence" of some property, even if it is not the only minimal occurrence.

Exercise 3.1. Suppose that $\leq$ is a well-founded relation on A and let $B \subseteq A$ be a non-empty subset. Show that if B has only one minimal element, then it has a minimum element.

Theorem 3.2. *Suppose that $\langle A, \leq_A\rangle$ and $\langle B, \leq_B\rangle$ are well-founded relations, then both $\leq_{\mathrm{pw}}$ and $\leq_{\mathrm{Lex}}$ are well-founded relations on $A \times B$.*

Proof. The proof in both cases is the same. Suppose that $X \subseteq A \times B$ is non-empty. Let $X_A = \operatorname{dom}(X) = \{a \in A \mid \text{There is some } b, \langle a, b\rangle \in X\}$, then $X_A \neq \varnothing$ and therefore has a $\leq_A$-minimal element, a. Let $X^a = \{b \in B \mid \langle a, b\rangle \in X\}$, then X^a is a non-empty subset of B and so has a $\leq_B$-minimal b. It is now a routine verification of the definitions of $\leq_{\mathrm{pw}}$ and $\leq_{\mathrm{Lex}}$ that $\langle a, b\rangle$ is a minimal element in X. □

Exercise 3.2. Complete the proof for the case of $\leq_{\mathrm{Lex}}$.

Theorem 3.3. *If $F \colon \langle A, \leq_A \rangle \to \langle B, \leq_B \rangle$ is an embedding and B is well-founded, then A is well-founded.*

Proof. Suppose that $X \subseteq A$ is non-empty, then $F[X]$ is a non-empty subset of B, so it has a $\leq_B$-minimal element, say b. Since all the elements of $F[X]$ have the form $F(a)$ for some $a \in X$, we claim that a for which $F(a) = b$ is $\leq_A$-minimal in X.

To see this, suppose that $x \in X$ was such that $x \leq_A a$, then $F(x) \in F[X]$, and since F is an embedding, $F(x) \leq_B F(a) = b$. But since b was $\leq_B$-minimal in $F[X]$, it means that $F(x) = b = F(a)$, so $x = a$ as wanted. □

3.2 Finiteness done "wrong"

Definition 3.4. We say that A is a *finite set* if $\langle \mathcal{P}(A), \subseteq \rangle$ is well-founded. If A is not finite, we say that it is *infinite*.

Remark

This is very far and very different from the standard definition of a finite set, which we will see later. However, this is equivalent to the standard definition nonetheless. The benefit of this definition is that it does not rely on any external understanding or intuition related to the natural numbers.

Theorem 3.5. *If $\langle A, \leq \rangle$ is a partial order and A is finite, then $\leq$ is well-founded.*

Proof. We saw that $\langle A, \leq \rangle$ embeds into $\langle \mathcal{P}(A), \subseteq \rangle$, and therefore it is well-founded. □

Exercise 3.3. If $\langle A, \leq \rangle$ is a finite partial order, then it has a maximal element.

Exercise 3.4. Show that $\mathbb{N}$ is infinite. (Visit solution)

Theorem 3.6. *A subset of a finite set is finite.*

Proof. Suppose that A is finite and $B \subseteq A$. Since $\mathrm{id} \colon B \to A$ is injective, we have an embedding of $\mathcal{P}(B)$ into $\mathcal{P}(A)$, and therefore $\mathcal{P}(B)$ is well-founded. □

Theorem 3.7. *The union of two finite sets is finite.*

Proof. Suppose that A and B are finite, then $A \cup B = A \cup (B \setminus A)$ and $A \cap (B \setminus A) = \varnothing$, so we may assume without loss of generality that $A \cap B = \varnothing$. Therefore $\mathcal{P}(A \cup B)$ is isomorphic to the pointwise product of $\mathcal{P}(A)$ and $\mathcal{P}(B)$ and therefore it is well-founded. □

Definition 3.8. Suppose that $f \colon X \to X$ is a function, we say that $Y \subseteq X$ is *f-closed* if whenever $x \in Y$, $f(x) \in Y$.

Exercise 3.5. Suppose that $f \colon X \to X$ is a function and that $\mathcal{T} \subseteq \mathcal{P}(X)$ is such that for every $T \in \mathcal{T}$, T is f-closed. Then $\bigcap \mathcal{T}$ is f-closed.

Lemma 3.9. *Suppose that $f \colon X \to X$ is a function and $x \in X$, then there is a smallest f-closed set, T, such that $x \in T$. Moreover, if $y \in T$, then either $x = y$ or else there is some $z \in T$ such that $f(z) = y$.*

Proof. Let $\mathcal{T} = \{Y \subseteq X \mid x \in Y \text{ and } Y \text{ is } f\text{-closed}\}$, then $\mathcal{T}$ is non-empty, so $T = \bigcap \mathcal{T}$ is f-closed, and since $x \in Y$ for all $Y \in \mathcal{T}$, it must be that $x \in T$, and therefore T is the smallest f-closed set as wanted. Suppose that $y \in T$ and $y \neq x$, then $T' = T \setminus \{y\}$, is such that $x \in T'$ and $T' \subsetneq T$, and therefore T' is not f-closed. Therefore, there is some $z \in T'$ such that $f(z) \notin T'$. Since $z \in T$, it must be that $f(z) \in T \setminus T' = \{y\}$, as wanted. □

Theorem 3.10. *Suppose that A is a finite set and $f\colon A \to A$ is injective, then f is a bijection.*

Proof. Let us consider the function $F\colon \mathcal{P}(A) \to \mathcal{P}(A)$ given by $F(X) = f[X]$. By the very definition of F, if $X \subseteq Y$, then $F(X) \subseteq F(Y)$, but since f is injective, if $F(X) \subseteq F(Y)$, then it has to be the case that $X \subseteq Y$. To see this, note that if $x \in X$, then $f(x) \in F(X)$ and therefore $f(x) \in F(Y)$. Since f is injective, it must be that $x \in Y$ as well, since x is the unique element of A whose image is $f(x)$, so $X \subseteq Y$. Therefore F is an embedding.

Let $T \subseteq \mathcal{P}(A)$ be the smallest F-closed family such that $A \in T$, as in Theorem 3.9. Moreover, note that $\{B \subseteq A \mid F(B) \subseteq B\}$ is both F-closed and has A as an element, it must be that if $B \in T$, then $F(B) \subseteq B$. Finally, we will argue that $T = \{A\}$, which implies $F(A) = A$, and therefore f was surjective (and thus bijective).

Since A is finite and $T \neq \varnothing$, some $B \in T$ is minimal, since $F(B) \in T$ and $F(B) \subseteq B$, it must be that $B = F(B)$. Consequently, $T \setminus \{B\}$ is F-closed, and so it must be that $A \notin T \setminus \{B\}$, by the minimality of T. And so $T = \{A\}$ as wanted. □

Remark

Richard Dedekind gave a definition for finite sets which is based on the above theorem: "If $f\colon A \to A$ is injective, then f is surjective". This is known as *Dedekind-finiteness* and is equivalent to the standard definition of finiteness under the Axiom of Choice.

Exercise 3.6. If A is a finite set and $f\colon A \to B$ is a surjection, then B is a finite set. (Hint: look at $F(X) = f^{-1}[X] = \{a \in A \mid f(a) \in X\}$.) (Visit solution)

Exercise 3.7. If A is a finite set and $f\colon A \to A$ is a surjection, then f is a bijection.

3.3 Induction and recursion in general

The idea of recursion and induction is very central to mathematics. We define something by recursion or prove something by induction by "iterating" some construction or argument. The idea is generally presented in the context of the natural numbers when we want to prove a property holds for every natural number n we begin by proving it holds for 0, then we show that if it happened to hold for n, then it also holds for $n+1$, and the principle of mathematical induction tells us that it holds for all the natural numbers. The matter of fact is that the only thing needed for recursion and induction is well-foundedness. We will see soon how this relates to the more standard and traditional approach to induction and recursion which you may have seen before.

Theorem 3.11 (General Induction). *Suppose that $\langle A, \leq \rangle$ is a well-founded relation. If $S \subseteq A$ has the property: "For any $a \in A$, if every $b < a$ is in S, then $a \in S$", then $S = A$.*

Proof. Suppose that $S \neq A$, then $A \setminus S$ is non-empty, and therefore it has some $a \notin S$ which is minimal. Therefore, if $b < a$ it must be that $b \notin A \setminus S$, so $b \in S$. But by the property, that means that $a \in S$ in contradiction to how it was selected. □

Exercise 3.8. Suppose that $\langle A, \leq\rangle$ is a relation such that if S has the property above, $S = A$. Show that $\leq$ is well-founded. (Visit solution)

Theorem 3.12. *The finite union of finite sets is finite.*

Proof. Let $\mathcal{S}$ be a finite set whose elements are finite sets. We want to prove that $\bigcup\mathcal{S}$ is finite as well. By our assumption, $\mathcal{P}(\mathcal{S})$ is well-founded, so we can let $\mathcal{T} \subseteq \mathcal{P}(\mathcal{S})$ be the set $\{\mathcal{S}' \subseteq \mathcal{S} \mid \bigcup\mathcal{S}' \text{ is finite}\}$. We want to show that $\mathcal{T}$ has the property "for any $\mathcal{S}' \subseteq \mathcal{S}$, if every proper subset of $\mathcal{S}'$ is in $\mathcal{T}$, then $\mathcal{S}'$ is in $\mathcal{T}$", as by the General Induction Theorem this would mean that $\mathcal{T} = \mathcal{P}(\mathcal{S})$, so $\mathcal{S} \in \mathcal{T}$, and therefore $\bigcup\mathcal{S}$ is finite.

Let $\mathcal{S}' \subseteq \mathcal{S}$ be such that all the proper subsets of $\mathcal{S}'$ are in $\mathcal{T}$, we will show that $\mathcal{S} \in \mathcal{T}$. If $\mathcal{S}' = \varnothing$, then $\bigcup\varnothing = \varnothing$ and therefore is finite, since $\mathcal{P}(\varnothing) = \{\varnothing\}$ is well-founded. So $\varnothing \in \mathcal{T}$, and there is nothing to check. Otherwise, $\mathcal{S}'$ is not empty, let $X \in \mathcal{S}'$ be some element and let $\mathcal{S}'' = \mathcal{S}' \setminus \{X\}$. Since $\mathcal{S}'' \subsetneq \mathcal{S}'$, by the induction hypothesis we have that $\mathcal{S}'' \in \mathcal{T}$, so $Y = \bigcup\mathcal{S}''$ is finite. By our assumption X was a finite set, so by Theorem 3.7, $X \cup Y$ is a finite set. However, $\bigcup\mathcal{S}' = X \cup Y$, so the property is verified, as wanted. □

Theorem 3.13. *Suppose that X is an infinite set and A is a finite set. Then there is an injection $f\colon A \to X$.*

Proof. Let $\mathcal{T} \subseteq \mathcal{P}(A)$ be the set $\{B \in \mathcal{P}(A) \mid \text{There is an injection } B \to X\}$. Suppose that $A' \subseteq A$ is such that any proper subset of A' is an element of $\mathcal{T}$. Note that trivially, $\varnothing \in \mathcal{T}$, as $\varnothing\colon \varnothing \to X$ is an injection. If A' is non-empty, let $a \in A'$ be some element, and consider $B = A' \setminus \{a\}$. By the induction hypothesis, $B \in \mathcal{T}$, so there is some injective function from B into X. Let $f\colon B \to X$ be an injective function. Since X is infinite and B is finite, f is not surjective, and therefore we can pick $x \in X \setminus \operatorname{rng}(f)$ and let $g = f \cup \{\langle a, x\rangle\}$.

We claim that $g\colon A \to X$ is injective. First, note that since $a \notin \operatorname{dom} f$, g is indeed a function. To see that it is indeed injective it is enough to verify that for $a' \neq a$, $g(a') \neq g(a)$, since for any $a' \neq a$, $g(a') = f(a')$ and f was already injective. However, as $x \notin \operatorname{rng}(f)$, if $a' \neq a$, then $a' \in B$, and therefore $g(a') = f(a') \neq x = g(a)$. Therefore, by Theorem 3.11, we have that $\mathcal{T} = \mathcal{P}(A)$, so $A \in \mathcal{T}$, and therefore there is an injective function $A \to X$. □

Remark

The standard proofs you may have seen would be by induction on the number of sets which we unionise, using Theorem 3.7 as the "proof-bearing" theorem. The proof Theorem 3.7 itself may have used induction on the size of one of the sets (or maybe even both). The other theorems in this chapter are likely to have been proved similarly as well.

It is worth noting the difficulty in teaching these proofs. The proofs often require subtle clarity from the student in understanding where the induction hypotheses play various roles, or "why is this not just obvious to begin with". Using a fairly unusual definition we circumvent both the use of induction (in most cases) as well as the lack of clarity as to what we need to prove or where the subtle and non-obvious points lie in the proof.

More generally, if we want to prove something by induction for all finite sets, we can look at the class of all finite sets and note that it is well-founded under $\subseteq$, despite being a proper class. The generalisation of induction works for proper classes, although under an additional condition which holds in this case and we will not discuss. But this provides us with the following theorem.

Theorem 3.14 (Finite Induction). *Suppose that φ is a property that $\varphi(\varnothing)$ holds, and if $\varphi(A)$ holds for a set A and $a \notin A$, then $\varphi(A \cup \{a\})$ holds. Then φ holds for every finite set.*

Proof. Suppose that A is a finite set, then $\mathcal{P}(A)$ is well-founded. Let $\mathcal{T} = \{B \subseteq A \mid \neg\varphi(B)\} = \mathcal{P}(A) \setminus \{B \subseteq A \mid \varphi(B)\}$, then if $\mathcal{T} = \varnothing$, it means that $\varphi(A)$ holds as wanted. Otherwise, let $B \subseteq A$ be a minimal element of $\mathcal{T}$. Since $\varphi(\varnothing)$ holds by assumption, B is not empty, so we may find $a \in B$, and then $B' = B \setminus \{a\}$ is such that $B' \notin \mathcal{T}$, so $\varphi(B')$ holds. Therefore, $\varphi(B' \cup \{a\})$ holds as well, which is to say that $\varphi(B)$ holds. This in contradiction to the fact that $B \in \mathcal{T}$, so $\mathcal{T}$ must be empty and $\varphi(A)$ must hold. □

Exercise 3.9. If A and B are finite, then $A \times B$ is finite. (Hint: represent this as a finite union of finite sets.)

Exercise 3.10. If A is a finite, then $\mathcal{P}(A)$ is finite. (Hint: use Finite Induction.) (Visit solution)

Exercise 3.11. If A and B are finite, then A^B is finite. (Hint: use the previous two exercises.)

Exercise 3.12. Let $B = A \cup \{b\}$ for some $b \notin A$ and let $<_B$ be a strict linear ordering of B. Show that if there is an embedding of linear orders $F\colon A \to \mathbb{Q}$, where A is ordered by $<_B$ and $\mathbb{Q}$ with its standard ordering, then there is an embedding $F'\colon B \to \mathbb{Q}$ such that $F' \restriction A = F$. (Hint: analyse the cases where b is "with respect to A" in the linear ordering $<_B$ and use the density of $\mathbb{Q}$.)

Exercise 3.13. Let $\langle A, <_A\rangle$ be a finite strict linear order. Show that there is an embedding of orders $F\colon A \to \mathbb{Q}$. (Hint: use Finite Induction and the previous exercise.)

Exercise 3.14. Every finite set can be linearly ordered. (Hint: Use finite induction.)

Exercise 3.15. If $\langle A, \leq\rangle$ is a finite partially ordered set, then there is a linear ordering of A, $\preceq$, such that if $a \leq b$, then $a \preceq b$. (Visit solution)

3.4 Some general comments on general recursion

Notation 3.15. For sets A and B let $B^{\subseteq A}$ be the set $\{f \mid \text{For some } A' \subseteq A, f\colon A' \to B\} = \bigcup\{B^{A'} \mid A' \subseteq A\}$.

Theorem 3.16 (General Recursion). *Suppose that $\langle A, \leq\rangle$ is a well-founded relation, B a set, and $G\colon A \times B^{\subseteq A} \to B$. Then there is a unique function $F\colon A \to B$ such that $F(a) = G(a, F \restriction \{b \in A \mid b < a\})$.*

We will not prove this theorem, but it is worth taking the time to try and understand how it relates to the more familiar definition by recursion. We will prove a simpler version for a more restricted later. The idea, however, is that G is a "step function", telling us how to proceed, given all the previous information, and F is the "sequence" defined by recursion which allows us to think of it as "iterated applications". It is hard to understand how this works in abstraction, but it is worth trying to think of the following example of defining the factorial function on the natural numbers.

We can define $n!$ in two ways. Mathematically they are of course equivalent, but if one tries to implement the recursion via programming, the difference becomes notable. The first definition, which is the naive one, of course, by recursion over $\{0, \ldots, n\}$ with its standard order:

$$n! = \begin{cases} 1 & \text{if } n = 0, \\ (n-1)! \cdot n & \text{otherwise.} \end{cases}$$

The second definition requires us to define recursively "interval product", $\mathrm{IP}(n, m)$, which returns the product of all integers in an interval, and then defining $n! = \mathrm{IP}(1, n)$. To define IP, for two natural numbers $n < m$, let $i_{n,m} = \lfloor \frac{n+m}{2} \rfloor$ be their midpoint (rounded down), then

$$\mathrm{IP}(n, m) = \begin{cases} n & \text{if } n = m, \\ \mathrm{IP}(n, i_{n,m}) \cdot \mathrm{IP}(i_{n,m} + 1, m) & \text{otherwise.} \end{cases}$$

This recursive definition is actually a definition on the order of subintervals of some $\{0, \dots, n\}$, which is finite and therefore well-founded.

Many other proofs and definitions that are given by "complete induction" or "induction on the length" are much more naturally presented as induction and recursion over a well-founded order instead. Understanding this helps to provide clarity in these situations.

3.5 Well-orders

Definition 3.17. $\langle A, < \rangle$ is a *well-ordered set* if it is a well-founded strict totally ordered set.

The canonical example is $\mathbb{N}$ with its standard order. But, as we saw, lexicographic products preserve both well-foundedness and linearity, so $\mathbb{N} \times \mathbb{N}$ ordered by $<_{\mathrm{Lex}}$ is also a well-order.

Exercise 3.16. Every finite linear order is a well-order.

Notation 3.18. Suppose that $\langle A, < \rangle$ is a well-ordered set and $B \subseteq A$ is non-empty. Since A is linearly ordered and B has a minimal element, this element is also the minimum of B. So there is no confusion in using the notation $\min B$. Similarly, we will use $\sup B$ (the supremum of B) to denote $\min\{a \in A \mid \text{For every } b \in B, b \leq a\}$ and in case $\sup B \in B$ we write $\max B$ (the maximum of B). Finally, if $a \in A$ and it has a successor, then this successor is unique and we write a' to denote it.

Definition 3.19. Let $\langle A, \leq \rangle$ be a totally ordered set and $a \in A$. We say that $B \subseteq A$ is an *initial segment* of A if whenever $b \in B$ and $a \leq b$, then $a \in B$ as well. We say that B is a *proper* initial segment if $B \neq A$. For $a \in A$ we write $I(a)$ to denote $\{b \in A \mid b < a\}$.

Definition 3.20. Let $\langle A, < \rangle$ be a well-ordered set. We say that $a \in A$ is a *limit point* if a is not $\min A$ and is not a successor point.

Remark

In some of the set theoretic literature a point is a limit point if and only if $\sup I(a) \notin I(a)$ (equivalently, $\sup I(a) < a$), in which case $\min A$ would also be a limit point. In any case, $a = b'$ if and only if $b = \max I(a)$.

Exercise 3.17. If $\langle A, < \rangle$ is well-ordered set then for $a \in A$, either $a = \max A$ or a' exists in A. Similarly, either $a = \min A$, a is a limit point, or a is a successor.

Remark

We did not formally define $\mathbb{N}$, but we have a good argument—based on Peano's axioms for arithmetic—to understand its order as a well-ordering which has no limit points. In the next chapter we will see how to understand that with just sets, but for the time being we can take these basic facts as properties of $\mathbb{N}$ with its standard order.

Exercise 3.18. Show that in $\langle \mathbb{N} \times \mathbb{N}, <_{\text{Lex}} \rangle$ there are infinitely many limit points. Conclude that the order is not isomorphic to $\mathbb{N}$ with the standard order. (Visit solution)

Exercise 3.19. If $\langle A, < \rangle$ is a well-ordered set with a limit point, then A is infinite. (Hint: Consider the reverse order and use Theorem 3.5.)

3.6 Induction and recursion for well-ordered sets

Using the $I(a)$ notation we can rephrase induction for well-ordered sets: If $T \subseteq A$ is such that $I(a) \subseteq T$ implies that $a \in T$, then $T = A$. We can also state the induction theorem for well-ordered set in an alternative way.

Theorem 3.21 (Induction for Well-Orders). *Let $\langle A, < \rangle$ be a well-ordered set and $T \subseteq A$ such that:*

1. $\min A \in T$;
2. *if $a \in T$, then $a' \in T$ if it exists; and*
3. *if a is a limit point and $I(a) \subseteq T$, then $a \in T$.*

Then $A = T$.

This is the generalisation of the "usual" definition of induction with a base case and a step. Although now since we also have to contend with limit points we need to add the third clause.

Proof. Suppose not, then $A \setminus T$ is non-empty, and therefore it has a minimal element, a. Since $\min A \in T$ by definition, $a \neq \min A$. If $a = b'$, then $b \in T$ by the minimality of a, but in that case $a \in T$ as well. Finally, if a is a limit point, then $I(a) \subseteq T$ by the minimality of a, and therefore $a \in T$. In either case we get that $a \in T$ and $a \notin T$, so $A \setminus T$ must be empty, as wanted. □

Definition 3.22. We say that two functions f and g are *compatible* if $f \cup g$ is a function. We say that a set F is a set of pairwise compatible functions if $\bigcup F$ is a function.

Exercise 3.20. Show that f and g are compatible if and only if $f \cap g = f \restriction (\operatorname{dom}(f) \cap \operatorname{dom}(g)) = g \restriction (\operatorname{dom}(f) \cap \operatorname{dom}(g))$.

Exercise 3.21. Show that if f and g are compatible and injective, then $f \cup g$ is an injective function if and only if $\operatorname{rng}(f) \cap \operatorname{rng}(g) = \operatorname{rng}(f \cap g)$.

Exercise 3.22. Show that if F is a set of functions and $\subseteq$ is linearly ordering F, then $\bigcup F$ is a function.

Theorem 3.23 (Definition by Recursion). *Suppose that $\langle A, < \rangle$ is a well-ordered set, B is a set, and $G\colon B^{\subseteq A} \to B$. Then there is a function $F\colon A \to B$ for which $F(a) = G(F \restriction I(a))$.*

Proof. We say that a function f is *good* if for some $a \in A$, $\operatorname{dom}(f) = I(a)$ or $\operatorname{dom} f = A$, and $f(x) = G(f \restriction I(x))$ for all $x \in \operatorname{dom} f$, and we say that $a \in A$ is a good point if there is a good function f such that $\operatorname{dom}(f) = I(a)$.

Let T be the set of good points. By induction we can show that if is a is a good point, there is exactly one function witnessing that it is good. We show by induction that $T = A$. Firstly,

$\min A$ is always good, since $I(\min A) = \varnothing$, and any $f \restriction \varnothing = \varnothing$. So $\min A \in T$. If a is a limit point and $I(a) \subseteq T$, then letting

$$f = \bigcup \{g \mid g \text{ witnesses that some } b < a \text{ is good}\}$$

is a good function witnessing that a is a good point. Finally, if $a \in T$, let f be a witness for that, then $f \cup \{\langle a, G(f)\rangle\}$ is a witness that a' is good. Therefore, by Theorem 3.21, $T = A$, so all points are good points. If $\max A$ exists, then we also need to define $F(a) = G(f)$ where f is the function witnessing that $\max A$ is good; otherwise, we simply take $\bigcup\{f \in A^{\subseteq B} \mid f \text{ is good}\}$. □

In the case of well-orders, we can also have a definition by recursion divided into the three cases. This leads us to the following theorem, which we will not prove.

Theorem 3.24 (Definition by Recursion II). *Suppose that $\langle A, <\rangle$ is a well-ordered set, B is a set, and $G_s \colon B \to B$ and $G_l \colon B^{\subseteq A} \to B$ are functions. Then for every $b \in B$ there is a unique $F_b \colon A \to B$ such that:*

1. *$F_b(\min A) = b$,*
2. *$F_b(a') = G_s(F_b(a))$,*
3. *$F_b(a) = G_l(F_b \restriction I(a))$ for a limit point a.*

We will often abuse the notation and use F itself in the implicit definition of G_s and G_l.

Theorem 3.25. *Suppose that $\langle A, <\rangle$ is a well-ordered set and $F \colon A \to A$ is an embedding. Then for all $a \in A$, $a \leq F(a)$.*

Proof. If this is not the case, let $a = \min\{x \in A \mid F(x) < x\}$. Then, since F is an embedding and $F(a) < a$, we have that $F(F(a)) < F(a) < a$ in contradiction to the minimality of a. □

Exercise 3.23. Conclude that if $\langle A, <\rangle$ is a finite well-ordered set, then every embedding is the identity. Show that there is an embedding $F \colon \mathbb{N} \to \mathbb{N}$ which is not id.

Theorem 3.26. *Let $\langle A, <\rangle$ be a well-ordered set. Then for all $a \in A$, $I(a) \ncong A$.*

Proof. If $F \colon A \to I(a)$ is an embedding, then for all $x \in A$, $F(x) < a$, in particular, $F(a) < a$, in contradiction to Theorem 3.25. □

Theorem 3.27 (Comparison Theorem). *Let $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ be two well-ordered sets. Then exactly one of the three options holds:*

1. *$A \cong B$.*
2. *There is some $a \in A$ such that $I_A(a) \cong B$.*
3. *There is some $b \in B$ such that $A \cong I_B(b)$.*

Proof. We "attempt" to build an embedding $F \colon A \to B$ by recursion: $F(a) = \min B \setminus F[I_A(a)]$.

This might fail for one of the following reasons: B is empty or for some $a \in A$, $F[I_A(a)] = B$. If B is empty, then either $A = B = \varnothing$, in which case (1) holds, or else $B \cong I_A(\min A))$. We will deal with the case where $B \neq \varnothing$ later, but for now let us assume that F was successfully defined for all $a \in A$.

First, let us show that $T = \{a \in A \mid F \upharpoonright I_A(a) \text{ is an embedding}\} = A$. Suppose that $I_A(a) \subseteq T$, we will show that $a \in T$ as well. If a is a limit point, then $I_A(a) = \bigcup\{I_A(x) \mid x < a\}$, so if F was not an embedding on $I_A(a)$, there would be some $x <_A y <_A z <_A a$ such that $F(y) \leq_B F(x)$, but since $x <_A y <_A z$, this means that $F \upharpoonright I_A(z)$ is not an embedding. However, $I_A(a) \subseteq T$, so no such z exists, so no such $x <_A y$ exist, and therefore $a \in T$. If $a = y'$, then $I(a) = I(y) \cup \{y\}$. It is therefore enough to check that $F(x) <_B F(y)$ for all $x <_A y$, but $F(y) = \min B \setminus F[I_A(y)]$, so by definition $F(x) <_B F(y)$, so $a \in T$, and therefore $T = A$.

So, indeed, the defined F is an embedding. If F was surjective, then (1) holds and we are done. Otherwise, let $b = \min B \setminus F[A]$. Then we claim that $F[A] = I_B(b)$. Otherwise, b is not the minimal point not in $F[A]$, and therefore (3) holds.

In the case that F was not defined on A, i.e. the recursion "failed", but $B \neq \varnothing$, let $a \in A$ be the least for which F could not be defined. Then the only way this could have happened is that $F[I_A(a)] = B$, in which case repeating the above argument for F being an embedding, shows that F is an isomorphism between $I_A(a)$ and B, so (2) holds, as wanted.

Finally, to show that exactly one of these cases can hold, note that if two hold at the same time, we would have a well-ordered set which embeds into one of its proper initial segments in contradiction to Theorem 3.26. □

Exercise 3.24. Find explicit functions that define F in the proof of Theorem 3.27 which would work for the two different definition by recursion formats. (Visit solution)

Exercise 3.25. Suppose that $\langle A, <_A\rangle \cong \langle B, <_B\rangle$ are two isomorphic well-ordered sets. Show that there is exactly one isomorphism between them. (Visit solution)

Exercise 3.26. Suppose that $\langle A, <\rangle$ is a well-ordered set and $B \subseteq A$. Show that there exists (a unique) initial segment of A which is isomorphic to B. Find an example where $B \subsetneq A$, but $B \cong A$.

Exercise 3.27. Show that if $\langle A, <\rangle$ is an infinite well-ordered set, then it has an initial segment isomorphic to $\mathbb{N}$. (Visit solution)

Exercise 3.28. A is a finite set if and only if it has a well-ordering $<$ such that its inverse is also a well-ordering. (Hint: In the one direction use Theorem 3.14, in the other, use the previous exercise.)

Theorem 3.28 (Hartogs' Theorem). *Suppose that X is a set, then there is a well-ordered set $\langle A, <\rangle$ such that there is no injection from A into X.*

Proof. Consider $W \subseteq \mathcal{P}(\mathcal{P}(X))$ where $C \in W$ if and only if $\langle C, \subsetneq\rangle$ is a well-ordered set. Let $\equiv$ be the order isomorphism equivalence relation on W. Namely, $C \equiv C'$ if and only if $\langle C, \subsetneq\rangle \cong \langle C', \subsetneq\rangle$. Let $A = W/\equiv$, since all the sets in W are well-ordered, we define an order on A given by $[C]_\equiv \prec [C']_\equiv$ if and only if C embeds into a proper initial segment of C'. By Theorem 3.27 this is a strict total order.

Let us verify that $\langle A, \prec\rangle$ is a well-order. Suppose that $B \subseteq A$ is non-empty, let $[C]_\equiv \in B$ be some equivalence class, consider $\{c \in C \mid [I_C(c)]_\equiv \in B\}$, then either this set has a minimum, c, in which case $[I_C(c)]_\equiv \in B$ and it has to be the minimum there, or else the set is empty in which case $[C]_\equiv$ has no proper initial segment whose equivalence class is in B, in which case $[C]_\equiv$ is itself the minimum of B.

Finally, if $f\colon A \to X$ was an injective function, then $C = \{f[I_A(a)] \mid a \in A\} \cup \{f[A]\}$ would be a chain of subsets such that $\langle C, \subsetneq\rangle \cong \langle A, \prec\rangle$. This would means that $I_A([C]_\equiv) \cong A$, in contradiction to Theorem 3.26. □

Chapter 4

Ordinal numbers

Chapter Goals

In this chapter we will learn about

- The von Neumann ordinals.
- The Burali-Forti Paradox.
- The Axioms of Replacement and Infinity.
- Induction and Recursion on the class of all ordinals.
- The natural numbers can be expressed as sets, as well as the integers, rational, and real numbers.
- The basics of ordinal arithmetic.

4.1 The von Neumann ordinals

4.1.1 Transitive sets

Definition 4.1. We say that a set A is a *transitive set* if for every $a \in A$, $a \subseteq A$.

Note that this definition is equivalent to saying that $A \subseteq \mathcal{P}(A)$.

Remark

This is the first time where we run into a problem if we refuse to accept the paradigm "everything is a set", since if x is not a set, is $\{x\}$ transitive or not? We can weaken the definition to mean that "for every $a \in A$, if a is a set, then $a \subseteq A$". We will not heavily rely on the assumption that everything is a set, however, in some of the exercises that follow, it will be implicit in the question that everything is a set.

Example 4.2. $\varnothing$ is a transitive set, as well as $\{\varnothing\}$. On the other hand, $\{\{\varnothing\}\}$ is not a transitive set, since $\{\varnothing\}$ is not a subset of $\{\{\varnothing\}\}$.

Exercise 4.1. Suppose that A is transitive and $B \subseteq A$, then $A \cup \{B\}$ is transitive.

Exercise 4.2. Suppose that A is transitive, then $\mathcal{P}(A)$ is transitive. (Visit solution)

Notation 4.3. If $\mathcal{F}$ is a set, we denote by $\bigcap \mathcal{F}$ the *intersection over* $\mathcal{F}$ which is the class

$$\{a \mid \text{For every } A \in \mathcal{F}, a \in A\}.$$

Exercise 4.3. If $\mathcal{F}$ is non-empty, then $\bigcap \mathcal{F}$ is a set. What is $\bigcap \varnothing$?

Exercise 4.4. Suppose $\mathcal{F} \neq \varnothing$ and every $A \in \mathcal{F}$ is transitive. Then $\bigcup \mathcal{F}$ and $\bigcap \mathcal{F}$ are transitive.

4.1.2 Neumann Janci dreamed a dream

Definition 4.4. An *ordinal* is a transitive set which is well-ordered by $\in$.

Remark

The goal of the ordinals is to provide us with robust representatives for the well-ordered sets. We will see that every well-ordered set is isomorphic to an ordinal. This definition was given by John von Neumann, and therefore will often be referred to as the "von Neumann ordinal assignment".

Notation 4.5. We will use Greek letters such as α, β, γ and so on to denote ordinals. We will not mention their well-order, as it is always $\in$. And we will generally write $\alpha < \beta$ to mean $\alpha \in \beta$ and $\alpha \leq \beta$ to mean that $\alpha \in \beta$ or $\alpha = \beta$.

Example 4.6. $\varnothing$ is an ordinal. It is certainly a transitive set, since $\varnothing \subseteq \mathcal{P}(\varnothing)$. And it is trivially well-ordered by $\in$. $\{\varnothing\}$ is also an ordinal: it is transitive, since its only member is $\varnothing$, and $\in$ is a well-ordering of $\{\varnothing\}$, since it is a singleton and it is not hard to check that every strict ordering on a singleton is a well-ordering (there is exactly one). Even less trivially, $\{\varnothing, \{\varnothing\}\}$ is an ordinal, as well as $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$.

Example 4.7. The set $\{\{\varnothing\}\}$ is not an ordinal. It is not a transitive set. Its only member is $\{\varnothing\}$, and $\varnothing \neq \{\varnothing\}$. Similarly, $\mathcal{P}(\{\varnothing, \{\varnothing\}\})$ is not an ordinal either: while it is a transitive set, $\varnothing \in \{\varnothing\} \in \{\{\varnothing\}\}$, but $\varnothing \notin \{\{\varnothing\}\}$, so $\in$ is not a well-ordering (another way to see this is that while $\{\{\varnothing\}\} \subseteq \{\varnothing, \{\varnothing\}\}$, it is not true that $\{\{\varnothing\}\} \in \{\varnothing, \{\varnothing\}\}$, nor the other direction of this $\in$ relation holds. So these two are incomparable members.

Remark

We would like to show counterexamples, perhaps, of a set which is linearly ordered by $\in$ but not well-ordered. This requires us to discuss an additional axiom, the Axiom of Foundation, which we will later on. On the surface level of it, both the axiom and its negation are consistent. So we cannot produce examples or counterexamples "by hand", nor we can prove that this never happens from the axioms we have so far.

Exercise 4.5. If α is an ordinal, then $\alpha \notin \alpha$. (Visit solution)

Exercise 4.6. If α is an ordinal, then $\alpha \cup \{\alpha\}$ is an ordinal.

Exercise 4.7. If α is an ordinal and $A \subseteq \alpha$, then A is an ordinal if and only if A is a transitive set.

Exercise 4.8. If α is an ordinal and $\beta < \alpha$, then $I_\alpha(\beta) = \beta$.

Theorem 4.8. *If $\alpha \cong \beta$, then $\alpha = \beta$.*

Proof. Since α is isomorphic to β, we prove by induction that $\alpha \subseteq \beta$. By the symmetry of the isomorphism relation, this will also show that $\beta \subseteq \alpha$, and therefore $\alpha = \beta$.

Let $A \subseteq \alpha$ be the set $\alpha \cap \beta$. We will show that if $\gamma \subseteq A$, then $\gamma \in A$ for all $\gamma < \alpha$. Suppose that $\gamma < \alpha$ and $\gamma \subseteq A$, then $\gamma \subseteq \beta$. Let $\delta \in \beta$ such that $I_\alpha(\gamma) = \mathrm{I}_\beta(\delta)$, i.e. the image of γ under the isomorphism between α and β. But by a previous exercise, $\gamma = I_\alpha(\gamma)$ and $I_\beta(\delta) = \delta$, so $\gamma = \delta$, and therefore $\gamma \in \beta$, so $\gamma \in A$. □

Exercise 4.9. If α and β are ordinals, then $\alpha \leq \beta$ or $\beta \leq \alpha$. (Hint: Use Theorem 3.27 and the previous theorem.)

Exercise 4.10. A is an ordinal if and only if it is a transitive set of ordinals. (Visit solution)

Exercise 4.11. If α and β are ordinals, then $\alpha \leq \beta$ if and only if $\alpha \subseteq \beta$.

Theorem 4.9 (Burali-Forti Paradox). *The class of all ordinals is a proper class.*

Proof. Note that if α is an ordinal, then itself is a set of ordinals. Therefore if A is the class of all ordinals and $\alpha \in A$, then $\alpha \subseteq A$. Therefore, if A is a set, it is a transitive set. Moreover, A is well-ordered by $\in$, therefore A is an ordinal, so $A \in A$, in contradiction to the fact that an ordinal cannot be an element of itself. □

Notation 4.10. We use Ord to denote the class of all ordinals. This class, despite being a proper class, is well-ordered by $\in$ as well, and in a sense it is a "universal well-order". So when we talk about sets of ordinals, $\sup A$ is always taken in the context of the class of ordinals. Similarly, when we write α' to denote the successor of α, we will mean that in the class of ordinals as well.

Theorem 4.11. *If A is a set of ordinals, then $\bigcup A$ is an ordinal, and moreover $\sup A = \bigcup A$. If $A \neq \varnothing$, then $\min A = \bigcap A$.*

Proof. Since A is a set of ordinals, it is a set of transitive sets, so $\bigcup A$ is a transitive set. Moreover, since an ordinal is a set of ordinals, $\bigcup A$ is a transitive set of ordinals, so it is an ordinal. To see that it is $\sup A$, note that if $\alpha \in A$, then $\alpha \subseteq \bigcup A$ and therefore either $\alpha = \bigcup A$ or $\alpha \in \bigcup A$, so $\bigcup A$ is the least ordinal, β, such that for every $\alpha \in A$, $\alpha \leq \beta$ which is the definition of supremum. Finally, if $A \neq \varnothing$ then $\bigcap A$ is also a transitive set of ordinals, and therefore an ordinal. Moreover, if $\alpha \in A$ then $\min A \subseteq \alpha$, so it is indeed $\min A$. □

Axiom: Replacement

Suppose that $\varphi(x, y)$ is a functional property. Namely, for every x there is exactly one y such that $\varphi(x, y)$ holds. Then for every set A, $\{b \mid \text{There is some } a \in A, \varphi(a, b)\}$ is a set.

Remark

The idea behind the Axiom of Replacement is that we can use functional properties as though they were actually functions. But another way of looking at it is by noting that if φ is a functional property and A is a set, then $\{\langle a, b\rangle \mid a \in A \text{ and } \varphi(a, b)\}$ is a set. Namely, functional properties define actual functions when the domains are sets.

Exercise 4.12. Assuming the Axiom of Replacement, prove the Axiom of Pairing is redundant.

Theorem 4.12. *Every well-ordered set is isomorphic to a unique ordinal.*

Proof. Let $\langle A, < \rangle$ be a well-ordered set and let $S = \{a \in A \mid I(a) \text{ is isomorphic to an ordinal}\}$. Clearly, $\min A \in S$, since $\varnothing \cong I(\min A)$. If $a \in S$ and $I(a) \cong \alpha$, then $I(a') = I(a) \cup \{a\} \cong \alpha' = \alpha \cup \{\alpha\}$ by simply taking $f\colon I(a) \to \alpha$ and considering $f \cup \{\langle a, \alpha \rangle\}$.

Finally, if a is a limit point of A and $I(a) \subseteq S$, consider $\varphi(x, \alpha)$ to mean "$I(x) \cong \alpha$". By the Axiom of Replacement, $X = \{\alpha \mid \text{There is some } x \in I(a), I(x) \cong \alpha\}$ is a set, and it is not hard to check that X is a transitive set of ordinals and that $X \cong I(a)$, so $a \in S$ as well. By Theorem 3.21 $S = A$.

So, considering now φ as before, applying Replacement to A itself, we get a set of ordinals which is transitive and isomorphic to A itself. □

Notation 4.13. Given a well-ordered set $\langle A, < \rangle$ we write $\operatorname{otp}(A, <)$ or $\operatorname{otp}(A)$ to denote the unique ordinal, α, isomorphic to it, and we say that α is the *order type* of A.

Definition 4.14. We say that α is a *successor ordinal* if $\alpha = \beta' = \beta \cup \{\beta\}$ for some ordinal β. If α is not $\varnothing$ or a successor, we say that it is a *limit ordinal*.

Exercise 4.13. Show that α is a limit ordinal if and only if $\alpha \neq \varnothing$ and $\bigcup \alpha = \alpha$.

Theorem 4.15 (Definition by Recursion on Ord). *Suppose that $\varphi(x, y)$ is a functional property, then there is a functional property, $\psi(x, y)$ such that whenever α is an ordinal $\psi(\alpha, y)$ holds if and only if $\varphi(\{\langle \beta, b \rangle \mid \psi(\beta, b) \text{ and } \beta < \alpha\}, y)$.*

In other words, we can define by recursion (and in fact, prove by induction) on the proper class of the ordinals. We will not prove this theorem.

Theorem 4.16 (Induction on Ord). *Suppose that $\varphi(x)$ is a property such that if every $\beta < \alpha$ satisfies φ, then $\varphi(\alpha)$ holds as well. Then φ holds for all the ordinals.*

Remark

Similarly to the statement of Theorems 3.21 and 3.24, we can also phrase induction and recursion over Ord by separating the cases of successor and limit ordinals.
So we can treat the proper class of Ord in very similar ways to how we treat any other well-ordered set despite it not being a set. Of course, we do need to exercise caution insofar that now our functions or collections that we use for the induction theorem(s) are not sets, but rather proper classes, and therefore given by properties instead of "objects".

Theorem 4.17. *Let $\langle A, < \rangle$ be a partial order. Then A is well-founded if and only if there exists a function $f\colon A \to \alpha$, for some ordinal α, such that whenever $a < b$, $f(a) < f(b)$.*

Proof. Suppose that $f\colon A \to \alpha$ exists and let $B \subseteq A$ be a non-empty set. Let $R = f[B]$, then R is a non-empty set of ordinals and therefore has a least member, β. Therefore, there is some $b \in B$ such that $f(b) = \beta$. If b is not minimal in B, then there is some $a \in B$ such that $a < b$. By the property of f, $f(a) < f(b) = \beta$. However, $\beta = \min R$, and therefore such a cannot exist, so b is minimal in B, and therefore A is well-founded.

In the other direction, we assume that A is well-founded let α be an ordinal such that α does not inject into A. Such ordinal exists as the order type of a well-ordering obtained through Theorem 3.28. Let $G\colon A \times \alpha^{\subseteq A} \to \alpha$ given by

$$G(a, f) = \sup\{f(x)' \mid x \in \operatorname{dom} f\}.$$

By Theorem 3.16 there exists a function $f\colon A \to \alpha$ such that $f(a) = G(a, f \restriction \{b \in A \mid b < a\})$, which is given by $f(a) = \sup\{f(b)' \mid b < a\}$. It is easy to see that f has the wanted property. If $a < b$, then $f(a) < f(a)' \leq f(b)$, as wanted. □

Definition 4.18. If $\langle A, <\rangle$ is a well-founded order, the *rank function* is the function defined recursion by $\mathrm{rank}_A(a) = \sup\{\mathrm{rank}_A(b) + 1 \mid b < a\}$. We say that the *(well-founded) rank of A* is α if $\alpha = \sup\{\mathrm{rank}_A(a) + 1 \mid a \in A\}$.

Exercise 4.14. If $\langle A, <\rangle$ is a well-founded partial order, then the range of its rank function is an ordinal.

Exercise 4.15. If α is an ordinal, then its rank is α.

4.2 The Natural Numbers

So far we have discussed $\mathbb{N}$ as being a set. But with the axioms we have so far, there is no means for us to prove that $\mathbb{N}$ is a set. Moreover, we have claimed that set theory is a good foundation for mathematics, so it should be able to interpret $\mathbb{N}$ by using sets.

Definition 4.19. A is an *inductive set* when $\varnothing \in A$ and if $x \in A$, then $x \cup \{x\} \in A$.

Axiom: Infinity

There exists an inductive set.

Remark

Now we see the futility of the Axiom of the Empty Set, if we are assuming the Axiom of Infinity. Firstly, by the existence of any set, the Axiom of Separation will provide us with the empty set by applying the formula $\varphi(x)$ given by $x \neq x$. Secondly, in a very explicit way, the Axiom of Infinity states that the empty set exists, being an element of an inductive set.

Exercise 4.16. If $\mathcal{F} \neq \varnothing$ and every $X \in \mathcal{F}$ is inductive, then $\bigcap \mathcal{F}$ is inductive. (Visit solution)

Definition 4.20. Let A be an inductive set and let $\omega = \bigcap\{B \subseteq A \mid B \text{ is inductive}\}$.

Theorem 4.21. *ω is well-defined and it is the minimum inductive set.*

Proof. To see that ω is well-defined we need to show that the definition does not depend on A. Indeed, let ω_A be the set given by taking A as our starting inductive set. If A and B are two inductive sets, then $\omega_A \cap \omega_B$ is an inductive set and $\omega_A \cap \omega_B \subseteq A$, so $\omega_A \subseteq \omega_A \cap \omega_B$, so $\omega_A \subseteq \omega_B$, and vice versa. It follows now that if A is any inductive set, then $\omega = \omega_A \subseteq A$, so it is the minimum. □

Theorem 4.22. *ω is an ordinal.*

Proof. Suppose that I is an inductive set, we claim that $I \cap \mathrm{Ord}$ is an inductive set as well. It is a set by Axiom of Separation, so it is enough to show that it is inductive. Since $\varnothing$ is an ordinal, $\varnothing \in I \cap \mathrm{Ord}$. Suppose that $\gamma \in I \cap \mathrm{Ord}$, then $\gamma \in I$, and since $\gamma \cup \{\gamma\} \in I$ as well, and $\gamma \cup \{\gamma\} = \gamma'$ is also an ordinal, it follows that $\gamma' \in I \cap \mathrm{Ord}$. Finally, let δ be the least ordinal which is not in I, then δ must be a limit ordinal by the above. And by definition an ordinal δ is an inductive set if and only if it is a limit ordinal. Therefore, $\delta = I \cap \mathrm{Ord}$ is a limit ordinal and $\omega \subseteq \delta$. Let $A \subseteq \delta$ be the set of all finite ordinals, as by the fact that δ is a limit ordinal it must be greater than all the finite ordinals.

We claim that A is an inductive set. Indeed, $\varnothing$ is an ordinal and it is a finite set, so $\varnothing \in A$, and if $\alpha \in A$, then $\alpha \cup \{\alpha\}$ is also a finite ordinal, so $\alpha' \in A$.

Therefore, $\omega \subseteq A$. In fact, $\omega = A$. To see that, note that $\varnothing \in \omega$, and if $\alpha \in \omega$, then by the fact that ω is an inductive set, $\alpha \cup \{\alpha\} = \alpha' \in \omega$. There is no need to check the case for a limit ordinal in A, since we know that a limit ordinal is infinite so no limit ordinals exist in A, and that case holds vacuously. So by Theorem 3.21 $\omega = A$. □

Exercise 4.17. ω is the smallest limit ordinal. (Visit solution)

Theorem 4.23. $\omega = \mathrm{otp}(\mathbb{N})$.

Proof. To see that this is the case, note that $\mathbb{N}$ does not have a maximal element, and every $n \in \mathbb{N}$ is either 0 or of the form $m+1$. Similarly, any ordinal in ω is finite, empty or of the form α' for some finite ordinal α, and ω being a limit ordinal does not have a maximum. So, we can define by recursion the isomorphism: $F(0) = \varnothing$ and $F(n+1) = F(n) \cup \{F(n)\} = F(n)'$. □

Now that we have seen that ω is order isomorphic to $\mathbb{N}$, we can start treating its elements as though they are (and always have been) the natural numbers. This means that $0 = \varnothing$ and $n+1 = n \cup \{n\}$, and so there is no more ambiguity as to which sets we mean when by 0 or 42.

Theorem 4.24. *Suppose that A is a set, then there is a transitive set B such that $A \subseteq B$.*

Proof. We define by recursion on ω a function, $F(0) = A$, $F(n+1) = \bigcup F(n)$.

Let $B = \bigcup F[\omega] = \bigcup\{F(n) \mid n < \omega\}$. Since $A = F(0) \in F[\omega]$, we have that $A \subseteq B$. To verify that B is a transitive set, suppose that $b \in B$, then there is some $n < \omega$ such that $b \in F(n)$, therefore $b \subseteq F(n+1) \subseteq B$ as wanted. □

Exercise 4.18. Show that B defined in the proof above is the smallest transitive set which contains A. In other words, if C is a transitive set and $A \subseteq C$, then $B \subseteq C$.

Notation 4.25. We write $\mathrm{tcl}(A)$ to denote the *transitive closure of* A, which is the smallest transitive set such that $A \subseteq \mathrm{tcl}(A)$.

Definition 4.26. Suppose that α and β are ordinals.

1. $\alpha + \beta$ is the order type of the concatenated well-order. Namely, $(\{0\} \times \alpha) \cup (\{1\} \times \beta)$.
2. $\alpha \cdot \beta$ is the order type of $\langle \beta \times \alpha, <_{\mathrm{Lex}} \rangle$.

Remark

The order on $\alpha + \beta$ is also lexicographic. We can treat it as a subset of $2 \times (\alpha \cup \beta)$ with the lexicographic order. Both of these operations can extend to longer sums and products, as well as to exponentiation whose definition is more complicated when given directly.

Theorem 4.27. *Fixing an ordinal α, addition and multiplication are equivalent to the following recursive definition.*

1. *$\alpha + 0 = \alpha$; $\alpha + (\beta') = (\alpha + \beta)'$; and $\alpha + \beta = \sup\{\alpha + \gamma \mid \gamma < \beta\}$ for a limit ordinal β.*
2. *$\alpha \cdot 0 = 0$; $\alpha \cdot (\beta') = (\alpha \cdot \beta) + \alpha$; and $\alpha \cdot \beta = \sup\{\alpha \cdot \gamma \mid \gamma < \beta\}$ for a limit ordinal β.*

We will not prove this theorem here, but the proof is done by fixing α and proving by induction on β.

Exercise 4.19. Show that $\alpha' = \alpha + 1$.

Remark

From this point onward, we will use $\alpha + 1$ to denote the successor of an ordinal α, and the $'$ notation will not be treated as special.

Exercise 4.20. Find α and β such that $\alpha + \beta \neq \beta + \alpha$. (Visit solution)

Exercise 4.21. An ordinal δ is a limit ordinal if and only if for all $\alpha < \delta$, $\alpha + 1 < \delta$. (Visit solution)

Exercise 4.22. If α is an infinite ordinal, then there is some $n < \omega$ and a limit ordinal δ such that $\alpha = \delta + n$.

Exercise 4.23. Show that $2 \cdot \omega = \omega < \omega + \omega = \omega \cdot 2$.

Exercise 4.24. Find a well-ordering of $\mathbb{N}$ whose order type is $\omega + \omega$. (Visit solution)

Remark

Without the Axiom of Replacement it is impossible to prove that the von Neumann ordinal $\omega + \omega$ exists, despite the fact we can prove Hartogs' theorem, which tells us that there are well-ordered sets which are much longer than it.

Finally, it is not hard to verify from the inductive definition of the arithmetic operations that the following theorem holds, giving even more credence to the claim that ω is truly a good and valid interpretation of $\mathbb{N}$ by sets.

Theorem 4.28. *Restricted to ω, $+$ and $\cdot$ are the standard arithmetic operations.* □

4.3 Rational approach to the real numbers that is too complex

We have seen that $\mathbb{N}$ can be represented by sets with its order and arithmetic. What about $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$? All of these can be, of course, interpreted as sets once we have established $\mathbb{N}$. These constructions are not particularly set theoretic either, but instead algebraic in nature.

Theorem 4.29. *Define the equivalence relation on $\mathbb{N} \times \mathbb{N}$ as follows: $\langle n, m\rangle \sim \langle n', m'\rangle$ if and only if $n + m' = n' + m$. Let $Z = \mathbb{N} \times \mathbb{N}/{\sim}$ and define*

1. $[\langle n, m\rangle]_\sim + [\langle n', m'\rangle]_\sim = [\langle n + n', m + m'\rangle]_\sim$,
2. $[\langle n, m\rangle]_\sim \cdot [\langle n', m'\rangle]_\sim = [\langle nn' + mm', nm' + n'm\rangle]_\sim$, *and*
3. $[\langle n, m\rangle]_\sim < [\langle n', m'\rangle]_\sim$ *if and only if* $n + m' < n' + m$.

Then Z with $+$, $\cdot$, and $<$ is an interpretation of $\mathbb{Z}$ with its standard operations. Moreover, the function $f\colon \mathbb{N} \to Z$ given by $f(n) = [\langle n, 0\rangle]_\sim$ is an order embedding such that $f(n + m) = f(n) + f(m)$ and $f(n \cdot m) = f(n) \cdot f(m)$.

The idea, of course, is that $\langle n, m\rangle$ represents $n - m$ and so we can introduce the additive inverses. So -4 would be the equivalence class of $\langle 0, 4\rangle$ as well as $\langle 1, 5\rangle$, etc. Now that we have obtained $\mathbb{Z}$, we get the rational numbers by introducing multiplicative inverses.

Theorem 4.30. *Define the equivalence relation on* $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ *as follows:* $\langle n, m\rangle \sim \langle n', m'\rangle$ *if and only if* $nm' = n'm$*. Let* $Q = \mathbb{Z} \times \mathbb{Z}/\sim$ *and define*

1. $[\langle n, m\rangle]_\sim + [\langle n', m'\rangle]_\sim = [\langle nm' + n'm, mm'\rangle]_\sim$,
2. $[\langle n, m\rangle]_\sim \cdot [\langle n', m'\rangle]_\sim = [\langle nn', mm'\rangle]_\sim$, *and*
3. $[\langle n, m\rangle]_\sim < [\langle n', m'\rangle]_\sim$ *if and only if* $nm' < n'm$.

Then Q *with* $+$, $\cdot$, *and* $<$ *is an interpretation of* $\mathbb{Q}$ *with its standard operations. Moreover, the function* $f\colon \mathbb{Z} \to Q$ *given by* $f(n) = [\langle n, 1\rangle]_\sim$ *is an order embedding such that* $f(n + m) = f(n) + f(m)$ *and* $f(n \cdot m) = f(n) \cdot f(m)$.

Arriving at $\mathbb{Q}$, we can now define the real numbers as well.

Theorem 4.31. *Define* R *as the set of non-empty, proper initial segment, without a maximal element of* $\mathbb{Q}$ *and define*

1. $D + D' = \{q + q' \mid q \in D \text{ and } q' \in D'\}$,
2. $-D = \{q - q' \mid q < 0 \text{ and } q' \in \mathbb{Q} \setminus D\}$ *is the additive inverse.*
3. $D < D'$ *if and only if* $D \subsetneq D'$.
4. $D \cdot D'$ *requires a breakdown into cases.*
   (a) *If* $D, D' \geq 0$*, then* $D \cdot D' = \{q \cdot q' \mid q \in D, q' \in D', \text{ and } q, q' \geq 0\} \cup \{q \mid q < 0\}$.
   (b) *Otherwise, using* $x \cdot y = -((-x) \cdot y) = -(x \cdot (-y)) = (-x) \cdot (-y)$ *we can replace* D *or* D'*, if necessary, and reduce to the previous case.*

Then R *with* $+$, $\cdot$, *and* $<$ *is an interpretation of* $\mathbb{R}$ *with its standard operations. Moreover, the function* $f\colon \mathbb{Q} \to R$ *given by* $f(q) = \{q' \in \mathbb{Q} \mid q' < q\}$ *is an order embedding such that* $f(q + q') = f(q) + f(q')$ *and* $f(q \cdot q') = f(q) \cdot f(q')$.

Remark

We may notice that these interpretations give us the situation that $\mathbb{N} \not\subseteq \mathbb{Z} \not\subseteq \mathbb{Q} \not\subseteq \mathbb{R}$, but instead there is an embedding of each into the following, and indeed this embedding is unique (once we require the arithmetic to be preserved). However, this is just one way of producing these interpretations, and we may very well do it in other ways, or even in a different order (e.g., first define the non-negative rationals and then $\mathbb{Q}$ rather than going through $\mathbb{Z}$).

Exercise 4.25. Define the complex numbers, $\mathbb{C}$, using $\mathbb{R}$.

Chapter 5

Cardinal numbers

Chapter Goals

In this chapter we will learn about

- The concept of cardinality, and in what sense do infinite sets have the same "number of elements".
- The basics of cardinal arithmetic.
- Countable sets.
- Cantor–Bernstein Theorem and the ordering on cardinals.
- Initial ordinals and the $\aleph$ numbers.

5.1 What do you mean "how many elements?"

5.1.1 Cardinality

Our goal in this chapter is to understand cardinal numbers as the concept of "how large is a set" as well the concepts and theorems around it. Walking into a football stadium, how can you tell if there are more seats than people in one of the sections? Easy, ask everyone to sit down, and see if there are any free seats left, if there are people standing, then there are more people, and if there are no free seats and no one is standing, then there are the same number of both.

In other words, see if you can map the set of people into the set of seats and if so, is this map surjective. This translates to the idea that is that if we can map two sets bijectively, then they must have the same size. To that end, if there is an injection $A \to B$, then it is a bijection of A with a subset of B. Certainly increasing the set will not decrease its size, so that means that injections and bijections give us a good measurement to compare sets.

Notation 5.1. If A and B are sets, we write $A \precsim B$ to mean "there exists an injective function $f\colon A \to B$". We write $A \sim B$ to mean "there exists a bijective function $f\colon A \to B$" in this case we say that A and B are *equipotent* (or *equipollent*) or *"have the same cardinality"*.

Exercise 5.1. The relation $A \sim B$ is an equivalence relation, and $A \precsim B$ is reflexive and transitive.

Exercise 5.2. If $A \precsim C$ and $B \precsim D$, then $A \times B \precsim C \times D$. Find an example showing that the converse might not be true. That is, $A \times B \precsim C \times D$ but either $A \not\precsim C$ or $B \not\precsim D$.

Exercise 5.3. If $A \cap B = C \cap D = \varnothing$, $A \precsim C$, and $B \precsim D$, then $A \cup B \precsim C \cup D$.

Exercise 5.4. If $A \precsim C$ and $B \precsim D$ are all non-empty, then $A^B \precsim C^D$. Show that if $A \sim C$ and $B \sim D$, then $A^B \sim C^D$. Explore the edge cases where some (or all) the sets are empty.

Exercise 5.5. $(A^B)^C \sim A^{B \times C}$ and $A^B \times A^C \sim A^{B \cup C}$ when $B \cap C = \varnothing$.

Exercise 5.6. $\mathcal{P}(A) \sim 2^A$. (Hint: For every $B \subseteq A$ consider the indicator function $\chi_B \colon A \to 2$ defined by $\chi_B(a) = 1$ if and only if $a \in B$.) (Visit solution)

Exercise 5.7. $A \times A \sim A^2$.

Proposition 5.2. *Suppose that A and B are sets, then there is some B' such that $A \cap B' = \varnothing$ and $B \sim B'$.*

Proof. Let $T = \{\langle S, a \rangle \in A \mid S \subseteq A, a \in \operatorname{rng}(A), \langle S, a \rangle \notin S\}$, and let $B' = \{T\} \times B$. The function $f(b) = \langle T, b \rangle$ is a bijection witnessing that $B \sim B'$, and we claim that $B' \cap A = \varnothing$. To see that, suppose that $\langle T, b \rangle \in B' \cap A$, then if $\langle T, b \rangle \in T$ then by definition $b \in \operatorname{rng}(A)$ so $\langle T, b \rangle \in T$ if and only if $\langle T, b \rangle \notin T$ which is impossible. □

5.1.2 Countable sets

Definition 5.3. We say that A is *countable* if $A \precsim \omega$, if A is also infinite, we say that it is *countably infinite*. If A is not countable, we say that it is *uncountable*.

Proposition 5.4. *If A is finite, then for some $n < \omega$, $A \sim n$, and in particular A is countable.*

Proof. We saw that every finite set can be well-ordered. So if A is a finite set, there is some order $<$ such that $\langle A, < \rangle$ is well-ordered, and therefore it is isomorphic to a finite ordinal n, and since ω is the least infinite ordinal, $n < \omega$ so A is countable. □

Remark

This is the standard definition of finiteness: there is some $n < \omega$ such that $A \sim n$.

Theorem 5.5. *Suppose that $A \subseteq \omega$ is infinite, then $A \sim \omega$. In particular, if X is a countable set, then X is finite or $X \sim \omega$.*

Proof. If $A \subseteq \omega$, then A is well-ordered as a subset of ω, and therefore it compares to ω as a well-ordered set. Since any proper initial segment of ω is a finite ordinal and A is infinite, A cannot be isomorphic to a proper initial segment of ω. On the other hand, since A embeds into ω, it is impossible that ω is isomorphic to a proper initial segment of A either. Therefore $\operatorname{otp}(A) = \omega$, and therefore $A \sim \omega$.

If X is a countable set, then let $f \colon X \to \omega$ be an injection and let $A = \operatorname{rng}(f)$, then $X \sim A$. So either X is finite, or else A is an infinite subset of ω and therefore $X \sim A \sim \omega$. □

Remark

In many places we can see the use of the term *denumerable* to mean countably infinite, or countable will be used for the infinite case whereas denumerable includes finite, or that countable means countably infinite and "at most countable" means "finite or countable".

Theorem 5.6. *Suppose that A and B are countable, then $A \cup B$ is countable.*

Proof. We may assume that $A \cap B = \varnothing$, since $A \cup B = (A \setminus B) \cup B$. Since A and B are both countable, let $f\colon A \to \omega$ and $g\colon B \to \omega$ be two injections. Define $h\colon A \cup B \to \omega$ by

$$h(x) = \begin{cases} 2f(x) & \text{if } x \in A, \\ 2g(x) + 1 & \text{if } x \in B \end{cases}$$

It is not hard to check that this function is indeed an injection from $A \cup B$ into ω as wanted. □

Theorem 5.7. *Suppose that A and B are countable, then $A \times B$ is countable.*

Proof. We will show that $\omega \times \omega$ is countable, since $A \times B \precsim \omega \times \omega$. Let $F(n, m) = 2^n(2m+1) - 1$, we claim that $F\colon \omega \times \omega \to \omega$ is a bijection.

F is injective: if $F(n, m) = F(i, j)$, then $2^n(2m + 1) - 1 = 2^i(2j + 1) - 1$, and therefore $2^n(2m + 1) = 2^i(2j + 1) = k$. Since $2m + 1$ and $2j + 1$ are odd, k is divisible by 2^n and 2^i, but by the uniqueness of the prime decomposition of k, $n = i$. Therefore $2m + 1 = 2j + 1$, and so $m = j$ as wanted.

F is surjective: given any $k < \omega$ let n be the maximal power of 2 which divides $k + 1$, then we write $k + 1 = 2^n k'$ where k' is a positive odd number, and therefore $k' = 2m + 1$ for some $m < \omega$. In other words, $F(n, m) = k$, so F is a bijection as wanted. □

Exercise 5.8. Show that for any $n < \omega$, if $A_1, \dots, A_n$ are countable sets, then $A_1 \cup \dots \cup A_n$ and $A_1 \times \dots \times A_n$ are countable.

Exercise 5.9. Show that $C(n, m) = \frac{(n+m)(n+m+1)}{2} + m$ is also a bijection $\omega \times \omega \to \omega$.

Theorem 5.8. *Suppose that $A \neq \varnothing$. The following are equivalent:*

1. *There exists an injection $f\colon A \to \omega$.*
2. *There exists a surjection $g\colon \omega \to A$.*

Proof. Suppose that there is an injection $f\colon A \to \omega$, as A is non-empty, fix some $a_0 \in A$. We define $g\colon \omega \to A$ as follows:

$$g(n) = \begin{cases} a & \text{if } f(a) = n \\ a_0 & \text{if } n \notin \operatorname{rng}(f) \end{cases}$$

It is not hard to see that $f^{-1} \subseteq g$, so g is surjective.

Suppose that $g\colon \omega \to A$ is surjective, we let $f(a) = \min\{n < \omega \mid g(n) = a\}$. Then f is injective, since if $f(a) = f(b) = n$, then $g(n) = a$ and $g(n) = b$, but g is a function, so $a = b$. □

Notation 5.9. We write $A \precsim^* B$ if either $A = \varnothing$ or there is a surjective function $f\colon B \to A$.

Using the new notation, we write rephrase the last theorem as "$A \precsim \omega$ if and only if $A \precsim^* \omega$".

Exercise 5.10. Show that $\mathbb{Z}$ and $\mathbb{Q}$ are countable. (Visit solution)

Exercise 5.11. Suppose that $\langle A, \prec\rangle$ is a linear order and A is countable. Show that A embeds into $\mathbb{Q}$. (Hint: Use Exercise 3.10.) (Visit solution)

Theorem 5.10. *The set $\operatorname{fin}(\omega) = \{A \subseteq \omega \mid A$ is finite$\}$ is countable.*

Proof. For every $n < \omega$ let $\mathcal{P}(n)$ be equipped with the lexicographic order: $A <_{\text{Lex}} B$ if and only if $\min(A \triangle B) \in B$. It is immediate that this is a strict totally ordered set, so it is well-ordered, since $\mathcal{P}(n)$ is finite. This allows us to canonically enumerate the elements of $\mathcal{P}(n)$ by the integers below 2^n. Therefore, we can map $\omega \times \omega$ onto $\text{fin}(\omega)$ by mapping $\langle n, m\rangle$ to the mth subset of n if $m < 2^n$, or else to $\varnothing$. □

Exercise 5.12. The set $\omega^{<\omega} = \bigcup\{\omega^n \mid n < \omega\}$ is countable. (Hint: use Theorems 5.7 and 5.10.) (Visit solution)

Exercise 5.13. Show that if $A \precsim B$, then $A \precsim^* B$.

Exercise 5.14. Show that if $A \precsim^* B$, then $\mathcal{P}(A) \precsim \mathcal{P}(B)$.

5.1.3 Some uncountable sets

Theorem 5.11. *Given any set A, $\mathcal{P}(A) \not\precsim A$ and there is an ordinal α such that $\alpha \not\precsim A$.*

Proof. By Theorem 1.38 we have that $\mathcal{P}(A) \not\precsim^* A$ and therefore $\mathcal{P}(A) \not\precsim A$ either. By Theorem 3.28 there is some well-ordered set which does not inject into A, and by Theorem 4.12 this well-ordered set is isomorphic to a unique ordinal α, therefore $\alpha \not\precsim A$. □

As an immediate corollary we have that $\mathcal{P}(\omega)$ as well as some ordinals are uncountable. This extends to another paradox of naive set theory, Cantor's Paradox, which states that the power set of the set of all sets cannot be strictly larger, which means that the notion of "the set of all sets" is inconsistent (as long as Cantor's theorem is provable).

Theorem 5.12 (Cantor's Paradox). *There is no set V such that for all A, $A \precsim V$.* □

Theorem 5.13 (Cantor–Bernstein Theorem). *If $A \precsim B$ and $B \precsim A$, then $A \sim B$.*

Proof. Since $A \precsim B$, there is some $B_0 \subseteq B$ such that $A \sim B$, that is, $\text{rng}(f)$ for some injective $f\colon A \to B$. Note that the composition $f \circ g$ is a function from B into B_0. So, we may assume that $A = B_0$ and $f = \text{id}_A$ without loss of generality. Therefore, we are in the situation where $A \subseteq B$ and $g\colon B \to A$ is an injective function.

If g was a bijection, there is nothing to do. Otherwise, define by recursion

$$E_0 = A \setminus \text{rng}(g), \quad E_{n+1} = g[E_n].$$

Next, let $E = \bigcup\{E_n \mid n < \omega\}$. We note that $a \in E$ if and only if $g(a) \in E$, since if $a \in E$, then $a \in E_n$ for some $n < \omega$ and then $g(a) \in E_{n+1}$, and if $g(a) \in E$, then by the choice of E_0, it cannot be that $g(a) \in E_0$. Therefore $g(a) \in E_{n+1}$ for some $n < \omega$, and so $a \in E_n$.

Now define $h\colon B \to A$ by

$$h(b) = \begin{cases} b & \text{if } b \in E \\ g(b) & \text{if } b \notin E \end{cases}$$

and let us see that h is a bijection.

If $b_1 \neq b_2$, if they are both in E, then $h(b_1) = b_1 \neq b_2 = h(b_2)$, and if they are both outside of E, then $h(b_i) = g(b_i)$ and since g is injective, $h(b_1) = g(b_1) \neq g(b_2) = h(b_2)$. In the case where exactly one of them is in E, say b_1, we have that $h(b_1) = b_1 \in E$ and $h(b_2) = g(b_2)$, but since $b_2 \notin E$, $g(b_2) \notin E$ as well, so $h(b_1) \neq h(b_2)$.

If $a \in A$, then if $a \in E$, then $h(a) = a$ and therefore $a \in \text{rng}(h)$. If $a \notin E$, then $a \notin E_0$ and therefore $a \in \text{rng}(g)$, that is $a = g(b)$ for some b. Moreover, it is impossible for b to be in E, since $g(b) \notin E$. And therefore $h(b) = g(b) = a$, and h is a bijection as wanted. □

Theorem 5.14. $\mathbb{R}$ *is uncountable.*

Proof. Suppose that $f\colon \omega \to \mathbb{R}$ is any function, let us find some $r_f \in \mathbb{R} \setminus \operatorname{rng}(f)$. This will show that no function $\omega \to \mathbb{R}$ can be a bijection. Let $r_n = f(n)$ and let d^n_m be the mth digit in the decimal expansion of r_n, choosing if need be, the one which has infinitely many 0.

We define d^f_m to be 7 in the case that $d^m_m \leq 4$, and $d^f_m = 2$ if $d^m_m \geq 5$. Let r_f be the real number whose decimal expansion is given by the d^f_m. Then $r_f \notin \operatorname{rng}(f)$. If it were, then for some n, $r_f = f(n)$, but then $d^n_n = d^f_n$. However this is clearly impossible by the choice of d^f_n. □

We can do better than prove that $\mathbb{R}$ is uncountable.

Theorem 5.15. $\mathbb{R} \sim \mathcal{P}(\omega)$.

Proof. It is enough to show that $\mathbb{R} \precsim \mathcal{P}(\omega)$ and $\mathcal{P}(\omega) \precsim \mathbb{R}$. For the first one, note that $\omega \sim \mathbb{Q}$ and therefore $\mathcal{P}(\omega) \sim \mathcal{P}(\mathbb{Q})$. We saw that $\mathbb{R}$ can be interpreted as non-empty, proper initial segments of $\mathbb{Q}$ without maximal elements. Therefore, in that interpretation, $\mathbb{R} \subseteq \mathcal{P}(\mathbb{Q})$, which gives us the first part.

In the other part, given $A \subseteq \omega$ we let

$$r_A = \sum_{n \in A} \frac{2}{3^{n+1}} = \sum_{n=0}^{\infty} \frac{2\chi_A(n)}{3^{n+1}},$$

where $\chi_A\colon \omega \to 2$ is given by $\chi_A(n) = 0$ when $n \notin A$ and otherwise $\chi_A(n) = 1$. Note that this is a convergent series, so r_A is a well-defined real number. If $A \neq B$, let $n = \min A \triangle B$, without loss of generality $n \in A$. Then $r_A - r_B > \frac{2}{3^{n+1}}$ and therefore the two are different and $r(A) = r_A$ is an injective function from $\mathcal{P}(\omega)$ into $\mathbb{R}$. □

Exercise 5.15. Show that $\mathbb{R} \sim (0,1) \sim [0,1)$. Conclude that $\mathbb{R} \sim \mathbb{R} \times 2$.

Exercise 5.16. $\mathbb{R}^\omega \sim \mathbb{R}$. (Hint: Use the rules of exponentiation and $\omega \times \omega \sim \omega$.) (Visit solution)

Remark

At this point in the story, Cantor wondered, is there any set $X \subseteq \mathbb{R}$ such that $\omega \prec X \prec \mathbb{R}$? Cantor conjectured that the answer is negative. That while $\mathbb{R}$ was uncountable, it was the smallest uncountable size. He managed to prove that such X, if it exists at all, cannot be a closed set. This conjecture, called now *The Continuum Hypothesis*, was put by David Hilbert as the first problem in his famous list of 23 problems. It was eventually shown to be irrefutable by Kurt Gödel in 1938, and unprovable in 1963 by Paul J. Cohen.

5.2 Cardinalities of well-ordered sets

5.2.1 Initial ordinals

Definition 5.16. We say that an ordinal α is an *initial ordinal* if for every $\beta < \alpha$, $\alpha \not\precsim \beta$.

Exercise 5.17. Every finite ordinal is initial, as well as ω itself. (Visit solution)

Exercise 5.18. If α is an ordinal, then there is an initial ordinal $\beta \leq \alpha$ such that $\alpha \sim \beta$.

Exercise 5.19. Show that $\omega + 1$ is not an initial ordinal.

Remark

The idea behind an initial ordinal is that if α is an initial ordinal, then α is the least order type we can impose on α as a set. So we can re-order ω to have an order type $\omega+\omega$, but we cannot do it in any way that results in an order type smaller than ω. And so α is an initial ordinal if and only if every proper initial segment of α is not equipotent with α itself.

Proposition 5.17. *If A is a set of initial ordinals, then $\sup A$ is an initial ordinal.*

Proof. Let $\beta = \sup A$, if $\beta \in A$, i.e. $\beta = \max A$, then by our assumption β is an initial ordinal. If $\beta \notin A$, suppose that $\gamma < \beta$, then there is some $\alpha \in A$ such that $\gamma < \alpha$, but since α is an initial ordinal $\alpha \not\precsim \gamma$, so in particular, as $\alpha \subseteq \beta$, there is no injection $\beta \to \gamma$, so $\beta \not\precsim \gamma$. □

Exercise 5.20. Show that if α is an initial ordinal, then either α is finite or α is a limit ordinal. Find a limit ordinal which is not an initial ordinal. (Visit solution)

Exercise 5.21. If α and β are initial ordinals, then $\alpha \precsim \beta$ holds if and only if $\alpha \leq \beta$. In particular, in this case if $\alpha \sim \beta$ it holds that $\alpha = \beta$.

Notation 5.18. For a set A we denote by $\aleph(A)$ (*pronounced "aleph (of A)"*) the least ordinal α such that $\alpha \not\precsim A$. We call $\aleph(A)$ the "Hartogs number of A".

Exercise 5.22. For every set A, $\aleph(A)$ exists and it is an initial ordinal.

Definition 5.19. We define by recursion the infinite initial ordinals: $\omega_0 = \omega$; $\omega_{\alpha+1} = \aleph(\omega_\alpha)$; if α is a limit ordinal and ω_β were defined for all $\beta < \alpha$, $\omega_\alpha = \sup\{\omega_\beta \mid \beta < \alpha\}$.

So, much like how ω was the least infinite ordinal, ω_1 is the least uncountable ordinal, ω_2 is the next uncountable initial ordinal, etc.

Theorem 5.20. *If δ is an initial ordinal then δ is a finite ordinal or else there is some α such that $\delta = \omega_\alpha$.*

Proof. Suppose this is not true and let δ be the least initial ordinal which is infinite and not ω_α for any α, clearly $\delta > \omega = \omega_0$. Let $D \subseteq \delta$ be the set of initial ordinals, if $\sup D = \delta$, by the minimality of δ each of the ordinals in D is of the form ω_β for some β and therefore $\delta = \omega_\alpha$ for some α. If $\sup D = \omega_\alpha < \delta$, we claim that $\delta = \omega_{\alpha+1}$ since given any $\gamma < \delta$ if $\omega_\alpha < \gamma$, then either $\gamma \precsim \omega_\alpha$ or else there is an initial ordinal δ' such that $\omega_\alpha < \delta' < \gamma < \delta$, but by assumption δ' is ω_β, so $\delta' \leq \omega_\alpha$. □

Theorem 5.21. *If α is an infinite ordinal, then $\alpha \sim \alpha \times \alpha$.*

Proof. It is enough to prove the claim for initial ordinals, since if $\beta \leq \alpha$ is an initial ordinal such that $\alpha \sim \beta$, then $\beta \times \beta \sim \alpha \times \alpha$. Define an order on pairs of ordinals, $\langle\alpha,\beta\rangle <_G \langle\gamma,\delta\rangle$ if and only if either $\max\{\alpha,\beta\} < \max\{\gamma,\delta\}$, or $\max\{\alpha,\beta\} = \max\{\gamma,\delta\}$ and $\langle\alpha,\beta\rangle <_{\text{Lex}} \langle\gamma,\delta\rangle$.

We can check that $<_G$ is a well-ordering of $\delta \times \delta$ for any ordinal δ. We now claim that if α is an ordinal then $\operatorname{otp}(\omega_\alpha \times \omega_\alpha, <_G) = \omega_\alpha$ which provides us with a bijection between $\omega_\alpha \times \omega_\alpha$ and ω_α. We prove this by induction on α. We will show that any proper initial segment of $\langle\omega_\alpha \times \omega_\alpha, <_G\rangle$ has an order type smaller than ω_α.

First, if $\gamma < \delta < \omega_\alpha$ then $\langle\gamma,\delta\rangle, \langle\delta,\gamma\rangle <_G \langle\delta,\delta\rangle$. In particular, if $\alpha = 0$ so that $\omega_\alpha = \omega$, and $n < \omega$, then $I_{<_G}(\langle n,n\rangle)$ can only have at most n^2 different points, so it is finite. Therefore, $\langle\omega \times \omega, <_G\rangle$ is a well-order whose proper initial segments are finite, so it is isomorphic to ω.

Next, suppose that for every $\beta < \alpha$, $\operatorname{otp}(\omega_\beta \times \omega_\beta, <_G) = \omega_\beta$. We will show that the same holds for ω_α. For every $\delta < \omega_\alpha$ there is a bijection between $I_{<_G}(\langle 0, \delta+1\rangle)$ and $(\delta+1) \times (\delta+1)$. Since $\delta < \omega_\alpha$, it must be that $\delta + 1 \sim \omega_\beta$ for some $\beta < \alpha$, by the induction hypothesis $\omega_\beta \times \omega_\beta \sim \omega_\beta$ and therefore $\operatorname{otp}(\delta \times \delta, <_G) < \omega_\alpha$. So, $\langle \omega_\alpha \times \omega_\alpha, <_G\rangle$ is a well-order whose proper initial segments all have order type strictly below ω_α, and ω_α embeds into the order, and therefore the two must be isomorphic. □

5.3 Cardinal numbers

Numbers should represent some quantity. In the case of the ordinal numbers, we represent "how long is a well-order". In the sense of cardinality we want to represent the idea of how many elements a set has. Since injections and bijections compare the sizes of sets, we can take $A \sim B$ to mean "same number of elements", so the cardinal numbers should represent that.

Notation 5.22. We write $|A|$ to denote *the cardinal number* of a set A. We require that $A \sim B$ if and only if $|A| = |B|$. We will write $|A| \leq |B|$ if $A \precsim B$ and $|A| \leq^* |B|$ if $A \precsim^* B$.

We want to have a good way of interpreting the cardinal numbers with specific sets. Much like how we chose canonical representatives for each well-ordered set, we want to have the cardinal number of A be a set with the same number of elements as A, if possible. Unfortunately, in general, this requires the Axiom of Choice, and so we do have to make some concessions. However, in the case of sets which can be well-ordered we can use the initial ordinals.

Notation 5.23. If A can be well-ordered, then $|A|$ is the initial ordinal equipotent with A.

We will write $|A| + |B|$ to denote the cardinal of $A \times \{0\} \cup B \times \{1\}$. Similarly, $|A| \cdot |B|$ is the cardinal number of $A \times B$, and $|A|^{|B|}$ is the cardinal number of A^B.

Notation 5.24. To differentiate between the ordinal arithmetic and cardinal arithmetic we will use $\aleph_\alpha$ when discussing the initial ordinal ω_α in its role as a cardinal. Sometimes we will use Greek letters such as κ to denote an $\aleph$ number, and we will be clear what arithmetical operations are being used.

Using this notation we now have that $|\mathbb{Q}| = \aleph_0$ and $|\mathbb{R}| = 2^{\aleph_0}$.

Remark

Cantor, who introduced the $\aleph$ notation, had Jewish roots, and chose $\aleph$ as it was the first letter of the Hebrew alphabet. Promptly after this, he rephrased the Continuum Hypothesis as $2^{\aleph_0} = \aleph_1$. Based on Cohen's work, Robert Solovay proved in 1970 that it might be that $\aleph_1 \neq 2^{\aleph_0}$, but the Continuum Hypothesis—*in its original formulation*—still holds. This relates to the Axiom of Choice, which we will discuss in the next chapter.

We say that a cardinal $\aleph_\alpha$ is a *successor cardinal* if α is a successor ordinal, and otherwise it is a *limit cardinal*. If κ is an $\aleph$ number, we write κ^+ to denote its successor, namely, if $\kappa = \aleph_\alpha$, then $\kappa^+ = \aleph_{\alpha+1}$. For example, $\aleph_\omega$ is a limit cardinal, whereas $\aleph_1$ is a successor.

Theorem 5.25. *If $\aleph_\beta \leq \aleph_\alpha$, then $\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \aleph_\alpha$.*

Proof. $\aleph_\alpha \leq \aleph_\alpha + \aleph_\beta \leq \aleph_\alpha + \aleph_\alpha = \aleph_\alpha \cdot 2 \leq \aleph_\alpha \cdot \aleph_\beta \leq \aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$. □

Theorem 5.26. *If κ is an $\aleph$ number, then $\kappa^\kappa = 2^\kappa$.*

Proof. First, $2 < \kappa < 2^\kappa$. Therefore $2^\kappa \leq \kappa^\kappa \leq (2^\kappa)^\kappa = 2^{\kappa\cdot\kappa} = 2^\kappa$. Therefore $2^\kappa = \kappa^\kappa$. □

Exercise 5.23. Simplify $\aleph_3^{\aleph_5} \cdot \aleph_2$. (Visit solution)

Exercise 5.24. Let $[\kappa]^\lambda$ denote $\{A \subseteq \kappa \mid |A| = \lambda\}$. Show that if $\kappa \geq \lambda$, then $\kappa^\lambda \precsim [\kappa]^\lambda \precsim^* \kappa^\lambda$. (Visit solution)

Chapter 6

The Axiom of Choice

Chapter Goals

In this chapter we will learn about

- The Axiom of Choice.
- Countable unions of countable sets are countable.
- Several equivalences of the Axiom of Choice of which *Zorn's Lemma* and the *Well-Ordering Principle* are the most famous.
- Examples of applications of the Axiom of Choice. Most importantly, filters and ultrafilters.

6.1 "I choose to believe what I was programmed to believe!"

When we have a non-empty set, we normally just "choose" an element by stating "let $a \in A$ be some element". Formally, this is understood as an inference rule called "existential instantiation" which allows us to add a new constant to the language and assume it has a certain property, as long as we are in a scenario that we can prove that there is something satisfying the property. This, of course, can be iterated finitely many times, but what if we need to be able to choose members from infinitely many sets at once? In this case, unfortunately, we cannot do this "by hand", but instead we need to have some kind of a rule that "computes" an element of the set. Of course, this is not computation per se, but more closely related to the idea of an oracle. We want the choice to be uniform, and we want it to remain "constant" in the sense that we choose the same element when presented with the same set. This, upon further inspection, leads us to the concept of a choice function defined next.

Definition 6.1. Let A be a set of non-empty sets, we say that f is a *choice function (on A)* if $\operatorname{dom}(f) = A$ and for all $a \in A$, $f(a) \in a$. We will often dispense with the requirement that all the members of A are non-empty sets by implicitly requiring the choice function to be defined on $\{a \in A \mid a \text{ is a non-empty set}\}$.

Exercise 6.1. Show that if $\mathcal{F}$ is a finite family of non-empty sets, then $\mathcal{F}$ admits a choice function. (Hint: Use finite induction.) (Visit solution)

Exercise 6.2. Show that if $\mathcal{F}$ is a family of finite sets of real numbers, then $\mathcal{F}$ admits a choice function. (Visit solution)

Axiom: Choice

Every set of non-empty sets admits a choice function.

Proposition 6.2. *The countable union of countable sets is countable.*

Proof. Suppose that $\{A_n \mid n < \omega\}$ are countable sets, we may assume that none of them is empty. So for each n, the set $F_n = \{f : \omega \to A_n \mid f \text{ is surjective}\}$ is non-empty. Using the Axiom of Choice we get a sequence $f_n \in F_n$. Define now $f : \omega \times \omega \to \bigcup\{A_n \mid n < \omega\}$ by $f(n, m) = f_n(m)$, then this is a surjection. □

Corollary 6.3. *$A \subseteq \omega_1$ is countable if and only if $\sup A < \omega_1$.* □

6.2 The Big Six

The following theorems are all consequences of the axioms we have so far, but we can instead prove something stronger: these are all just equivalent to the Axiom of Choice.

Theorem 6.4 (Well-Ordering Principle). *Every set can be well-ordered.*

Theorem 6.5 (Zorn's Lemma). *Let $\langle P, \leq\rangle$ be a partially ordered set such that if $C \subseteq P$ is a chain, then C has an upper bound, namely, some $p \in P$ such that for all $q \in C$, $q \leq p$. Then $\langle P, \leq\rangle$ has a maximal element.*

Theorem 6.6 (Teichmüller–Tukey Lemma). *Let $\mathcal{F}$ be a family with a finite character. Namely, $A \in \mathcal{F}$ if and only if $\operatorname{fin}(A) \subseteq \mathcal{F}$. Then $\mathcal{F}$ has a $\subseteq$-maximal element.*

Theorem 6.7 (Existence of Representatives). *Let E be an equivalence relation on a set A, then there is some $R \subseteq A$ such that for every $a \in A$ there is a unique $r \in R$ such that $a \mathrel{E} r$.*

Theorem 6.8 (Cardinal Comparability). *For any two sets A and B, $|A| \leq |B|$ or $|B| \leq |A|$.*

Theorem 6.9. *The following are equivalent.*

1. *The Axiom of Choice.*
2. *The Well-Ordering Principle.*
3. *Zorn's Lemma.*
4. *Teichmüller–Tukey Lemma.*
5. *Existence of Representatives.*
6. *Cardinal Comparability.*

Proof. **(1) implies (2):** Assuming the Axiom of Choice, let A be a non-empty set (if A is empty, then it is already well-ordered by itself). We fix a choice function on $\mathcal{P}(A)$, f. Define by recursion on $\aleph(A)$ a function $g : \aleph(A) \to A$, $g(\alpha) = f(A \setminus \operatorname{rng}(g \restriction \alpha))$. Of course, the recursion can only proceed as long as $g \restriction \alpha$ is not surjective, otherwise $A \setminus \operatorname{rng}(g \restriction \alpha) = \varnothing$.

The function g is injective: if $\beta < \gamma$, then $g(\gamma)$ was chosen from a set which did not include $g(\beta)$, so the two must be distinct. Therefore the recursion must have failed at some $\alpha < \aleph(A)$. Let α be the least point for which $g \restriction \alpha$ is surjective. Then $g : \alpha \to A$ is a bijection, and therefore we can define a well-ordering on A given by $a <_g b \iff g^{-1}(a) < g^{-1}(b)$. Easily, $\langle A, <_g\rangle$ is a well-ordering, since g is an isomorphism between it and α.

(2) implies (1): Let $\mathcal{F}$ be a family of non-empty sets. We fix a $\prec$ to be a well-ordering of $\bigcup\mathcal{F}$, and define $f(A) = \min A$ given from this well-ordering.

(1) implies (3): Let $\langle P, \leq\rangle$ be a partially ordered set satisfying the conditions of Zorn's Lemma. We fix a choice function c on $\mathcal{P}(P)$, and we define by recursion a function

$$g(\alpha) = c(\{p \in P \mid \text{For all } q \in \operatorname{rng}(g \restriction \alpha), q < p\}).$$

By induction we can see that $\operatorname{rng}(g \restriction \alpha)$ is a chain for every α, and therefore the recursion will continue. However, much like in the previous case, this g is injective, so the recursion must fail somewhere below $\aleph(P)$. If α is the least ordinal for which $g(\alpha)$ cannot be defined, then $g \restriction \alpha$ is a chain such that if $p \in P$ is an upper bound of the chain, which exists by our assumption on $\langle P, \leq\rangle$, then $p \in \operatorname{rng}(g)$. In this case, we claim that p is a maximal element. Otherwise, there will be some $p' \in P$ such that $p < p'$, but since p is an upper bound, so must be p', and since p is an upper bound, if $p' \in \operatorname{rng}(g)$, then $p' \leq p$, in contradiction to our assumption. Therefore p must be maximal as wanted.

(3) implies (1): Let $\mathcal{F}$ be a family of non-empty sets. Let P be the set of all partial choice functions. Namely, all f such that there is some $\mathcal{F}' \subseteq \mathcal{F}$ which f is a choice function on $\mathcal{F}'$. And let P be ordered by $\subseteq$. Let us check that the condition for Zorn's Lemma holds. Given a chain $C \subseteq P$, $\bigcup C = f$ is a function, since a chain of functions are a family of compatible functions, and clearly if $g \in C$, then $g \subseteq f$ by definition. Let $\mathcal{F}' = \bigcup\{\operatorname{dom} g \mid g \in C\}$, then $\mathcal{F}' \subseteq \mathcal{F}$ by the definition of P, and f is a choice function on $\mathcal{F}'$, since if $A \in \mathcal{F}'$, then there is some $g \in C$ such that $A \in \operatorname{dom} g$ and $f(A) = g(A) \in A$. Therefore f satisfies the condition for being in P, and so it is an upper bound. By Zorn's Lemma, there is a maximal element, f. Suppose that $\operatorname{dom}(f) \neq \mathcal{F}$, then there is some $A \in \mathcal{F} \setminus \operatorname{dom}(f)$. Pick such A and pick some $a \in A$ and let $g = f \cup \{\langle A, a\rangle\}$, then g is also a choice function from a subset of $\mathcal{F}$, so $g \in P$, and $f \subsetneq g$. This is impossible since f was maximal, and so $\operatorname{dom}(f) = \mathcal{F}$ and therefore $\mathcal{F}$ admits a choice function.

(3) implies (4): Suppose that $\mathcal{F}$ is a family with a finite character. We claim that $\langle \mathcal{F}, \subseteq\rangle$ satisfies the condition for Zorn's Lemma. If $C \subseteq \mathcal{F}$ is a chain, then either it has a maximal element, in which case it serves as an upper bound, or else it does not. If C does not have a maximal element, we claim that $A = \bigcup C$ is an upper bound of C. Clearly, every $B \in C$ satisfies $B \subseteq A$, so we need only to check that $A \in \mathcal{F}$. Since $\mathcal{F}$ has a finite character, it is enough to check that every finite subset of A is in $\mathcal{F}$. However, if $\{a_0, \dots, a_{n-1}\} \subseteq A$, then for each $i < n$, $A_i = \{B \in C \mid a_i \in B\}$ is non-empty. Therefore $\{A_i \mid i < n\}$ is a finite family of non-empty sets. Let $B_i \in A_i$ be a choice from each A_i, then $\{B_i \mid i < n\} \subseteq C$ is a finite chain and there is some j such that $B_i \subseteq B_j$ for all $i < n$. Therefore, $\{a_0, \dots, a_{n-1}\}$ is a finite subset of B_j, and since $B_j \in \mathcal{F}$, it follows that this finite set is there as well. By Zorn's Lemma, $\mathcal{F}$ must have a maximal element.

(4) implies (5): Suppose that E is an equivalence relation on A. We say that T is a *set of representatives* if for every $a, b \in T$, $a\ E\ b$ if and only if $a = b$. In other words, T meets each equivalence class on *at most* one element. Let $\mathcal{F}$ be the family of all sets of representatives, we claim that it has a finite character. Indeed, if $T \notin \mathcal{F}$, then there is some $\{a, b\} \subseteq T$ such that $a\ E\ b$ and $a \neq b$. So if all finite subsets of T are themselves sets of representatives, T must be as well. The other part is trivial, since if T is a set of representatives and $S \subseteq T$, then S must be a set of representatives as well. By the Teichmüller–Tukey Lemma, $\mathcal{F}$ has a maximal element, R. We claim that R is a system of representatives. Otherwise, there is some $a \in A$ such that $R \cap a/E = \varnothing$, then $R \cup \{a\}$ is a set of representatives, so $R \cup \{a\} \in \mathcal{F}$ and therefore R cannot be maximal there.

(5) implies (1): Let $\mathcal{F}$ be a family of non-empty sets. Consider the set $\{\langle A, a\rangle \mid a \in A \in \mathcal{F}\}$ and the equivalence relation $\langle A, a\rangle \ E \ \langle B, b\rangle$ if and only if $A = B$. Let R be a system of representatives, then for every $A \in \mathcal{F}$ there is a unique $a \in A$ such that $\langle A, a\rangle \in R$. In other words, R is a choice function.

(2) implies (6): Since every two sets can be well-ordered, it must be that they can be compared, since if $A \sim \alpha$ and $B \sim \beta$, then $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$, so $|A| \leq |B|$ or $|B| \leq |A|$ must hold.

(6) implies (2): Let A be an arbitrary set, then by definition $\aleph(A) \nleq |A|$, and therefore it must be that $|A| < \aleph(A)$. Therefore, there is an injection from A into some ordinal which lets us define a well-ordering on A. □

Remark

There are many more formulations and equivalents of the Axiom of Choice. Enough that a whole book was written about them with hundreds of equivalents.

Notation 6.10. Let $\{A_i \mid i \in I\}$ be a family of sets, we write $\prod_{i\in I} A_i$ to denote the set of all functions $f : I \to \bigcup\{A_i \mid i \in I\}$ such that $f(i) \in A_i$.

Exercise 6.3. Assume the Axiom of Choice. $\prod_{i\in I} A_i = \varnothing$ if and only if for some $i \in I$, $A_i = \varnothing$.

Exercise 6.4. Show that Zorn's Lemma is equivalent to the statement "If $\langle P, \leq\rangle$ is a partial order satisfying the condition of Zorn's Lemma, then for every $p \in P$ there is some $q \in P$ such that $p \leq q$ and q is a maximal element.

Exercise 6.5. Show that if $\langle P, \leq\rangle$ is a partial order, then it contains a maximal chain. Show that this principle is in fact also equivalent to the Axiom of Choice. (Hint: Use Zorn's Lemma.) (Visit solution)

Exercise 6.6. A is uncountable if and only if $\aleph_1 \leq |A|$. (Visit solution)

6.3 Applications

6.3.1 Vector spaces

Recall that a vector space over a field F is a set V equipped with an addition $+$ and scalar multiplication operators. Fixing V to be a vector space over some field F, we say that $v_0, \dots, v_n$ are *linearly dependent* if there are $\lambda_i \in F \setminus \{0\}$ such that $\lambda_0 v_0 + \dots + \lambda_n v_n = 0$, and otherwise we say that the vectors are linearly independent. We say that $B \subseteq V$ is a linearly independent set if every finite subset is linearly independent.

We say that $B \subseteq V$ is a spanning set if for every $v \in V$ there is some $v_0, \dots, v_n \in B$ and $\lambda_0, \dots, \lambda_n \in F$ such that $v = \lambda_0 v_0 + \dots + \lambda_n v_n$. We say that B is a basis if it is a linearly independent and spanning set.

Exercise 6.7. Show that B is a basis if and only if it is a maximal linearly independent set if and only if it is a minimal spanning set.

Theorem 6.11. *Every vector space has a basis.*

Proof. Let V be a vector space over a field F. Let $\mathcal{F}$ be the family of all linearly independent subsets of V, then $\mathcal{F}$ has a finite character, so by the Teichmüller–Tukey Lemma, $\mathcal{F}$ contains a maximal element, B, which is a maximal linearly independent set. □

Remark

Interestingly enough, this statement is equivalent to the Axiom of Choice, but in order to prove that we require one more additional axiom that $\in$ is a well-founded relation on every set. We will touch upon this axiom when we return to set theory later on.

6.3.2 Filters and ultrafilters

Given a set A we want to have an abstraction to the idea of "practically all the elements of A are in B". In a sense, this tells us when is a subset of A "large", but it is better to think of it as "almost everything". Two things we require is that $\varnothing$ is not almost everything, and that A itself is everything and therefore "almost everything". We also want that two things which are "almost everything" intersect on something that is "almost everything". Finally, we want that if X is almost everything and $X \subseteq Y$, then Y is almost everything as well. This leads us to the following definition.

Definition 6.12. Let A be a non-empty set. We say that $\mathcal{F}$ is a *filter (on A)* if $\mathcal{F} \subseteq \mathcal{P}(A)$ satisfies:

1. $A \in \mathcal{F}$ and $\varnothing \notin \mathcal{F}$,
2. $X, Y \in \mathcal{F}$ implies $X \cap Y \in \mathcal{F}$,
3. $X \in \mathcal{F}$ and $X \subseteq Y \subseteq A$ imply that $Y \in \mathcal{F}$.

We say that $\mathcal{F}$ is an *ultrafilter* if for every $X \subseteq A$, either $X \in \mathcal{F}$ or $A \setminus X \in \mathcal{F}$.

Exercise 6.8. Suppose that A is a non-empty set, fix some $a \in A$, $\mathcal{F}_a = \{B \subseteq A \mid a \in B\}$ is an ultrafilter.

Exercise 6.9. Suppose that A is an infinite set, then $\mathcal{F}_{\text{fin}} = \{B \subseteq A \mid A \setminus B \in \text{fin}(A)\}$ is a filter.

Exercise 6.10. Suppose that $\mathcal{F}$ is a filter on A, then $\mathcal{F} \cup \{X\}$ extends to a filter if and only if for every $Y \in \mathcal{F}$, $X \cap Y \neq \varnothing$. (Visit solution)

Exercise 6.11. Suppose that $\mathcal{F}$ is a filter on A, then $\mathcal{F}$ is an ultrafilter if and only if $\mathcal{F}$ is a maximal filter under $\subseteq$.

Definition 6.13. Suppose that $\mathcal{F}$ is a filter on a set A. We say that $\mathcal{F}$ is a *principal* filter if $\bigcap \mathcal{F} \in \mathcal{F}$. Otherwise, we say that $\mathcal{F}$ is a *free (or non-principal)* filter.

Exercise 6.12. If $\mathcal{F}$ is a principal ultrafilter on A, then there is some $a \in A$ such that $\mathcal{F} = \mathcal{F}_a$. (Visit solution)

Exercise 6.13. If $\mathcal{F}$ is a filter on a finite set, then $\mathcal{F}$ is principal.

Exercise 6.14. Let $\mathcal{F}$ be a filter on a set X. If $\mathcal{F}_{\text{fin}} \subseteq \mathcal{F}$, then $\mathcal{F}$ is free. Show that if $\mathcal{F}$ is an ultrafilter, then the other implication holds as well, namely, it is free if and only if $\mathcal{F}_{\text{fin}} \subseteq \mathcal{F}$.

Theorem 6.14. *Every filter extends to an ultrafilter.*

Proof. Let A be a non-empty set and let $\mathcal{F}$ be a filter on A. Let P be the set of all filters on A which extend $\mathcal{F}$, ordered by $\subseteq$. We will show that $\langle P, \subseteq \rangle$ satisfies the conditions of Zorn's Lemma, then it has a maximal element which is an ultrafilter.

Suppose that $C \subseteq P$ is a chain if $C = \varnothing$, then $\mathcal{F}$ is an upper bound for it, so we may assume $C \neq \varnothing$. We claim that $\mathcal{U} = \bigcup C$ is a filter on A and that $\mathcal{F} \subseteq \mathcal{U}$, and therefore $\mathcal{U}$ is an upper bound for C. It is easy to see that $\mathcal{F} \subseteq \mathcal{U}$, since C is non-empty, and therefore there is some $\mathcal{F}' \in C$ and $\mathcal{F} \subseteq \mathcal{F}' \subseteq \mathcal{U}$. To see that $\mathcal{U}$ is a filter on A, it is certainly the case that $A \in \mathcal{U}$ as it is in all the elements of C, and $\varnothing \notin \mathcal{U}$ as it is in none of them. Suppose that $X \subseteq Y$ and $X \in \mathcal{U}$, then there is some $\mathcal{F}' \in C$ such that $X \in \mathcal{F}'$, and therefore $Y \in \mathcal{F}'$ as well. Similarly, if $X, Y \in \mathcal{U}$ there is some $\mathcal{F}_X, \mathcal{F}_Y \in C$ such that $X \in \mathcal{F}_X$ and $Y \in \mathcal{F}_Y$. Since C was a chain, $\mathcal{F}_X \subseteq \mathcal{F}_Y$ or vice versa, and so without loss of generality $X, Y \in \mathcal{F}_X$, and therefore $X \cap Y \in \mathcal{F}_X$ as well. Therefore the conditions for Zorn's Lemma hold and there is a maximal filter extending $\mathcal{F}$ which is an ultrafilter. □

Exercise 6.15. Show that there is a free ultrafilter on ω. (Visit solution)

Remark

The principle "Every filter extends to a maximal ultrafilter" is, surprisingly enough, weaker than the Axiom of Choice!

6.3.3 Injections and surjections

Theorem 6.15. *Suppose that $f \colon A \to B$ is a surjective function, then there is some $g \colon B \to A$ which is injective and $f \circ g = \mathrm{id}_B$.*

Proof. For each $b \in B$, let $A_b = \{\langle b, a\rangle \in B \times A \mid f(a) = b\}$. By the Existence of Representatives we have that there is $g \subseteq \bigcup\{A_b \mid b \in B\}$ which is a system of representatives. In particular, for every $b \in B$ there is a unique $a \in A$ such that $\langle b, a\rangle \in g$, and therefore g is a function. To see that g is injective, suppose that $g(b) = a$, then $f(a) = b$, and so if $g(b) = g(b') = a$, then $b = f(a) = b'$ and so $b = b'$. Finally, by the very definition, $f \circ g(b) = f(g(b)) = b$. □

Exercise 6.16. Show that Theorem 6.15 is equivalent to the Axiom of Choice.

Theorem 6.16 (The Partition Principle). *If $|A| \leq^* |B|$, then $|A| \leq |B|$.*

Proof. If $|A| \leq^* |B|$, then either $A = \varnothing$, in which case $|A| \leq |B|$, or else there is a surjection $f \colon B \to A$. By Theorem 6.15, there is an injective function $g \colon A \to B$, so $|A| \leq |B|$. □

Remark

We do not know if The Partition Principle is equivalent to the Axiom of Choice. As of 2023, this problem has been open for over 120 years.

6.3.4 Unspecified recursion

Recall the definition of Dedekind-finite sets was "If $f \colon A \to A$ is an injective function, then f is a bijection". We want to show now that this is equivalent to the standard definition of finiteness. First, let us prove the following equivalence.

Proposition 6.17. *A is Dedekind-finite if and only if $\aleph_0 \nleq |A|$.*

Proof. Suppose that $f\colon \omega \to A$ is an injective function, define $g\colon A \to A$ as follows:

$$g(a) = \begin{cases} f(n+1) & \text{if } a = f(n) \text{ for some } n < \omega, \\ a & \text{otherwise.} \end{cases}$$

It is not hard to check that g is injective: if $g(a) = g(b)$, in the case where both $a, b \notin \operatorname{rng}(f)$ this is trivial; as is the case exactly one of them is in $\operatorname{rng}(f)$. If $a = f(n)$ and $b = f(m)$, then by the injectivity of f, $n = m$, so $a = b$. At the same time, g is not surjective, as $f(0) \notin \operatorname{rng}(g)$.

In the other direction, let $f\colon A \to A$ be injective and not surjective and let $a_0 \in A \setminus \operatorname{rng}(f)$. We define by recursion a function $g\colon \omega \to A$ by $g(0) = a_0$, $g(n+1) = f(g(n))$. If g is not injective, let n be the least such that for some $m < n$, $g(n) = g(m)$. It is impossible that $m = 0$, as $g(0) \notin \operatorname{rng}(f)$ and $g(k) \in \operatorname{rng}(f)$ for all $k > 0$, so $g(n) = f(g(n-1)) = f(g(m-1))$, but this means that $m - 1 < n - 1$ is a also a counterexample to injectivity, in contradiction to the choice of n as minimal. □

So, it is enough to show now, assuming the Axiom of Choice, that every infinite set has a countably infinite subset. The naive approach would be to "pick an element, pick another, pick another, etc." and eventually form a countably infinite subset. But this approach is based on a recursion argument which does not specify the step function, but instead relies on the weaker situation "the pool of candidates for the next step is non-empty". In order to apply the definition by recursion we need to convert this into a specific function, which is exactly where the Axiom of Choice is being used. This approach can be seen in the proofs of (2) and (3) from (1) in Theorem 6.9. And we can make it more explicit.

Theorem 6.18 (Unspecified Recursion). *Suppose that α is an ordinal, B is a set, and R is a relation such that $\operatorname{dom}(R) = \bigcup\{B^\beta \mid \beta \leq \alpha\}$ and $\operatorname{rng}(R) \subseteq B$. Then there is a function $f\colon \alpha \to B$ such that for all $\beta < \alpha$, $f \restriction \beta \mathrel{R} f(\beta)$.*

Proof. Fix a choice function c on $\mathcal{P}(B)$. For $h\colon \beta \to B$ for some $\beta < \alpha$, define C_h to be $\{b \in B \mid h \mathrel{R} b\}$. Now consider the function $G\colon B^{\subseteq\alpha} \to B$ given by

$$G(h) = \begin{cases} c(C_h) & h \in B^\beta \text{ for some } \beta < \alpha, \\ c(B) & \text{otherwise.} \end{cases}$$

By Theorem 3.23 there exists a function f such that $f(\beta) = G(f \restriction \beta)$, and by the definition of G, $f \restriction \beta \mathrel{R} f(\beta)$ as wanted. □

Theorem 6.19. *Every infinite set is Dedekind-infinite.*

Proof. Let A be an infinite set and let $R = \{\langle f, a\rangle \mid \text{For some } n < \omega, f \in A^n \text{ and } a \notin \operatorname{rng}(f)\}$. Since A is infinite, for any $n < \omega$ and $f\colon n \to A$, $\operatorname{rng}(f) \neq A$. So $\operatorname{dom}(R) = A^{<\omega}$, and by the Unspecified Recursion theorem, there is a function $f\colon \omega \to A$ such that $f \restriction n \mathrel{R} f(n)$.

We claim that this f must be injective, and indeed, if $m < n$, then $f(n) \notin \operatorname{rng}(f \restriction n)$ but $f(m) \in \operatorname{rng}(f \restriction n)$, so the two must be distinct. □

Chapter 7

First-Order Logic

Chapter Goals

In this chapter we will learn about

- The syntax of first-order logic. This includes terms, formulas, sentences, and theories.
- The semantics of first-order logic. This includes interpretation, assignment, and truth.
- The concept of a structure, a theory, and a model.
- Embeddings and isomorphisms.
- The theory of dense linear orders without endpoints has a unique (up to isomorphism) countable model.
- The concept of a complete theory.

7.1 Syntax

We want to describe a loose collection of rules that given some symbols, we can treat them as a language and form expressions. For brevity we want to have some symbols appear in any and all situations: $=$, variables, connectives, and quantifiers.

First and foremost, the equality symbol $=$ is always going to be understood as what equality is: a binary relation where two things are equal if they are the same. In all of our contexts, the collection of variables is countably infinite and will usually be denoted by $\{x_n \mid n < \omega\}$, although we will often, when clarity allows, use other letters to denote variables as well, such as x, y, z and more.

The connectives are $\wedge$ for "and" (or conjunction), $\vee$ for "or" (or disjunction), $\neg$ for negation, and finally $\rightarrow$ for "implies". The quantifiers are $\forall$ for "every" and $\exists$ for "some".

To form statements with meaning, discussing relations and functions, we need to have a language with those symbols. So a language is going to be a collection of symbols. It is going to be an implicit assumption that a symbol cannot have more than one meaning, i.e. if x_n is a variable, then we will not use it to denote a function.

Formally, a language is a tuple $\langle \mathcal{R}, \mathcal{F}, \mathcal{C}, \text{arity} \rangle$ where $\mathcal{R}$, $\mathcal{F}$, and $\mathcal{C}$ are pairwise disjoint sets and arity is a function, $\text{arity}\colon \mathcal{R} \cup \mathcal{F} \rightarrow \omega$. We call the members of $\mathcal{R}$ "*relation symbols*", the members of $\mathcal{F}$ "*function symbols*", and the members of $\mathcal{C}$ "*constant symbols*". If $\text{arity}(f) = n$, we say that f is an "n-ary" function or relation symbol. If $\text{arity}(f) = 1$ we say it is "unary", if $\text{arity}(f) = 2$ we say it is "binary", and if $\text{arity}(f) = 3$ we say it is "ternary".

Remark

We will often write a language as a set with some symbols explaining what is the type of each symbol and its arity. For example we will say "$\{<\}$ *is the language where* $<$ *is a binary relation symbol*" to mean that $\mathcal{F} = \mathcal{C} = \varnothing$ and $\mathcal{R} = \{<\}$ and $\text{arity}(<) = 2$.

Next we define terms and formulas. Terms are syntactic constructs that allow us to discuss "objects in the structure" whereas formulas describe "properties of objects in the structure".

Let us consider as a "running example" for the following definitions the language $\{F, R, c\}$ where F is a binary function symbol, R is a ternary relation symbol, and c is a constant symbol. So, $\mathcal{F} = \{F\}$, $\mathcal{R} = \{R\}$, and $\mathcal{C} = \{c\}$ with $\text{arity}(F) = 2$ and $\text{arity}(R) = 3$.

Definition 7.1 (Terms). Given a language $\mathcal{L}$ we define a *term* as a sequence of symbols that is formed in one of the following recursive rules:

1. If x is a variable, then x is a term.
2. If c is a constant symbol, then c is a term.
3. If f is an n-ary function symbol and $t_0, \ldots, t_{n-1}$ are terms, then $f(t_0, \ldots, t_{n-1})$ is term.

Clauses (1) and (2) are called *atomic terms*. We denote by $\text{Term}_{\mathcal{L}}$ the set of all terms in the language $\mathcal{L}$, and we omit the $\mathcal{L}$ when it is clear from context.

For example, $F(c, F(x_0, x_1))$ is a term, but $R = F(x_0)$ is not.

Definition 7.2 (Formulas). Given a language $\mathcal{L}$ we define a *formula* as a sequence of symbols that is formed in one of the following recursive rules:

1. If t_0, t_1 are terms, then $t_0 = t_1$ is a formula.
2. If R is an n-ary relation symbol and $t_0, \ldots, t_{n-1}$ are terms, then $R(t_0, \ldots, t_{n-1})$ is a formula.
3. If φ is a formula, $\neg\varphi$ is a formula.
4. If φ, ψ are formulas, then $\varphi * \psi$ is a formula, when $* \in \{\wedge, \vee, \rightarrow\}$.
5. If φ is a formula and x is a variable, then $\exists x\varphi$ and $\forall x\varphi$ are formulas.

Clauses (1) and (2) are called *atomic formulas*. We denote by $\text{Form}_{\mathcal{L}}$ the set of all formulas in the language $\mathcal{L}$, and we omit the $\mathcal{L}$ when it is clear from context.

For example, $R(c, F(c, c), F(F(c, c), c))$ is an atomic formula, and $\exists x R(c, x, F(x, c))$ is a formula, but not an atomic one. On the other hand, $F(R(x, y), c)$ is not a valid formula nor a term.

Exercise 7.1. In our example language, write the sets of all the atomic terms and all the atomic formulas and calculate their cardinalities. (Visit solution)

Definition 7.3. If φ is a formula and x is a variable, we define the notion of x being a *free variable of* φ by recursion:

1. If φ is atomic and x appears in φ, then x is a free variable of φ.
2. If φ is a formula and x is a free variable of φ, then it is a free variable of $\neg\varphi$.
3. If φ and ψ are formulas and x is a free variable of φ or ψ, then x is a free variable of $\varphi * \psi$ for $* \in \{\wedge, \vee, \to\}$.
4. If φ is a formula and x is a variable, then x is **not** a free variable of $\exists x\varphi$ and $\forall x\varphi$.

If x is not free in φ we say that it is *bound* there. When we write $\varphi(x_0, \dots, x_{n-1})$ we understand this to mean that if x is a free variable of φ, then $x \in \{x_0, \dots, x_{n-1}\}$, although when we simply write φ we do not necessarily mean that φ does not have free variables.

If φ does not have any free variables we say that it is a *sentence* and we write $\mathrm{Sent}_{\mathcal{L}}$ as the set of all sentences.

For example, $F(c, c) = c$ has no free variables. However, $R(x_0, x_1, c)$ has two. On the other hand, $\forall x_1(R(x_0, x_1, x_1) \to x_0 = x_1)$ only has x_0 as a free variable.

Remark

All the definitions we saw so far were essentially recursive: a base case was given (e.g., atomic terms) and rules of how to produce new things from it via "finite means". This lends itself for a "subformula" or "subterm" order, which is the transitive closure of the relation defined by these rules (for example, φ is a subformula of $\psi \to (\varphi \wedge \varphi')$, even though there are two steps from φ to it). Since terms and formulas are finite objects, it is easy to check that these is a well-founded relation. Therefore we can (and judicially will) use recursion and induction on these orders, usually referring to it as "induction on the structure of the formula".

Exercise 7.2. Define by recursion the function $\mathrm{FV}\colon \mathrm{Form} \to \{x_n \mid n < \omega\}$ such that $\mathrm{FV}(\varphi)$ is exactly the set of free variables of φ.

As we said, a sentence is a formula with no free variables. In that case, the formula is no longer discussing the properties of members of the structure and their relationship, but instead the properties of the structure itself.

Remark

Our logic is called *first-order logic* since we only can apply quantifiers to "first-order variables" which are intended to represent objects of the world. We also have *second-order logic* where we can quantify over functions and relations, even if they are not specified in the language. There are *infinitary logics* where we allow the disjunction or conjunction of infinitely many formulas at once, or quantification over infinitely many variables at once. Some logics will have different quantifiers, and others will have syntax rules defined in very different ways. We will focus on first-order logic for this course, at least unless specified otherwise.

Traditionally, we would spend some time on understanding a proof system which allows us to mechanically manipulate sequences of sentences in order to deduce one sentence from another. There are multiple proof systems, but they are all equivalent. One of the features of model theory, however, is that it works hard to avoid any and all need for the proof system, as the focus is much more on the semantics sides of the logic.

7.2 Semantics

Syntax, on its own, is meaningless. It is a list of symbols which we can manipulate according to various rules, and while this is interesting on its own, it is just letters on a page. Much like that the text you are reading right now will not make any sense for someone who does not speak English (and they might not even be able to recognise a "text" to begin with), the agreed-upon meaning of the symbols on the page is how we communicate. Semantics is the side of "meaning" in logic. In the context of mathematics, this means the actual mathematical structure we are studying, such as $\langle A, <\rangle$ or $\langle \mathbb{R}, +, \cdot, 0, 1, <\rangle$, etc.

Let us fix some language $\mathcal{L}$. Formally speaking, a structure is a pair $\langle M, \Sigma\rangle$ where M is a set and Σ is an interpretation function whose domain is $\mathcal{F} \cup \mathcal{R} \cup \mathcal{C}$ and:

1. For every $F \in \mathcal{F}$, $\Sigma(F)\colon M^{\operatorname{arity}(F)} \to M$ is an arity(F)-ary function defined on M.
2. For every $R \in \mathcal{R}$, $\Sigma(R) \subseteq M^{\operatorname{arity}(R)}$ is an arity(R)-ary relation defined on M.
3. For every $c \in \mathcal{C}$, $\Sigma(c) \in M$ is an element of M.

We will usually write F^M, R^M, and c^M as a shorthand to denote $\Sigma(F)$, $\Sigma(R)$, $\Sigma(c)$. And we will often omit the superscript if it is clear from context. And much like the shorthand in the case of the language, we will often simply write $\langle M, F, R, c\rangle$ to denote a structure, e.g. $\langle \mathbb{N}, +, \cdot, 0\rangle$.

Next, we are ready to discuss the truth. Specifically, we want to evaluate things like $x < y$ or $F(x) = y$, etc. But of course, these are meaningless until we assign "meaning" to the variables.

Definition 7.4. Let M be an $\mathcal{L}$-structure. If $\sigma\colon \{x_n \mid n < \omega\} \to M$ we say that σ is an *assignment*. If x is a variable and $m \in M$, we denote by $\sigma[x/m]$ the assignment given by

$$\sigma[x/m](x_n) = \begin{cases} m & x_n = x, \\ \sigma(x_n) & x_n \neq x. \end{cases}$$

Definition 7.5. Let M be an $\mathcal{L}$-structure and σ an assignment. We extend the domain of σ to Term by recursion:

1. If $c \in \mathcal{C}$, then $\sigma(c) = c^M$.
2. If $t = F(t_0, \dots, t_{n-1})$ for some n-ary function symbol F and terms $t_0, \dots, t_{n-1}$, then $\sigma(t) = F^M(\sigma(t_0), \dots, \sigma(t_{n-1}))$.

Definition 7.6. Let M be an $\mathcal{L}$-structure and let σ be an assignment. We define the *satisfaction relationship*, $M \models_\sigma \varphi$, by recursion on Form.

1. If φ is $t_0 = t_1$, $M \models_\sigma \varphi$ if and only if $\sigma(t_0) = \sigma(t_1)$.
2. If φ is $R(t_0, \dots, t_{n-1})$ for some relation symbol R, $M \models_\sigma R(\sigma(t_0), \dots, \sigma(t_{n-1}))$ if and only if $\langle \sigma(t_0), \dots, \sigma(t_{n-1})\rangle \in R^M$.
3. If $\varphi = \neg\psi$, then $M \models_\sigma \varphi$ if and only if $M \not\models_\sigma \psi$.
4. If $\varphi = \varphi_0 \wedge \varphi_1$, then $M \models_\sigma \varphi$ if and only if $M \models_\sigma \varphi_0$ and $M \models_\sigma \varphi_1$.
5. If $\varphi = \varphi_0 \vee \varphi_1$, then $M \models_\sigma \varphi$ if and only if $M \models_\sigma \varphi_0$ or $M \models_\sigma \varphi_1$.

6. If $\varphi = \varphi_0 \to \varphi_1$, then $M \models_\sigma \varphi$ if and only if $M \models_\sigma \neg\varphi_0$ or $M \models_\sigma \varphi_1$.
7. If $\varphi = \exists x\psi$, then $M \models_\sigma \varphi$ if and only if there exists some $m \in M$ such that $M \models_{\sigma[x/m]} \psi$.
8. If $\varphi = \forall x\psi$, then $M \models_\sigma \varphi$ if and only if for every $m \in M$, $M \models_{\sigma[x/m]} \psi$.

Exercise 7.3. Let $\varphi(x_0, \ldots, x_{n-1})$ be a formula. Show that if $\sigma_0(x_i) = \sigma_1(x_i)$ for all $i < n$, then $M \models_{\sigma_0} \varphi$ if and only if $M \models_{\sigma_1} \varphi$. (Visit solution)

As a consequence of the exercise, if φ is a sentence, then $M \models \varphi$ independently of σ, and we can simply omit it from the notation. We will often abuse the notation and write $M \models \varphi(m)$ to mean that for some assignment, σ, for which $\sigma(x) = m$, $M \models_\sigma \varphi(x)$.

When $M \models \varphi$, we say that φ is *true in* M, so otherwise we say that it is false.

Exercise 7.4. Suppose that M is an $\mathcal{L}$-structure. Show that given any $\varphi \in \mathrm{Sent}$, φ is either true or false in M. (Hint: Use the previous exercise and the definition of negation.)

Exercise 7.5. Write in the language of $\{<\}$, where $<$ is a binary relation symbol the axioms for a partial order, a strict linear order. Write the following formulas in this language:

1. There exists a minimum element,
2. there are no maximal elements,
3. x is the successor of y. (Visit solution)

Exercise 7.6. Write down a sentence φ_n in the empty language such that $M \models \varphi_n$ if and only if $|M| \geq n$. (Visit solution)

Exercise 7.7. Show that $M \models \exists x\varphi(x)$ if and only if $M \models \neg\forall x\neg\varphi$ and $M \models \forall x\varphi$ if and only if $M \models \neg\exists x\neg\varphi$.

7.3 Structures, theories, and models

Notation 7.7. From this point onwards, we will use $\bar{x}$ notation to denote n-tuples of "the appropriate length" and we understand that if $\bar{t}$ is a tuple, then its members have the form t_i. We will often abuse the notation and write $\bar{a} \in A$ to mean that each $a_i \in A$. If $\bar{a}$ is a tuple and f is a function, we will write $f(\bar{a})$ to mean the tuple whose members are $f(a_i)$.

Definition 7.8. Suppose that M is an $\mathcal{L}$-structure. We say that N is a *substructure* of M if $N \subseteq M$, for every constant symbol, c, $c^M \in N$ and $c^N = c^M$; for every n-ary function symbol F, $F^N = F^M \restriction N^{\mathrm{arity}(F)} \colon N^{\mathrm{arity}(F)} \to N$; and for every n-ary relation symbol R, $R^N = R^M \cap N^{\mathrm{arity}(R)}$.

Definition 7.9. Suppose that M and N are two $\mathcal{L}$-structures, then $e \colon N \to M$ is an *embedding of $\mathcal{L}$-structures* if

1. for every constant symbol c, $e(c^N) = c^M$;
2. for every n-ary function symbol F, $e(F^N(\bar{x})) = F^M(e(\bar{x}))$; and
3. for every n-ary relation symbol R, $\bar{x} \in R^N$ if and only if $e(\bar{x}) \in R^M$.

We implicitly require that e is injective, that is to say, that it also preserves $=$, which is not in the language, but it is part of the underlying logic. If e is surjective, then we say that e is an *isomorphism of $\mathcal{L}$-structures*, and we write $N \cong M$.

For example, we saw the definition of embeddings in the context of partially ordered sets already, and we have seen that every countable linear order embeds into $\mathbb{Q}$. In the empty language, i.e. the one without any constant, function, or relation symbols, an injection is an embedding.

Exercise 7.8. Suppose that M is an $\mathcal{L}$-structure and $f\colon M \to N$ is a bijection. Then there is an interpretation function making N an $\mathcal{L}$-structure such that f is an isomorphism. (Visit solution)

Remark

The above exercise is known as "transport of structure" where we copy the structure from M to N by using a bijection. So as a corollary, every countably infinite set can be well-ordered in order type $\omega \cdot \omega + 5$, or can be given an equivalence relation where every equivalence class has exactly 53 elements. Every set of size $2^{\aleph_0}$ can be made isomorphic to $\mathcal{P}(\omega)$ or to $\operatorname{fin}(\mathbb{R})$, both ordered by $\subseteq$; or to the complex numbers or the real numbers.

Exercise 7.9. Suppose that $e\colon N \to M$ is an isomorphism, then $M \models \varphi$ if and only if $N \models \varphi$. Moreover, if σ is an assignment in N, then $e \circ \sigma$ is an assignment in M and $M \models_{e\circ\sigma} \varphi$ if and only if $N \models_\sigma \varphi$.

Exercise 7.10. Find an embedding from $\langle \mathbb{R}, + \rangle$ into $\langle \mathbb{R}^+, \cdot \rangle$ where $\mathbb{R}^+ = \{r \in \mathbb{R} \mid 0 < r\}$ and $+$ and $\cdot$ are the standard operations. Prove or disprove: the embedding you have found is an isomorphism. (Visit solution)

Often we care not about "arbitrary interpretation of the syntax" but instead specific ones. We have an intended meaning to the symbols and we want to make sure that they behave "as wanted". For this we need the notion of a theory and a model.

Definition 7.10. If $\mathcal{L}$ is a language we say that T is a *theory* (or an $\mathcal{L}$-theory) if $T \subseteq \mathrm{Sent}_{\mathcal{L}}$.

If T is a theory, we write $M \models T$ to mean that for every $\varphi \in T$, $M \models \varphi$. We will say, in this case that M is a *model of T*.

Whereas we will often want to study and understand structures an interpretations of languages to understand what can or cannot be expressed in the language, in the case of model theory we want to focus on interpretations which have some intended meaning.

For example, we can consider the language $\{+, 0\}$ where $+$ is a binary function symbol and 0 is a constant symbol. These could mean practically anything, but we want to understand them from a point of view of a very specific theory. Perhaps it is the theory of abelian groups. Namely,

φ_1: $\forall x \forall y \forall z (x + (y + z) = (x + y) + z)$,

φ_2: $\forall x \exists y (x + y = 0)$,

φ_3: $\forall x (x + 0 = x)$, and

φ_4: $\forall x \forall y (x + y = y + x)$.

So, now we are not concerned with *all* structures in our language, but just those which satisfy the above sentences, or axioms, which give the symbols some *intended meaning*.

Definition 7.11. Let T be an $\mathcal{L}$-theory and let M be a model of T. We say that $N \subseteq M$ is a *submodel* if it is a substructure of M which is also a model of T.

Exercise 7.11. Let T be the theory of abelian groups. Show that $\langle \mathbb{Z}, + \rangle$ is a submodel of $\langle \mathbb{Q}, + \rangle$. Find some φ in the language of abelian groups that holds in one structure and not in the other.

Remark

Recall that a field is a structure $\langle F, +, \cdot, 0, 1 \rangle$ where $+$ and $\cdot$ are commutative, with $\cdot$ distributive over $+$. 0 is the identity for $+$ and 1 is the identity for $\cdot$. Every element has an additive inverse and any non-zero element has a multiplicative inverse.

Exercise 7.12. Let $\mathcal{L}$ be the language of fields, $\{+, \cdot, 0, 1\}$ with $+$ and $\cdot$ being binary function symbols, 0 and 1 being constant symbols. Suppose that F is a field such that for every $n < \omega$, F satisfies that the n-fold addition $1 + \cdots + 1 \neq 0$. Show that there is an embedding $e \colon \mathbb{Q} \to F$, where $\mathbb{Q}$ is the field of rational numbers with its standard operations.

Exercise 7.13. Let $\mathcal{L}$ be the language with a single binary relation symbol, $\{<\}$. Write the theory for dense linear orders without endpoints. (Visit solution)

Theorem 7.12. *Suppose that $\langle A, < \rangle$ is a dense linear order such that A is countably infinite and does not have endpoints. Then $A \cong \mathbb{Q}$.*

Proof. We enumerate $A = \{a_n \mid n < \omega\}$ and $\mathbb{Q} = \{q_n \mid n < \omega\}$. We proceed to define an isomorphism, f, by recursion in two steps. Let k be the least such that $a_k \notin \operatorname{dom} f$, and let n be the least such that $f \cup \{\langle a_k, q_n \rangle\}$ is still an isomorphism, we know that such n exists, so there is a minimal n. Next, let k be the least such that $q_k \notin \operatorname{rng} f$, and let n be the least such that $f \cup \{\langle a_n, q_k \rangle\}$ is still an isomorphism, and we know that such n exists, so there is a minimal n.

Let us check that f is an isomorphism: if $a_k < a_n$, then by step $\max\{k, n\}$ of the recursion we are guaranteed that both a_k and a_n are in the domain of f which is order preserving, so $f(a_k) < f(a_n)$. Similarly, if $f(a_k) < f(a_n)$, then by step $\max\{k, n\}$ both are in the domain and therefore $a_k < a_n$. To see that f is surjective, note that if $n < \omega$, then by the nth step of the recursion we have that $q_n \in \operatorname{rng} f$, and so f must be surjective. ☐

This argument is known as Cantor's "*back and forth argument*". Let us see how the argument works by enumerating $\mathbb{Q}$ twice. If $\bar{p} = \langle 0, \frac{1}{2}, -10, \frac{1}{4}, \dots \rangle$ and $\bar{q} = \langle -1, -2, 3, \frac{1}{3}, \dots \rangle$, then $f(0) = -1$, then $f(-10) = -2$, then $f(\frac{1}{2}) = 3$, then $f(\frac{1}{4}) = \frac{1}{3}$, and so on.

Exercise 7.14. Suppose that $\bar{p}$ and $\bar{q}$ are two n-tuples of rational numbers such that $f(p_i) = q_i$ is an isomorphism between them. Show that there is some isomorphism $F \colon \mathbb{Q} \to \mathbb{Q}$ such that $F(p_i) = q_i$. (Visit solution)

7.4 Complete theories

Definition 7.13. Let T be an $\mathcal{L}$-theory and $\varphi \in \mathrm{Sent}_{\mathcal{L}}$. We say that $T \models \varphi$ if whenever $M \models T$, $M \models \varphi$. We say that T is a *complete theory* if for every φ either $T \models \varphi$ or $T \models \neg\varphi$. We say that T is *consistent* if there is no φ such that $T \models \varphi$ and $T \models \neg\varphi$.

Remark

Usually we define consistent in terms of the syntax. Namely, T is consistent if it does not prove a falsehood, i.e. $\varphi \land \neg\varphi$. But since we are not talking about the proof system, we will just mention the *Completeness and Soundness Theorems* which say that for first-order logic "T proves φ" is exactly the same as $T \models \varphi$. From this we can conclude that "consistent" is the same as "has a model".

Exercise 7.15. Let M be some $\mathcal{L}$-structure and let $T = \{\varphi \mid M \models \varphi\}$. Show that T is a consistent and complete theory. (Visit solution)

Exercise 7.16. If T is a consistent and complete theory and $M \models T$, then $T \models \varphi$ if and only if $M \models \varphi$. (Visit solution)

For an $\mathcal{L}$-structure M, let $\mathrm{Th}_{\mathcal{L}}(M) = \{\varphi \in \mathrm{Sent}_{\mathcal{L}} \mid M \models \varphi\}$ be *the theory of* M. The first exercise is claiming that $\mathrm{Th}_{\mathcal{L}}(M)$ is a consistent and complete theory. The second exercise shows that if T is consistent and complete, then $\{\varphi \mid T \models \varphi\}$ is $\mathrm{Th}_{\mathcal{L}}(M)$ for any $M \models T$.

Implicitly, from here on end, we will assume that our theories are consistent, unless explicitly proved or stated otherwise.

Exercise 7.17. Suppose that any two models of T are isomorphic, then T is complete.

Exercise 7.18. Show that in the empty language, $\{\exists x \forall y (x = y)\}$ is complete. In contrast, show that $\{\exists x (x = x)\}$ is not complete. (Visit solution)

While it might be tempting to think that T is complete if and only if all of its models are isomorphic, this is not the case. We will see later a refinement of this idea which does in fact hold, as well as the reason for which the naive approach fails.

Chapter 8

Elementary, my dear Moose.

Chapter Goals

In this chapter we will learn about

- Elementary equivalence and elementary substructures, as well as the difference between the two notions.
- The Tarski–Vaught Criterion for elementarity.
- Ultraproducts, Łoś's Theorem, and ultrapowers.
- The Compactness Theorem for first-order logic.
- Various applications of these theorems in proving that certain ideas cannot be expressed in some first-order languages.
- An original and curious proof that $\sqrt{2}$ is irrational.

8.1 Elementarity

Much like in the case where we wanted to distinguish a structure from a model, and a substructure from a submodel, sometimes we want to require more. We want the theories of the two structures to be the same, and even more.

Definition 8.1. Suppose that M and N are two $\mathcal{L}$-structures. We say that M is *elementarily equivalent* to N, denoted by $M \equiv N$, if $\mathrm{Th}(M) = \mathrm{Th}(N)$.

For example, any two isomorphic structures are elementarily equivalent. And sometimes this is the other way around.

Exercise 8.1. Suppose that $M \equiv N$ and $N = \{x\}$, show that $M \cong N$.

Exercise 8.2. Show that the following are equivalent:

1. T is complete.
2. Any two models of T are elementarily equivalent. (Visit solution)

Sometimes we can say more.

Proposition 8.2. *Suppose that α is an ordinal and $\alpha \equiv \omega$. Then $\alpha = \omega$.*

Proof. First, note that ω does not have a maximal element, i.e. $\omega \models \forall x \exists y(x < y)$, and therefore α must be a limit ordinal as well. Secondly, ω satisfies that every element is either 0, which is the minimum element, or it is a successor. So,

$$\omega \models \forall x(\forall y(x < y \vee x = y) \vee \exists y(y < x \wedge \forall z(z < x \to y < z \vee y = z))).$$

By elementarity α must satisfy the same. Therefore, α is a limit ordinal, so $\omega \leq \alpha$ and if $\beta < \alpha$, then $\beta = 0$ or a successor. In particular, if $\omega < \alpha$, then ω must be a successor, which is not the case, so $\omega = \alpha$. □

More generally, two structures that are elementarily equivalent will have similar structures when it comes to finite things that we can express. For example, "there is exactly one element satisfying φ" or "There are exactly five things which can satisfy ψ".

The above shows, for example, that if $M \subseteq \omega$ is a substructure of ω (in the language $\{<\}$), and M is both elementarily equivalent to ω and happened to be a transitive set, then $M = \omega$, as well as that if $\omega \subseteq M$ and they are elementarily equivalent and M is transitive, then $M = \omega$. This might not be too surprising, since the requirements that M is transitive is quite strong. But maybe it is not necessary? The following exercise shows otherwise.

Exercise 8.3. Show that if $M \subseteq \omega$ is any infinite set, then $M \equiv \omega$. (Hint: Isomorphisms preserve truth.)

To resolve this, we introduce the notion of an elementary substructure.

Definition 8.3. Let M be an $\mathcal{L}$-structure. We say that $N \subseteq M$ is an *elementary substructure or elementary submodel (of M)* if it is a substructure and whenever $\varphi(\bar{x})$ is an $\mathcal{L}$-formula, and $\bar{a} \in N$, then $M \models \varphi(\bar{a})$ if and only if $N \models \varphi(\bar{a})$. We write $N \prec M$ to mean that N is an elementary submodel of M. In the other direction we say that M is an *elementary extension* of N.

In other words, being an elementary submodel means that not only you are an elementary equivalent structure, but whenever $a_0, \ldots, a_{n-1}$ are members of the substructure, they will have the same "relations and properties that the language can express" in both the small and the large structure.

Exercise 8.4. If $N \prec M$, then $N \equiv M$.

Proposition 8.4. *Suppose that $M \subseteq \omega$ is an elementary submodel of ω, then $M = \omega$.*

Proof. We know that M must be infinite, otherwise they will not be elementary equivalent to begin with. We prove this by induction. If $0 \notin M$, then let $n = \min M$, then $M \models$ "n is the minimum", whereas ω does not. Therefore $0 \in M$. Suppose that $n \in M$, then ω satisfies that n is not the maximum, and that $n + 1$ is the successor of n. Therefore M must satisfy that n is not the maximum, but then let $k = \min\{m \in M \mid n < m\}$, then M satisfies that "k is the successor of n" and by elementarity ω will too, so $k = n + 1$, and therefore if $n \in M$, then $n + 1 \in M$ as well. By induction $M = \omega$. □

We can use the following theorem to determine whether a substurcture is elementary or not.

Theorem 8.5 (Tarski–Vaught Criterion). *Let N be an $\mathcal{L}$-structure and let $M \subseteq N$ be a substructure. Then $M \prec N$ if and only if whenever $\bar{m} \in M$ such that $N \models \exists x\varphi(x, \bar{m})$, then there is $x \in M$ such that $N \models \varphi(x, \bar{m})$.*

Proof. One direction is trivial. If $M \prec N$, then whenever $N \models \exists x\varphi(x, \bar{m})$, it holds that $M \models \exists x\varphi(x, \bar{m})$, so there is some $x \in M$ such that $M \models \varphi(x, \bar{m})$, and therefore by elementarity $N \models \varphi(x, \bar{m})$.

The other direction is proved by induction on the structure of φ that $M \models \varphi(\bar{m})$ if and only if $N \models \varphi(\bar{m})$. For quantifier free formulas this is true simply because M is a substructure. For negation and connectives this follows by verifying the truth tables. Finally, it is enough to prove this for $\exists x\varphi$, since $\forall x\varphi$ is equivalent to $\neg\exists x\neg\varphi$.

If $M \models \exists x\varphi(\bar{m})$, then there is some $x \in M$ such that $M \models \varphi(x, \bar{m})$, and by the induction hypothesis on φ, $N \models \varphi(x, \bar{m})$, and therefore $N \models \exists x\varphi(\bar{m})$ as wanted. In the other direction, suppose that $N \models \exists x\varphi(x, \bar{m})$, by the assumption we made, there is some $x \in M$ such that $N \models \varphi(x, \bar{m})$, and by the induction hypothesis, $M \models \varphi(x, \bar{m})$ and so $M \models \exists x\varphi(x, \bar{m})$ as wanted. $\square$

Theorem 8.6. *Let $\mathcal{L} = \{E\}$ be the language with a single binary relation symbol E and let T be the theory stating that E is an equivalence relation with two infinite equivalence classes. That is, T includes the axioms of equivalence relations as well as $\exists x\exists y(x \not E\ y \wedge \forall z(x\ E\ z \vee y\ E\ z))$ and the sentences $\forall \bar{x}\exists y(\bigwedge_{i<n} x_i \neq y \wedge \bigvee_{i<n} x_i\ E\ y)$, for every n-tuple, $\bar{x}$.*

If $N \models T$ and $M \subseteq N$ is a substructure such that $M \models T$, then $M \prec N$.

Proof. Note that if $M \models T$, then really we are just saying that M has infinitely many elements from each equivalence class.

Suppose that $N \models \exists x\varphi(x, \bar{m})$ for some n-tuple $\bar{m} \in M$. Let $n \in N$ be such that $\varphi(n, \bar{m})$ holds, then n appears in exactly one of the two equivalence classes. Let $m \in M$ be a point such that $m \neq m_i$ for all i and $m\ E\ n$. We can find such m since M satisfies that each equivalence class is infinite.

Now considering the function $f\colon M \to M$ which simply exchanges n and m, that is

$$f(x) = \begin{cases} m & \text{when } x = n, \\ n & \text{when } x = m, \\ x & \text{when } x \notin \{m, n\}. \end{cases}$$

Then f is an isomorphism of M with itself, since it is a bijection and it preserves E, the only two points moved by f are in the same equivalence class. Therefore, $N \models \varphi(f(n), \bar{m})$ which is to say $N \models \varphi(m, \bar{m})$, and so by the Tarski–Vaught Criterion $M \prec N$ as wanted. $\square$

8.2 Łoś's Theorem

Definition 8.7. Let $\{M_i \mid i \in I\}$ be a collection of sets and let $\mathcal{F}$ be a filter on I. The reduced product $\prod_{i\in I} M_i/\mathcal{F}$ is the set of equivalence classes of the relation $\prod_{i\in I} M_i$ defined by $f \sim_{\mathcal{F}} g$ if and only if $\{i \mid f(i) = g(i)\} \in \mathcal{F}$. In the case where $\mathcal{F}$ is an ultrafilter we call it an *ultraproduct.*

Exercise 8.5. Show that $\sim_{\mathcal{F}}$ is an equivalence relation.

We will use $[f]_{\mathcal{F}}$ to denote each equivalence class in the context of ultraproducts.

Definition 8.8. Let $\mathcal{L}$ be a language and suppose that $\{M_i \mid i \in I\}$ are $\mathcal{L}$-structures and $\mathcal{U}$ is an ultrafilter on I, then the *ultraproduct*, $M = \prod_{i\in I} M_i/\mathcal{U}$, has a naturally defined interpretation as an $\mathcal{L}$-structure given by the following:

1. For a constant symbol c, $c^M = [f]_\mathcal{U}$ where $f(i) = c^{M_i}$.
2. For an n-ary function symbol F, $F^M([\bar{f}]_\mathcal{U}) = [g]_\mathcal{U} \iff \{i \in I \mid F^{M_i}(\bar{f}(i)) = g(i)\} \in \mathcal{U}$.
3. For an n-ary relation symbol R, $[\bar{f}]_\mathcal{U} \in R^M \iff \{i \in I \mid \bar{f}(i) \in R^{M_i}\} \in \mathcal{U}$.

Theorem 8.9 (Łoś's Theorem). *Suppose that $\{M_i \mid i \in I\}$ are $\mathcal{L}$-structures, $\mathcal{U}$ is an ultrafilter on I, and M is the ultraproduct. Then for every formula $\varphi(\bar{x})$,*

$$M \models \varphi([\bar{f}]_\mathcal{U}) \text{ if and only if } \{i \in I \mid M_i \models \varphi(\bar{f}(i))\} \in \mathcal{U}.$$

Proof. We prove this by induction on the complexity of φ, for readability we will omit the free variables from the notation. For atomic formulas this just the definition of the ultraproduct structure, for negation we use the fact that $\mathcal{U}$ is an ultrafilter and therefore either $\{i \mid M_i \models \varphi\}$ or $\{i \mid M_i \models \neg\varphi\}$ must be in $\mathcal{U}$. For the connectives we use the closure of $\mathcal{U}$ under finite intersections as well as the fact that $A \in \mathcal{U}$ implies that for all B, $A \cup B \in \mathcal{U}$.

Suppose that for φ the conclusion holds, then $M \models \exists x\varphi(x)$ if and only if there exists some $[f]_\mathcal{U}$ such that $M \models \varphi([f]_\mathcal{U})$. By the induction hypothesis, $\{i \in I \mid M_i \models \varphi(f(i))\} \in \mathcal{U}$ and therefore for each such i, $M_i \models \exists x\varphi(x)$, so $\{i \in I \mid M_i \models \exists x\varphi(x)\} \in \mathcal{U}$. In the other direction, suppose that $\{i \in I \mid M_i \models \exists x\varphi(x)\} \in \mathcal{U}$, then for every such i, there is some $m_i \in M_i$ such that $M_i \models \varphi(m_i)$. Fix some $h \in \prod_{i\in I} M_i$ and define

$$f(i) = \begin{cases} m_i & \text{if } M_i \models \varphi(m_i) \\ h(i) & \text{if no such } m_i \text{ exists.} \end{cases}$$

Then $\{i \in I \mid M_i \models \varphi(f(i))\} \in \mathcal{U}$, so $M \models \varphi([f]_\mathcal{U})$ and therefore $M \models \exists x\varphi(x)$. Finally, since $M \models \forall x\varphi$ if and only if $M \models \neg\exists x\neg\varphi(x)$, then the case of $\forall x\varphi(x)$ follows from $\exists x\varphi$ and negation. □

Remark

Łoś's Theorem might feel somewhat trivial. However, once we have defined the interpretation function making M into an $\mathcal{L}$-structure, the satisfaction relation is now set in stone. What Łoś's Theorem tells us is that it behaves in a very democratic way: if the majority of structures (i.e. a set of indices in $\mathcal{U}$) agrees that something is true, then it is true in the ultraproduct. This is a very important criterion, since it can be easily checked in many cases, whereas checking the explicit truth definition might be more complicated.

Theorem 8.10. *Let $\mathcal{L}$ be a language and let T be a theory. Suppose that for every $n < \omega$ there is a model of T with at least n elements. Then T has an infinite model.*

Proof. Let M_n be a model of T such that $|M_n| \geq n$ and let $\mathcal{U}$ be a free ultrafilter on ω, and let M be the ultraproduct. Let φ_n be the statement that there are at least n distinct elements in the structure, then $M_k \models \varphi_n$ whenever $k > n$, and therefore $\omega \setminus \{k < \omega \mid M_k \models \varphi_n\}$ is a finite set, so $M \models \varphi_n$ for all n, and therefore M must be infinite. Moreover, since $M_n \models T$ for all $n < \omega$, clearly $M \models T$ as well. □

Exercise 8.6. Show that there is an infinite group in a very roundabout way.

Exercise 8.7. Find a language $\mathcal{L}$ and a theory T which has infinite models, but its finite models are exactly of sizes 2 and 4. (Visit solution)

Proposition 8.11. *$\sqrt{2}$ is irrational.*

Proof. Let $\mathcal{L}$ be the language of fields, $\{+, \cdot, 0, 1\}$ where $+$ and $\cdot$ are binary function symbols and 0 and 1 are constants. We will need the following fact. For infinitely many prime numbers, p, $\mathbb{F}_p \models \neg\exists x(x \cdot x = 1 + 1)$, where $\mathbb{F}_p$ is the unique field with p elements.

Let $I = \{p < \omega \mid p \text{ is prime and } \mathbb{F}_p \models \neg\exists x(x \cdot x = 1 + 1)\}$, which is an infinite set. Let $\mathcal{U}$ be a free ultrafilter on I, and let F be the ultraproduct $\prod_{p \in I} \mathbb{F}_p/\mathcal{U}$, then $F \models \neg\exists x(x \cdot x = 1 + 1)$. If $\sqrt{2}$ was rational, that is, $\mathbb{Q} \models \exists x(x \cdot x = 1 + 1)$ with $\sqrt{2}$ witnessing this, and $e\colon \mathbb{Q} \to F$ was an embedding, then $F \models e(\sqrt{2}) \cdot e(\sqrt{2}) = e(1) + e(1) = 1 + 1 = 2$, which would contradict Łoś's theorem. So it is enough to show that such e exists.

For this it is enough to see that for any $n < \omega$, the n-fold addition $1 + \cdots + 1 \neq 0$ in F. Note that in $\mathbb{F}_p$, if $n < p$, then the n-fold addition $1 + \cdots + 1 \neq 0$. Since I was infinite, for every $n < \omega$, all but finitely many $p \in I$, will be larger than n. By the same argument as the previous theorem, we get that $F \models 1 + \cdots + 1 \neq 0$. Therefore there is an embedding $e\colon \mathbb{Q} \to F$, and therefore $\sqrt{2}$ is irrational. □

Definition 8.12. We say that an ultraproduct $\prod_{i \in I} M_i/\mathcal{U}$ is an *ultrapower* if there is some M such that for all $i \in I$, $M_i \cong M$. We then simply write $M^I/\mathcal{U}$.

Exercise 8.8. The embedding $j\colon M \to M^I/\mathcal{U}$, where $\mathcal{U}$ is an ultrafilter on I, given by $j(m) = [c_m]_\mathcal{U}$ where $c_m(i) = m$ for all $i \in I$ is an elementary embedding. That is, $j[M] \prec M^I/\mathcal{U}$. This j is called *the ultrapower embedding*. (Visit solution)

Exercise 8.9. Suppose that $\mathcal{U}$ is a principal ultrafilter on I, show that $M \cong M_i$ for some $i \in I$. Find an example where $\mathcal{U}$ is a free ultrafilter, but $M \cong M^I/\mathcal{U}$.

Theorem 8.13. *Let $\mathcal{L}$ be a language with a binary relation symbol $<$. Then there is no $\mathcal{L}$-theory T with infinite models such that $M \models T$ if and only if $<$ is a well-ordering of M.*

Proof. Let M be an infinite model of T where $<$ is a well-ordering of M. Let us write $0 = \min M$, and if $n < \omega$ was defined, $n + 1$ is its successor. As M is infinite it has an initial segment of order type ω and this is simply the canonical embedding of ω as an initial segment of M, so we can just assume that $\omega \subseteq M$ for simplicity.

Let $\mathcal{U}$ be a free ultrafilter on ω and consider $N = M^\omega/\mathcal{U}$, then $N \models T$. We will show that $<^N$ is not a well-ordering of N. For $n < \omega$, let $f_n\colon \omega \to \omega$ given by $f_n(k) = \max\{0, k - n\}$. That is, $f_n(k) = 0$ for $k \leq n$, and for $k > n$, $f_n(k) = k - n$. If $m \leq n$, then for all $i < \omega$, $f_n(i) \leq f_m(i)$, and therefore $\{[f_n]_\mathcal{U} \mid n < \omega\}$ does not have a minimal element. Therefore $<^N$ is not a well-ordering of N. □

Theorem 8.14. *There is an elementary extension of $\mathbb{R}$ as an ordered field where there exists ε such that $\varepsilon > 0$ but for all $n \in \mathbb{N}$, $\varepsilon < \frac{1}{n}$. Moreover, there is one which has the same cardinality as $\mathbb{R}$.*

Proof. Let $\mathcal{U}$ be a free ultrafilter on ω and let $M = \mathbb{R}^\omega/\mathcal{U}$. Since $|\mathbb{R}^\omega| = 2^{\aleph_0}$ and $j\colon \mathbb{R} \to M$ is an elementary embedding, we get that $2^{\aleph_0} \leq |M| \leq 2^{\aleph_0}$ so the cardinality requirement is satisfied.

We let ε be $[f]_\mathcal{U}$ where $f(n) = \frac{1}{n}$. Then for all $n < \omega$, $0 < f(n)$, so $M \models 0 < [f]_\mathcal{U}$. On the other hand, for all $n < \omega$, $\{k < \omega \mid f(k) < \frac{1}{n}\} = \{k < \omega \mid k > n\}$ which is in $\mathcal{U}$ as its complement is a finite set. Therefore $M \models [f]_\mathcal{U} < j(\frac{1}{n})$, as wanted. □

Exercise 8.10. In the context of the previous theorem show that $j[\mathbb{R}]$ has an upper bound but does not have a supremum in $\mathbb{R}^\omega/\mathcal{U}$. Conclude from that the completeness property of $\langle\mathbb{R},<\rangle$, namely "if A has an upper bound, then $\sup A$ exists" is not expressible in first-order logic. (Visit solution)

Remark

The model M in the previous theorem is often referred to as "(the) hyperreals". While all hyperreal fields have the same theory, indeed they are elementarily equivalent as ultrapowers of the same structure, the question of whether or not they are isomorphic is not provable from the standard axioms of set theory.

We can apply similar approach to models of arithmetic, i.e. $\langle\mathbb{N},+,\cdot,0,<\rangle$ and obtain what are known as "non-standard models" of arithmetic. The study of non-standard models of arithmetic is broader and more intricate than simply taking ultrapowers of $\mathbb{N}$, and it is often the case that we want to find models which are not elementarily equivalent to $\mathbb{N}$.

8.3 Compactness

Theorem 8.15 (Compactness Theorem). *Suppose that T is an $\mathcal{L}$-theory such that for every finite $T_0 \subseteq T$ there is a model $M \models T_0$. Then T has a model.*

Proof. For every $T_0 \in \operatorname{fin}(T)$, let $M_{T_0} \models T_0$. Formally speaking, we are using the Axiom of Replacement to find a set large enough which contains enough models for all the $T_0 \in \operatorname{fin}(T)$, and then we are using the Axiom of Choice to choose one for each T_0. For every $\varphi \in T$ let $U_\varphi = \{T_0 \in \operatorname{fin}(T) \mid \varphi \in T_0\}$, then for any finite $T_0 \in \operatorname{fin}(T)$ we have that $\bigcap_{\varphi\in T_0} U_\varphi$ is non-empty, and it is in fact $\{T_1 \in \operatorname{fin}(T) \mid T_0 \subseteq T_1\}$.

Let $\mathcal{F}$ be the filter $\{X \subseteq \operatorname{fin}(T) \mid \exists\varphi \in T \text{ such that } U_\varphi \subseteq X\}$ and let $\mathcal{U}$ be an extension of $\mathcal{F}$ to an ultrafilter. Let $M = \prod_{T_0\in\operatorname{fin}(T)} M_{T_0}/\mathcal{U}$, then for every $\varphi \in T$ the set $\{T_0 \mid \varphi \in T_0\} \in \mathcal{U}$, so for every $\varphi \in T$, $\{T_0 \in \operatorname{fin}(T) \mid M_{T_0} \models \varphi\} \in \mathcal{U}$, so $M \models \varphi$. Therefore $M \models T$ as wanted. □

Theorem 8.16. *Suppose that $T \models \varphi$, then there is a finite $T_0 \subseteq T$ such that $T_0 \models \varphi$. In particular, T is consistent if and only if every finite $T_0 \subseteq T$ is consistent.*

Proof. Suppose that for every finite $T_0 \subseteq T$, $T_0 \not\models \varphi$. Then there is a model, $M_0 \models T_0 \cup \{\neg\varphi\}$. In particular, every finite subset of $T \cup \{\neg\varphi\}$ has a model. By the Compactness Theorem, $T \cup \{\neg\varphi\}$ has a model, so $T \not\models \varphi$. □

Theorem 8.17. *Let M be an infinite model of some theory T in a language $\mathcal{L}$. Then for every cardinal $\kappa > |M|$ there is a model of T, N, such that $|N| \geq \kappa$.*

Proof. Let $\mathcal{L}^+$ be an expanded language where we add to $\mathcal{L}$ constant symbols $\{c_\alpha \mid \alpha < \kappa\}$, and let T^+ be the theory $T \cup \{c_\alpha \neq c_\beta \mid \alpha < \beta < \kappa\}$. For any finite $T_0 \subseteq T$, only finitely many of the new axioms appear in T_0 and therefore only finitely many of the new constants. Let $M_0 \subseteq M$ be a large enough finite set so that we can interpret $c_\alpha^M \in M_0$ for all α for which c_α appears in one of the axioms of T_0. Now $M \models T_0$ since any axiom of T which lies in T_0 already holds in M, and the new constants are interpreted in distinct ways.

By the Compactness Theorem, T^+ has a model, N, and since $c_\alpha^N \neq c_\beta^N$ for all $\alpha < \beta < \kappa$, the function $f(\alpha) = c_\alpha^N$ is an injection from κ into N as wanted. □

Exercise 8.11. For a cardinal κ, find a language $\mathcal{L}$ and a theory T that if $M \models T$, then $|M| \geq \kappa$. (Visit solution)

Proposition 8.18. *Assuming the Compactness Theorem, every set can be linearly ordered.*

Proof. Let A be a set, and without loss of generality it is an infinite set, as finite sets can be linearly ordered. Let $\mathcal{L}$ be the language $\{<\} \cup \{c_a \mid a \in A\}$ where $<$ is a binary relation symbol and c_a is a constant symbol for each $a \in A$. Let T be the theory with the following axioms:

1. $\forall x \neg(x < x) \wedge \forall x \forall y \forall z(x < y \wedge y < z \rightarrow x < z)$.
2. For every $a, b \in A$ such that $a \neq b$, $c_a \neq c_b$.

We will show that every finite $T_0 \subseteq T$ has a model and therefore T has a model. To see that, note that if $T_0 \subseteq T$ is finite there are only finitely many axioms of the form $c_a \neq c_b$, so the set $A_0 = \{a \in A \mid c_a \text{ appears in some axioms of } T_0\}$ is finite. Therefore we can linearly order A_0 by picking a bijection between A_0 and some $n < \omega$.

By the compactness theorem, T has a model, M, and the injection $i(a) = c_a^M$ is injective since $M \models c_a^M \neq c_b^M$ for all $a \neq b$. Since M is linearly ordered by $<$, we can use this to define a linear ordering on A: $a < b$ if and only if $M \models c_a^M < c_b^M$. □

Remark

The previous proof works without the Axiom of Choice, as long as the Compactness Theorem holds. One can ask, naturally, is the Compactness Theorem equivalent to the Axiom of Choice? The answer is that it is weaker than the Axiom of Choice. Indeed, the Compactness Theorem is equivalent to Theorem 6.14. One can then ask, does "Every set can be linearly ordered" imply the Compactness Theorem, and the answer is again negative. It is, indeed, a strictly stronger principle.

Chapter 9

Skolem in the House of Lions

Chapter Goals

In this chapter we will learn about

- The Löwenheim–Skolem theorems.
- The notion of a definable element, relation, and function.
- Automorphisms of a structure.
- How to use automorphisms to prove the lack of definability of an element.

9.1 Going down?

We saw in Theorem 8.17 that if T has an infinite model, then it has arbitrarily large infinite models. What about very small models? This is not always possible, of course, since if $\mathcal{L}$ contains uncountably many constants, and T proves that they are all distinct, then any model of T must be uncountable. But we can still say something interesting.

Theorem 9.1 (Downward Löwenheim–Skolem). *Suppose that T is a theory with an infinite model M and $X \subseteq M$. Then there is $X \subseteq N \prec M$ such that $|N| \leq |X| + |\mathcal{L}| + \aleph_0$.*

Before we prove this theorem, let us introduce a helpful concept.

Definition 9.2. Let M be a $\mathcal{L}$-structure and let $\varphi(\bar{x}, y)$ be a formula. We say that $f\colon M^n \to M$, where n is the length of $\bar{x}$, is a *Skolem function (for φ)* if whenever $M \models \exists y\varphi(\bar{x}, y)$, then $M \models \varphi(\bar{x}, f(\bar{x}))$.

Note that we do not care about the value of f when $\exists y\varphi(\bar{x}, y)$ is false in M.

Exercise 9.1. Use the Axiom of Choice to show that Skolem functions exist.

Proof. We begin by fixing a Skolem function, f_φ, for any formula of the form $\varphi(\bar{x}, y)$. Let N be the smallest subset of M such that $X \subseteq N$ and N is closed under all the Skolem functions. That is, if $\bar{x} \in N$, then $f_\varphi(\bar{x}) \in N$ whenever it is defined. We claim that $N \prec M$.

Suppose that $M \models \exists y\varphi(\bar{x}, y)$ for some $\bar{x} \in N$, then by the definition of Skolem functions, $M \models \varphi(\bar{x}, f_\varphi(\bar{x}))$. As we have that N is closed under the Skolem functions, $f_\varphi(\bar{x}) \in N$. By the Tarski–Vaught Criterion, $N \prec M$.

Let us compute the cardinality of N. We can write $N = \bigcup_{k<\omega} N_k$ where $N_0 = X$ and $N_{k+1} = N_k \cup \{f_\varphi(\bar{x}) \mid \bar{x} \in N_k, f_\varphi \text{ is a Skolem function}\}$. The cardinality of N_1 is that of X along with at most one element for each $f_\varphi(\bar{x})$. As there are at most $|\text{Form}_\mathcal{L}|$ Skolem functions, the upper bound is $|\text{Form}_\mathcal{L}| \cdot |X^{<\omega}|$. If X is finite, then $|X|^{<\omega}$ is countable, and so the cardinals are just $|\text{Form}_\mathcal{L}| = |\mathcal{L}| + \aleph_0$, and if X is infinite, then $|X| = |X^{<\omega}|$. By induction, this upper bound holds for the cardinality of N_k for all k. Therefore, $|N| \leq |X| + |\text{Form}_\mathcal{L}| = |X| + |\mathcal{L}| + \aleph_0$ as wanted. □

Remark

One immediate corollary of this is that $\mathbb{R}$, in the language of ordered fields, has a countable elementary submodel. This is somewhat counterintuitive, but we can in fact find such elementary submodel explicitly: $\{r \in \mathbb{R} \mid \exists p \in \mathbb{Q}[x] \text{ such that } p(r) = 0\}$.

More strangely, if set theory is consistent, and it has a model, then it has a countable model, M. Since set theory is a foundation of mathematics, we can understand the real numbers $\mathbb{R}$ inside of M, but since M is countable, it can only contain countably many real numbers. The solution to this paradox, called *Skolem's paradox*, is that the notion of "countable" and "uncountable" are relative to the universe of mathematics in which you are working. In other words, in M, there is no bijection between the object M thinks is ω and the object it thinks is the real numbers. Whether or not such bijection exists outside of M is meaningless.

Proposition 9.3. *There is a countable linear order $\langle N, <\rangle$, which is not a well-order, which is an elementary extension of ω.*

Proof. We already saw that any ordinal which is elementarily equivalent to ω must be ω itself, so any proper elementary extension of ω is not going to be well-ordered, as it cannot be isomorphic to any ordinal.

In Theorem 8.13 we saw that if $\mathcal{U}$ is a free ultrafilter on ω, then $\omega^\omega/\mathcal{U}$ is an elementary extension of ω which is not well-ordered. Let $j\colon \omega \to \omega^\omega/\mathcal{U}$ be the ultrapower embedding, and let $c \in \omega^\omega/\mathcal{U} \setminus j[\omega]$. Using Theorem 9.1, take $X = j[\omega] \cup \{c\}$ and let $N \prec \omega^\omega/\mathcal{U}$ be the elementary submodel generated by X. Since the language is finite and X is countable, we get that N is countable. Of course, N is not isomorphic to ω, and therefore not well-ordered, since $c \in N$, and since c is not $j[\omega]$. By replacing $j[\omega]$ with ω itself, we get that N is an elementary extension of ω. □

Exercise 9.2. Suppose that $M_0 \prec N$ and $M_1 \prec N$. If $M_0 \subseteq M_1$, then $M_0 \prec M_1$. Similarly, if $M_0 \prec M_1$ and $M_1 \prec N$, then $M_0 \prec N$. (Visit solution)

Exercise 9.3. Let $\langle I, <\rangle$ be a totally ordered set and $\{M_i \mid i \in I\}$ be $\mathcal{L}$-structures such that for $i < j$, $M_i \prec M_j$. Let $M = \bigcup\{M_i \mid i \in I\}$, then M is an $\mathcal{L}$-structure and $M_i \prec M$ for all $i \in I$.

Proposition 9.4. *Suppose that $\mathcal{L}$ is a countable language and ω_1 is a $\mathcal{L}$-structure. Then there is $\alpha < \omega_1$ such that $\alpha \prec \omega_1$.*

Proof. We define two sequences by recursion, $\alpha_0 = 0$ and M_0 is an elementary substructure generated by α_0, note that it is countable since $\mathcal{L}$ was countable and α_0 was countable. Suppose that M_n was defined and is countable. Since M_n is a countable subset of ω_1, $\sup M_0 = \alpha_{n+1}$ is a countable ordinal, so we can let M_{n+1} be a countable elementary substructure generated by α_{n+1}, and it is countable as well.

Using the Tarski–Vaught Criterion we also have that $M_n \prec M_{n+1}$. To see this, note that if $M_{n+1} \models \exists x\varphi(x, \bar{m})$ for $\bar{m} \in M_n$, then by elementarity, $\omega_1 \models \exists x\varphi(x, \bar{m})$. Since $M_n \prec \omega_1$, it is true that there is some $m \in M_n$ such that $\omega_1 \models \varphi(m, \bar{m})$, and therefore $M_{n+1} \models \varphi(m, \bar{m})$, and elementarity follows.

Note that the sequence we defined satisfies that $\alpha_n \subseteq M_n \subseteq \alpha_{n+1}$. Let

$$\alpha = \sup\{\alpha_n \mid n < \omega\} = \bigcup\{M_n \mid n < \omega\}.$$

Since α is a countable union of countable ordinals, it is a countable ordinal as well, and therefore $\alpha < \omega_1$. By the previous exercise it is also the union of a chain of elementary extensions and therefore $\alpha \prec \omega_1$ as wanted. □

Exercise 9.4. Find an example of a finite language $\mathcal{L}$ and an interpretation function such that ω_2 is an $\mathcal{L}$-structure, and if $\alpha \prec \omega_2$, then $\alpha > \omega_1$. (Visit solution)

Theorem 9.5. *The theory of dense linear orders without endpoints is complete.*

Proof. Suppose that $\langle A, <\rangle$ is a dense linear order without endpoints and therefore infinite. Let $M \prec A$ be a countable elementary submodel, then M is a countable dense linear order and therefore $M \cong \mathbb{Q}$. Given any φ in the language of $\{<\}$, either $\mathbb{Q} \models \varphi$ or $\mathbb{Q} \models \neg\varphi$. Since $\mathbb{Q} \cong M \prec N$, $N \models \varphi$ if and only if $M \models \varphi$ if and only if $\mathbb{Q} \models \varphi$. Therefore $T \models \varphi$ if and only if $\mathbb{Q} \models \varphi$, where T is the theory of dense linear orders. □

Exercise 9.5. If T is a (consistent) $\mathcal{L}$-theory without finite models such that every two countable models of T are isomorphic, then T is complete.

Definition 9.6. Let $\mathcal{L} = \{E\}$ be the language where E is a binary relation, and let T be the theory of graphs (stating that E is symmetric and irreflexive). Recall that a graph $\langle G, E\rangle$ is called a random graph if whenever $A, B \in \mathrm{fin}(G)$ are disjoint, then there is $g \in G \setminus (A \cup B)$ such that $g\ E\ a$ for all $a \in A$ and $g \not\!E\ b$ for all $b \in B$.

Exercise 9.6. Show that the theory of random graphs can be expressed in first-order logic. (Visit solution)

Exercise 9.7. Show that any two countable random graphs are isomorphic, conclude that the theory of random graphs is complete. (Hint: the back-and-forth method is useful here.)

9.2 Going up?

Theorem 9.7 (Upward Löwenheim–Skolem). *Suppose that T is an $\mathcal{L}$-theory with an infinite model M. Then for every $\kappa > |M| + |\mathcal{L}|$ there is an elementary extension of M of size exactly κ.*

Proof. Let $\mathcal{L}^+$ be the augmented language $\mathcal{L}$ along with $\{c_m \mid m \in M\}$ where c_m is a constant symbol. For every $\varphi(\bar{x}) \in \mathrm{Form}_\mathcal{L}$, if $M \models_\mathcal{L} \varphi(\bar{m})$, we let $\varphi^+_{\bar{m}}$ be the sentence where every free occurrence of x_i is replaced by the constant symbol c_{m_i}. We define T^+ as $T \cup \{\varphi^+_{\bar{m}} \mid M \models \varphi(\bar{m})\}$.

It is not hard to check that $M \models_{\mathcal{L}^+} T^+$ when we interpret $c^M_m = m$. Using Theorem 8.17 let N be a model of T^+ such that $|N_0| \geq \kappa$. We claim that $j\colon M \to N_0$ given by $j(m) = c^{N_0}_m$ is an elementary embedding as $\mathcal{L}$-structures, since if $\varphi(\bar{x}) \in \mathrm{Form}_\mathcal{L}$ such that $M \models \varphi(\bar{m})$, then $\varphi^+_{\bar{m}} \in T^+$, therefore $N \models \varphi(\bar{c}^N_{m_i})$, and similarly if $M \not\models \varphi(\bar{m})$, then $\neg\varphi^+_{\bar{m}} \in T^+$ and the same argument applies.

Therefore by replacing $j[M]$ with M itself we get that N is an elementary extension of M. Fix some $A \subseteq N$ such that $M \subseteq A$ and $|A| = \kappa$, then by Theorem 9.1 we have that there is an elementary submodel of N of size κ which contains A, and therefore M, and therefore is an elementary extension of M of size κ. □

The theory T^+ is often called *"the elementary diagram of M"* for elementary reasons.

Corollary 9.8. *There exists uncountable elementary extensions of ω in any cardinality.* □

> **Remark**
>
> Using this we can get a sort of "reverse Skolem paradox". Namely, if M is a model of set theory, then there is one where we can arrange for ω^M to be uncountable. In other words, not only we can get a model with countably many real numbers, we can also get a model with uncountably many natural numbers. The key point, of course, is that most of the objects that M perceives as natural numbers are not really natural numbers.

9.3 Definability and lack thereof

Definition 9.9. Let M be an $\mathcal{L}$-structure and let $m \in M$. We say that m is *definable* in M if there is some formula $\varphi(x)$ such that $M \models \varphi(a)$ if and only if $m = a$, and we will say that φ *defines* m. If no such φ exists, then we say that m is *undefinable*.

We say that $A \subseteq M^n$ is a *definable relation* (or subset if $n = 1$) when there is a formula $\varphi(\bar{x})$ such that $\langle \bar{m} \rangle \in A$ if and only if $M \models \varphi(\bar{m})$.

Definable elements are, in a sense, elements that can be specified. For example, if c is a constant symbol, then c^M is a definable element. As are any of the other relations and functions in the language. We often, in the mathematical practice, extend the language by adding new symbols for definable elements, definable functions, and definable relations.

We will sometimes talk about definability "with parameters" which means that the formula φ has some additional free variables, e.g. $\varphi(x, x_0)$, and we fix some elements of the structures in their place. So, given any $a \in A$, $\{x \in A \mid x \neq a\}$ is definable with the parameter a, even if a itself is not definable.

Exercise 9.8. Suppose that a is definable in some M. If A is definable with a as a parameter, then A is definable. (Visit solution)

Exercise 9.9. m is definable in M if and only if $\{m\}$ is a definable set.

Exercise 9.10. If A, B are definable in M, then $A \cup B$, $A \cap B$, and $M \setminus A$ are all definable as well. (Visit solution)

Remark

There is an important point to make about the past few exercises. The idea is that if an element m, or a relation $R \subseteq M^n$, is definable, then we can "expand the language and use it". Much like how $\varnothing$ is definable in the language of set theory as $\forall y(y \notin x)$, or the power set function $\mathcal{P}(a)$ is a definable function. We will not formalise this idea in this course, but it is one of the most important things that we can understand about it.

We can also think of definable elements or objects "from a theory". Namely, if $\varphi(x)$ is a formula in the language of T, e.g. the axioms of set theory, can we prove from T that there exists exactly one object satisfying φ? Consider the definition above for the empty set. In the language of a binary relation, the definition we gave above is simply that of a minimal element. But we can prove from the axioms we have that exactly one object will satisfy this definition.

Proposition 9.10. *Every $n < \omega$ is definable in $\langle \omega, < \rangle$.*

Proof. Let us denote by $\sigma(x, y)$ the formula stating that x is the unique successor of y. Let $\varphi_0(x)$ be the formula stating that x is the minimum element, then φ_0 defines 0. If φ_n defines n, then $\varphi_{n+1}(x)$ is, for example, $\exists y(\varphi_n(y) \wedge \sigma(x, y))$. By induction, we get that every element of ω is definable. □

In general, if a structure satisfies that every single element is definable in the structure, we say that it is *pointwise definable*.

Proposition 9.11. *Given a countable language $\mathcal{L}$ such that ω is an $\mathcal{L}$-structure, there exists some $A \subseteq \omega$ which is not definable.*

Proof. $\text{Form}_{\mathcal{L}}$ is a countable set, so there can be at most countably many subsets which are definable, in contrast, $\mathcal{P}(\omega)$ is uncountable. □

Proposition 9.12. *If M is an $\mathcal{L}$-structure, then there is an expansion of the language, $\mathcal{L}^+$ such that M is an $\mathcal{L}^+$-structure and every $m \in M$ is definable in $\mathcal{L}^+$.*

Proof. Let $\mathcal{L}^+$ be $\mathcal{L}$ with new constant symbols, $\{c_m \mid m \in M\}$ and let M be an $\mathcal{L}^+$-structure where the new constant symbols are interpreted by $c_m^M = m$ for all $m \in M$. Now every element of M is definable by $x = c_m$. □

Theorem 9.13. *Suppose that $M \prec N$ and $m \in N$ is definable, then $m \in M$.*

Proof. If m is definable, then there is some $\varphi(x)$ which defines m in N. Therefore $N \models \exists x \varphi(x)$, so by elementarity, $M \models \exists x \varphi(x)$. Suppose that $m' \in M$ such that $M \models \varphi(m')$, then by elementarity, again, $N \models \varphi(m')$, but since φ defines m in N, $m = m'$. □

Corollary 9.14. *Suppose that $N = M^I/\mathcal{U}$ is an ultrapower of M. If $[f]_{\mathcal{U}}$ is definable in N, then for some $m \in M$, $j(m) = [f]_{\mathcal{U}}$.*

Proof. We previously saw in an exercise that $j[M] \prec N$, so $[f]_{\mathcal{U}}$ must be in $j[M]$. □

Exercise 9.11. Show that the standard order is definable in $\langle \omega, + \rangle$. (Visit solution)

Proposition 9.15. *Let $\mathcal{L}$ be a language including a binary relation symbol $<$ and suppose that ω is an $\mathcal{L}$-structure with $<$ interpreted as the standard order. Suppose that $A \subseteq \omega$ and A is definable. Then $\min A$ is a definable element. Therefore, any elementary extension of ω must satisfy the same, that is, if A is definable, then $\min A$ exists and it is definable as well.*

Proof. If $\varphi(x)$ defines A, then $\min A$ is defined by $\varphi(x) \land \forall y(\varphi(y) \to x < y \lor x = y)$, so for any formula $\varphi(x)$,

$$\omega \models \exists x\varphi(x) \to (\exists x(\varphi(x) \land \forall y(\varphi(y) \to x < y \lor x = y)).$$

If N is an elementary extension of ω and $A \subseteq N$ is definable and non-empty, then $\min A$ is definable in N. □

Definition 9.16. Let M be an $\mathcal{L}$-structure. We say that $f\colon M \to M$ is an *automorphism (of $\mathcal{L}$-structures)* if it is an isomorphism of $\mathcal{L}$-structures from M to itself.

Exercise 9.12. Suppose that $M \models \varphi(x)$ and f is an automorphism of M, then $M \models \varphi(f(x))$.

Proposition 9.17. *The standard order is not definable in $\langle \mathbb{Z}, + \rangle$.*

Proof. Consider the function $f\colon \mathbb{Z} \to \mathbb{Z}$ given by $f(x) = -x$, then f is an isomorphism of $\langle \mathbb{Z}, + \rangle$, since $f(x + y) = -(x + y) = (-x) + (-y) = f(x) + f(y)$. However, $x < y$ if and only if $f(y) < f(x)$, so the order cannot be definable. □

Exercise 9.13. Addition is not definable in $\langle \mathbb{Z}, < \rangle$ where $<$ is the standard order. (Visit solution)

Proposition 9.18. *Suppose that $\mathcal{L}$ is a language which includes a binary relation symbol $<$, and suppose that ω is an $\mathcal{L}$-structure where $<$ is the standard order. If N is an elementary extension of ω as a $\mathcal{L}$-structure, then ω is not definable in N.*

Proof. We saw that any proper extension of ω is not well-ordered. And suppose that $\omega \prec N$ and $\varphi(x)$ was a formula which defined ω. In that case, $\neg\varphi$ defines $N \setminus \omega$. However, by Theorem 9.15 we get that $\min(N \setminus \omega)$ is definable. This is impossible: if $m = \min N \setminus \omega$, since ω satisfies that any non-zero element is a successor, by elementarity N must satisfy the same. So m must be the successor of some n, but by its minimality, $n \in \omega$, which is a contradiction since the successor of n must be in ω as well, by elementarity. □

9.4 More uses for automorphisms

Theorem 9.19. *Suppose that $A \subseteq B$ are two infinite sets, then as structures in the empty language, $A \prec B$.*

Proof. Since the language is empty, A is already a substructure of B. Let us check the elementarity condition.

We prove by induction on the complexity of φ that $A \models \varphi(\bar{a})$ if and only if $B \models \varphi(\bar{a})$ whenever $\bar{a} \in A$. If φ is atomic, then it has the form $x_i = x_j$ in which case this is trivial. In the case that φ is $\varphi_0 * \varphi_1$ for some connective $*$ or $\neg\varphi_0$, this is tantamount to checking the truth tables as well.

Finally, if φ is $\exists x\psi(x, \bar{x})$, then if $A \models \exists x\psi(x, \bar{a})$ then there is some $a \in A$ such that $A \models \psi(a, \bar{a})$ and since $A \subseteq B$, by the induction hypothesis, $B \models \psi(a, \bar{a})$, so $B \models \exists x\psi(x, \bar{a})$. In the other direction, suppose that $B \models \exists x\psi(x, \bar{a})$, let $b \in B$ be such that $B \models \psi(b, \bar{a})$, if $b \in A$

we are done by the induction hypothesis on ψ. Suppose that $b \notin A$, since A is infinite, let $a \in A$ be such that $a \neq a_i$ for all $i < n$. Consider the bijection $f\colon B \to B$ given by

$$f(x) = \begin{cases} a & \text{if } x = b, \\ b & \text{if } x = a, \\ x & \text{otherwise.} \end{cases}$$

Since f is a bijection, it is an automorphism of B as a structure in the empty language. Therefore $B \models \psi(f(b), f(\bar{a}))$, but since a, b are both not in $\bar{a}$, $f(\bar{a}) = \bar{a}$ and $f(b) = a$. So, in other words, $B \models \psi(a, \bar{a})$ and by the induction hypothesis we are done. The case of $\forall x\psi$ follows, as $\forall x\psi$ is equivalent to $\neg\exists x\neg\psi$. □

Remark

Note that the above proof is, in effect, a proof of an instance of Theorem 8.5.

Exercise 9.14. Let $\mathcal{L}$ be the language with constant symbols $\{c_n \mid n < \omega\}$ and no other symbol. Let $\omega + 1$ be an $\mathcal{L}$-structure where c_n is interpreted as n. Show that ω is the unique undefinable element in the structure. (Visit solution)

Exercise 9.15. Working in $\langle \mathbb{Q}, < \rangle$ let us fix some $a_0 < \cdots < a_{n-1}$. The type of $a \in \mathbb{Q}$ with respect to the a_i is $\{a_i \in \{a_0, \ldots, a_{n-1}\} \mid a_i < a\}$. Show that for $a, b \notin \{a_0, \ldots, a_{n-1}\}$, a and b have the same type with respect to the a_i if and only if there is an automorphism $f\colon \mathbb{Q} \to \mathbb{Q}$ such that $f(a) = b$ and $f(a_i) = a_i$ for all $i < n$.

Remark

Exchanging two given points in a structure is the key concept of homogeneity, which features prominently in model theory. The idea is that any points, or a tuple, which satisfy exactly the same formulas, can be exchanged.

This allows us to repeat proofs similar to one above. We can study the theory of $\langle \mathbb{Q}, < \rangle$, "Dense Linear Order", whose axioms are that $<$ is a strict linear order, without minimum or maximum, between any two points lies a third, and there are at least two points. It turns out that these axioms are enough to guarantee that the proof of the exercise go through.

Chapter 10

Where the Moose and the Lions dread to step

Chapter Goals

In this chapter we will learn about

- Second-order logic, infinitary logic, and additional quantifiers we can add or change.
- How all of these extended logics fail either the Löwenheim–Skolem theorems, Łoś's Theorem, or the Compactness Theorem.
- How various concepts can be formalised in these generalised logics, even if they cannot be formalised in first-order logic. For example, well-orders.

10.1 Second-order logic

Definition 10.1. *Second-order logic* is the logic given by adding to first-order logic a new type of variables which represents n-ary relations. The semantics of second-order logic are defined the same as before. We will use capital letter variables A, R, F, etc. to denote the second-order variables. To improve the readability of our second-order formulas we will often not specify the arity of a second-order variable explicitly, we will write $x \in A$ to mean $A(x)$ in the case of a unary predicate, and we will omit the part of a formula stating that F is a function, when quantifying over functions.

Theorem 10.2. *There is a second-order sentence in the empty language, whose models are exactly all the finite sets. Therefore Łoś's Theorem fails for second-order logic.*

Proof. Let φ be the following sentence:

$$\begin{aligned}
&\exists R(\\
&\quad \forall x \neg R(x,x) \wedge \forall x \forall y \forall z (R(x,y) \wedge R(y,z) \to R(x,z)) \wedge \forall x \forall y (R(x,y) \vee R(y,x) \vee x = y) \wedge \\
&\quad \forall A(\exists x (x \in A) \to \exists y \exists z (y \in A \wedge z \in A \wedge \forall x (x \in A \to y < x \vee x = y \wedge x < z \vee x = z))) \\
&\quad)
\end{aligned}$$

We saw that a set is finite if and only if it has a well-ordering whose inverse is also a well-ordering, this is to say that every non-empty set has a minimum and a maximum. Therefore

Theorem 8.10 fails, and therefore Łoś's Theorem must fail too, as φ has models of arbitrarily large finite size, but no infinite models. □

Remark

We can replace the statement by "Every relation is well-founded" or "Every injection is a bijection" just as well. But in both of these cases we need to use the Axiom of Choice to prove the claim. The above, however, works even without it.

Exercise 10.1. Write a second-order axiom saying that every subset of the universe satisfies φ or has a bijection with the universe itself. Find all the models of this axiom.

Theorem 10.3. *Let T be the second-order theory in the language $\{<\}$ containing a single binary relation stating that:*

1. $\forall x \neg(x < x) \wedge \forall x \forall y \forall z (x < y \wedge y < z \to x < z) \wedge \forall x \forall y (x < y \vee y < x \vee x = y)$.
2. $\forall x (\forall y (x < y \vee x = y) \vee \exists y (y < x \wedge \forall z (y < z \to x < z \vee x = z))) \wedge \forall x \exists y (x < y)$.
3. $\forall A (\exists x (x \in A) \to \exists x (x \in A \wedge \forall y (y \in A \to x < y \vee x = y)))$.

Then T has an infinite model, and any other infinite model is isomorphic to it. In particular, the Compactness Theorem fails for second-order logic.

Proof. The theory T states that $<$ is a well-ordered set without a maximum where every point is the minimum or a successor. We saw that any such well-ordered set is isomorphic to ω, therefore T has an infinite model, and it must be unique up to isomorphism. If the Compactness Theorem held for second-order logic, then Theorem 8.17 would be imply that T has arbitrarily large infinite models, and so Compactness must fail. □

Exercise 10.2. Write second-order theories that characterise $\omega + \omega$ and $\omega \cdot \omega$ in the language $\{<\}$. (Visit solution)

Theorem 10.4. *Let T be the second-order theory in the language of ordered fields, $\langle 0, 1, +, \cdot, < \rangle$ which includes the axioms of an ordered field, as well as the second order axiom*

$$\forall A (\exists a (a \in A) \wedge \exists x \varphi(A, x) \to \exists x (\varphi(A, x) \wedge \forall y (\varphi(A, y) \to x < y \vee x = y))),$$

where $\varphi(A, x)$ is the formula $\forall a (a \in A \to a < x \vee a = x)$. In other words, every non-empty bounded set has a supremum.

Then $\mathbb{R}$ is a model of T with its standard ordered field structure, and moreover any other model of T is isomorphic to $\mathbb{R}$, therefore the Downward Löwenheim–Skolem Theorem fails for second-order logic.

Proof. $\mathbb{R}$ is a model of T as it is an ordered field. Recall that we showed that the collection of non-empty and bounded initial segments of $\mathbb{Q}$ is a model of $\mathbb{R}$, working with that interpretation, if $A \neq \varnothing$ is a bounded subset of $\mathbb{R}$, then $\bigcup A = \sup A$ as it is an initial segment of $\mathbb{Q}$, it is bounded and non-empty, and it is not hard to check that it is indeed the supremum of A in $\mathbb{R}$.

On the other hand, if F is a model of T, then it is a field such that for every $n < \omega$, the n-fold addition $1 + \cdots + 1 \neq 0$. Therefore there is an embedding $e \colon \mathbb{Q} \to F$. We can now extend this embedding to $e^+ \colon \mathbb{R} \to F$ by setting $e^+(r) = \sup\{e(q) \mid q < r\}$, this is well-defined since in both F and $\mathbb{R}$ every bounded set has a supremum.

It remains to show that e^+ is an isomorphism. Suppose not, if we can find some $u \in F$ such that u is an upper bound of $\operatorname{rng}(e)$, then by the second-order axiom F is satisfying, there is some $z \in F$ such that $z = \sup \operatorname{rng}(e)$, in which case $z - 1 < e(q)$ for some $q \in \mathbb{Q}$, and therefore $z < e(q) + 1 = e(q+1)$ which is a contradiction. Suppose that $u \in F$ and $u \notin \operatorname{rng}(e^+)$, without loss of generality we may assume that $u > 0$, otherwise take $-u$. If u is an upper bound for $\operatorname{rng}(e)$ or $\operatorname{rng}(e^+)$, then we are done. Otherwise, let $r = \sup\{q \in \mathbb{Q} \mid e(q) < u\}$ and let $u' = u - e^+(r)$, since $e^+(r) = \sup\{e(q) \mid e(q) < u\}$ it follows that $u' > 0$, but there is no $q \in \mathbb{Q}$ such that $0 < e(q) < u'$, since otherwise $e^+(r) < e(q) < u$. Therefore $\frac{1}{u'}$ is an upper bound of $e(q)$, but that is impossible.

Therefore the Downward Löwenheim–Skolem Theorem must fail for second-order logic, since T has an infinite model, but it does not have a countably infinite model despite being a theory in a finite language. □

Exercise 10.3. Write a theory T in the language $\{+, <\}$ which characterises $\langle \mathbb{Z}, +, < \rangle$.

Remark

In both cases we could have easily worked in the empty language. Specifically, we simply need to posit that there exists a binary relation, or there exist $+$ and $\cdot$ and $<$, etc. Of course, allowing these to be part of the language makes the whole thing easier to state and explain, but the power of second-order logic is quite far-reaching in that we can conjure finite first-order languages out of the ether.

We can define stronger logics, such as third-order, fourth-order, and generally nth-order logics by allowing to quantify over sets of relations, etc. Each iteration puts us farther and farther away from first-order logic. We can also define weaker, intermediate logics, for example monadic second-order logic allows us to quantify over subsets, but not over relations in general.

However, set theory as a foundation of mathematics helps us recover from this issue. If we are working inside a universe of set theory, then all of the quantification is now a first-order quantifier on subsets of some set. Namely, "for every subset of $\mathbb{R}$" is simply quantifying over the elements of $\mathcal{P}(\mathbb{R})$, which is just a first-order quantifier as far as set theory is concerned. This is one the benefits of working in a set theoretic context.

Exercise 10.4. Write a single second-order axiom in the empty language whose models are exactly of size $\aleph_1$. (Hint: Use Theorem 10.3 to characterise all the subsets which are infinite and do not surject onto the whole structure.) (Visit solution)

10.2 Infinitary logic

Definition 10.5. Let κ and λ be infinite cardinals. The logic $\mathcal{L}_{\kappa,\lambda}$ is obtained by modifying first-order logic in the following way: we increase the set of free variables to $\{x_\alpha \mid \alpha < \lambda\}$. Next, we take the closure of the atomic formulas under $\wedge, \vee$ of any collection of fewer than κ many formulas, and allow quantifying over fewer than λ variables (as a single operation).

For example, $\mathcal{L}_{\omega_1,\omega}$ is obtained by allowing countable conjunctions and disjunctions, but only finitely many quantifiers at a time.

Theorem 10.6. *There is an $\mathcal{L}_{\omega_1,\omega}$ sentence φ in the language $\{<\}$ such that the only model of φ is ω.*

Proof. We saw that for every $n < \omega$ there is a first-order formula $\varphi_n(x)$ such that $\omega \models \varphi_n(x)$ if and only if $x = n$. Consider the sentence φ which is the conjunction of the axioms stating that $<$ is a strict linear order without a maximum and every point is the minimum or a succesor, with the infinitary sentence $\forall x \bigvee_{n<\omega} \varphi_n(x)$. □

It is not hard to see that if M is a pointwise definable model (of a first-order theory T) of size κ, then M can be fully characterised in $\mathcal{L}_{\kappa,\omega}$ by a similar argument as before.

Theorem 10.7. *The logic $\mathcal{L}_{\omega_1,\omega}$ is not compact.*

Proof. Consider the language $\{<\} \cup \{c_\alpha \mid \alpha < \omega_1\}$ and the sentence φ from Theorem 10.6, and let T be the theory $\{\varphi\} \cup \{c_\alpha \neq c_\beta \mid \alpha < \beta < \omega_1\}$. If $T_0 \subseteq T$ is a finite subset, then there are only finitely many axioms of the form $c_\alpha \neq c_\beta$, so we can interpret those finitely many constants as distinct natural numbers. However, T cannot possibly be consistent since any model of φ must be countable and any model of T must be uncountable. □

Exercise 10.5. Show that in the above proof we can use a single constant symbol instead of uncountably many constants to get that $\mathcal{L}_{\omega_1,\omega}$ is not compact. Consider your theory and find a finite theory which is equivalent to it and is inconsistent. (Visit solution)

The truth is that for some κ, $\mathcal{L}_{\kappa,\omega}$ and $\mathcal{L}_{\kappa,\kappa}$ can be somewhat compact. Here "somewhat compactness" means that if T is such that for all $S \subseteq T$, if $|S| < \kappa$, then S has a model, then T has a model as well. The existence of a cardinal κ for which $\mathcal{L}_{\kappa,\kappa}$ is somewhat compact is not provable from the axioms of set theory that we have seen. Indeed, we can even add more information by defining varying degrees of compactness through requirements such as the size of T being limited, or the existence of Łoś's Theorem for $\mathcal{L}_{\kappa,\kappa}$, etc.

Interestingly enough, the Downward Löwenheim–Skolem Theorem holds for $\mathcal{L}_{\omega_1,\omega}$, but the Upward fails, as the above clearly shows.

Theorem 10.8. *There is an $\mathcal{L}_{\omega_1,\omega_1}$-sentence in the language $\{<\}$ which characterises well-ordered sets.*

Proof. Let φ be the conjunction of the sentence that $<$ is a strict linear order along with

$$\forall x_0 \dots \forall x_n \dots \left(\bigwedge_{n<\omega} (x_{n+1} < x_n \vee x_{n+1} = x_n) \right) \to \bigvee_{n<\omega} \bigwedge_{m \geq n} x_m = x_n.$$

In other words, the statement says that if $x_{n+1} \leq x_n$ for all $n < \omega$, then there is some n such that $x_m = x_n$ for all $m \geq n$.

To see that this characterises well-ordered sets we need Theorem 10.9, which completes the proof. □

Lemma 10.9. *Suppose that $\langle A, < \rangle$ is a strict linearly ordered set. Then A is well-ordered if and only if there is no $f\colon \omega \to A$ such that $n < m$ if and only if $f(m) < f(n)$.*

Proof. Suppose that such a function exists, then $\operatorname{rng}(f)$ is a non-empty subset of A without a minimal element, so A is not well-ordered. In the other direction, let $B \subseteq A$ be a set without a minimal element. We fix a choice function c on $\mathcal{P}(B)$, i.e. for every non-empty subset of B, c selects an element of the set.

Define by recursion, $f(0) = c(B)$ and $f(n+1) = c(\{a \in B \mid a < f(n)\})$. Since B has no minimal element, $f(n+1)$ is well-defined, and by definition $f(n+1) < f(n)$, so by induction if $n < m$, we get $f(m) < f(n)$ as wanted. □

Remark

The use of the Axiom of Choice in the above proof is necessary, and indeed we can prove a similar lemma about any well-founded relation (with the caveat that not requiring the relation to be transitive, we need to require in the lemma that $f(n+1) < f(n)$ rather than $f(m) < f(n)$ for the general case). In the case of well-founded relations this turns out to be equivalent to the generalised recursion theorem for ω, as well as to the Downward Löwenheim–Skolem Theorem for countable languages.

Exercise 10.6. Show that in the language of fields, every rational number is definable (in every field).

Theorem 10.10. *$\langle \mathbb{R}, 0, 1, +, \cdot, < \rangle$ is completely characterisable in $\mathcal{L}_{\omega_1,\omega}$.*

Proof. Let φ_q be the first-order formula defining the rational number q. For each $r \in \mathbb{R}$ we let $\psi_r(x)$ be the $\mathcal{L}_{\omega_1,\omega}$ formula

$$\forall y \left(\bigwedge_{q<r, q\in\mathbb{Q}} \varphi_q(y) \to y < x \wedge \bigwedge_{r<q, q\in\mathbb{Q}} \varphi_q(y) \to x < y \right).$$

Let T be the first-order theory of ordered fields augmented by adding, for each $r \in \mathbb{R}$, the axiom $\exists x(\psi_r(x) \wedge \forall y(\psi_r(y) \to y = x))$. Namely, for each $r \in \mathbb{R}$, there is exactly one object which satisfies $\psi_r(x)$.

We claim that if F is a model of T, then $\mathbb{R} \cong F$. To define the isomorphism, clearly we map each $r \in \mathbb{R}$ to the unique witness for ψ_r in F. Easily, this is an embedding. Let us show that it is surjective. Suppose it was not surjective and let $t \in F$ be such that $F \models \neg\psi_r(t)$ for all $r \in \mathbb{R}$. Without loss of generality, $0 < t$, otherwise $-t$ will also be a solution. And without loss of generality, there is some $q \in \mathbb{Q}$ such that $t < q$, otherwise t is an upper bound to all the rational numbers in F, in which case $\frac{1}{t}$ will be a positive lower bound to all the positive rational numbers, while also satisfying $\neg\psi_r$ for all r. Finally, by considering $r = \sup\{q \in \mathbb{Q} \mid q < t\}$, we get that there are no rational numbers between r and t, and therefore if $q \in \mathbb{Q}$, then $q < r$ if and only if $q < t$ and $r < q$ if and only if $q < t$, therefore $\psi_r(t)$ holds, which is a contradiction to the assumption that $\neg\psi_r(t)$ holds. □

10.3 There exists a quantifier

10.3.1 Definable quantifiers

In some cases we want to have some "syntactic sugar" to make a statement more readable and to mean something that is easily expressible via a complicated statement.

Definition 10.11. We define the quantifier $\exists! x\varphi$ to be "there exists exactly one x for which φ holds". In other words, $M \models \exists! x\varphi(x)$ if and only if there is a unique $m \in M$ such that $M \models \varphi(m)$.

Proposition 10.12. *$\exists! x\varphi(x)$ is equivalent to $\exists x(\varphi(x) \wedge \forall y(\varphi(y) \to y = x))$.*

Proof. Suppose that $M \models \exists! x\varphi(x)$, and let $m \in M$ be the unique element such that $M \models \varphi(m)$. Then $M \models \exists x\varphi(x)$ by the definition of the satisfaction relation, and moreover, $M \models \forall y(\varphi(y) \to y = x)$ by the uniqueness of m.

In the other direction, if $M \models \exists x(\varphi(x) \wedge \forall y(\varphi(y) \to y = x))$, then $M \models \exists x \varphi(x)$, so there is some $m \in M$ such that $M \models \varphi(m)$. Since $M \models \forall y(\varphi(y) \to y = m)$, it follows that m is the unique element satisfying φ, so $M \models \exists! x \varphi(x)$. □

Exercise 10.7. Let $\mathcal{L}$ be the language containing one binary relation symbol $\{R\}$. Write a sentence φ using $\exists!$ such that $M \models \varphi$ if and only if $R^M : M \to M$ is a bijection.

Exercise 10.8. Let $\mathcal{L}$ be the language containing a unary function symbol F. Write a sentence φ using $\exists!$ such that $M \models \varphi$ if and only if $\operatorname{rng}(F^M) = M \setminus \{m\}$ for some $m \in M$.

If we are willing to look at other logics, we can define even more quantifiers, and quantifiers can take multiple variables to quantify over. In $\mathcal{L}_{\omega_1,\omega_1}$ we can define $\mathscr{Q}xy\varphi(x,y)$ by

$$M \models \mathscr{Q}xy\varphi(x,y) \text{ if and only if}$$
$$(\exists x_0 \dots)(\exists y_0 \dots) \bigwedge_{i<j<\omega} x_i \neq x_j \wedge y_i \neq y_j \wedge \bigwedge_{n<\omega} \varphi(x_n, y) \wedge \neg\varphi(x, y_n).$$

10.3.2 Undefinable quantifiers

Sometimes we want to add a quantifier that gives us a greater power of expressibility. In some cases we can show that it cannot be definable, since the logic we obtain by adding the quantifier is not compact, or does not have the Downward Löwenheim–Skolem Theorem to it.

Notation 10.13. If $\mathcal{L}$ is a logic (e.g. first-order logic or $\mathcal{L}_{\kappa,\lambda}$, etc.) and Q is a quantifier we write $\mathcal{L}(Q)$ to denote the logic obtained by adding the quantifier Q to the available quantifiers. In the case of first-order logic we will use $\mathcal{L}(Q)$ to denote the expansion.

Definition 10.14. Let $\mathcal{L}(\forall^\infty)$ be the logic obtained by adding to first-order logic the quantifier $\forall^\infty$ whose interpretation is given by $M \models \forall^\infty x\varphi(x)$ if and only if $\{m \in M \mid M \models \neg\varphi(m)\}$ is a finite set.

Theorem 10.15. *The logic $\mathcal{L}(\forall^\infty)$ is not compact, therefore the quantifier is not definable.*

Proof. Let us consider the sentence φ given by $\exists x \forall^\infty y(x = y)$. Then $M \models \varphi$ if and only if there exists some $m \in M$ such that $\{n \in M \mid n \neq m\}$ is a finite set. In other words, M is finite (and non-empty).

In particular, the sentence characterises non-empty finite sets. This shows that Theorem 8.10 fails for $\mathcal{L}(\forall^\infty)$, and therefore the Compactness fails. □

Exercise 10.9. Show that $\forall^\infty$ is definable in the logic $\mathcal{L}_{\omega_1,\omega}$. (Visit solution)

Theorem 10.16. *We can characterise $\langle \omega, < \rangle$ in $\mathcal{L}(\forall^\infty)$ with the language $\{<\}$.*

Proof. Consider the sentence saying that $<$ is a strict linear order without a maximal element and $\forall x \forall^\infty y(x < y)$, which states that every proper initial segment is finite. If $M \models \varphi$, then $<^M$ is a linear ordering without a maximal whose proper initial segments are finite, note that the size of initial segment determines the point uniquely, which by induction allows us to define an isomorphism with ω. □

Exercise 10.10. Show that $\omega \cdot \omega$ can also be characterised in $\mathcal{L}(\forall^\infty)$ in the language $\{<\}$. (Visit solution)

Exercise 10.11. Show that $\exists^\infty x\varphi$ stating "There are infinitely many x such that φ" is definable in $\mathcal{L}(\forall^\infty)$. Moreover, show that $\forall^\infty$ is definable in $\mathcal{L}(\exists^\infty)$.

This can be, of course, heavily generalised.

Definition 10.17. Let $\mathcal{F}$ be a filter on a set M. Then for any structure on the set M, the universal quantifier for $\mathcal{F}$ is $\forall^{\mathcal{F}}x\varphi(x)$ is interpreted as $\{m \in M \mid M \models \varphi(m)\} \in \mathcal{F}\}$. We will usually consider filters that are defined in the same way for every set, e.g. "all the sets whose complement is finite".

We can also define $\exists^{\mathcal{F}}x\varphi(m)$ by $\{m \in M \mid M \models \neg\varphi(m)\} \notin \mathcal{F}$. If we think of the sets in $\mathcal{F}$ as "almost everything" and their complements as "almost nothing", then $\forall^{\mathcal{F}}x\varphi$ is "almost everything satisfies φ" and $\exists^{\mathcal{F}}x\varphi$ is "the set of points satisfying φ is not nothing".

Exercise 10.12. Find filters whose universal quantifiers are $\forall$ and $\forall^\infty$. (Visit solution)

Definition 10.18. Let $Q_1x\varphi$ be the quantifier such that $\{m \in M \mid M \models \varphi(m)\}$ is uncountable.

Exercise 10.13. Show that Q_1 is definable in $\mathcal{L}_{\omega_1,\omega_1}$.

Theorem 10.19. *$\mathcal{L}(Q_1)$ is not compact.*

Proof. Consider the sentence φ in $\mathcal{L}(Q_1)$, $\forall x\neg Q_1y(x \neq y)$. If $M \models \varphi$, then for every $m \in M$, the set $M\setminus\{m\}$ is not uncountable. Therefore M must be countable, so Theorem 8.17 fails. $\square$

Exercise 10.14. Is there a filter such that Q_1 is $\forall^{\mathcal{F}}$ or $\exists^{\mathcal{F}}$?

Remark

We have discussed different types of logic here. Higher-order logics, infinitary logics, additional quantifiers. These, of course, can be mixed: We can consider the second-order $\mathcal{L}_{\omega_1,\omega_1}(Q)$ as our logic and use that to study and discuss all kind of structures, we can even allow more than two truth values. What we did not do was to define explicitly what is a logic. This definition goes beyond the scope of this course, but very informally, logic has syntax and semantics, and it has a way to connect the two.

First-order logic, as we have seen, has two wonderful properties: the Compactness Theorem, which allows us to amalgamate finitary pieces and "go up" in the size of models, and the Downward Löwenheim–Skolem theorem, which allows us to carve out small substructures and "go down" in the size of our models. It turns out that these two almost fully characterise first-order logic.

Per Lindström proved that any logic that has double negation elimination (that is, $\neg\neg\varphi \to \varphi$) and satisfies both the Compactness Theorem and the Downward Löwenheim–Skolem theorem is, essentially, first-order logic. This is why when we strengthen our logics, as we did above, these two properties break down.

Chapter 11

You and V

Chapter Goals

In this chapter we will learn about

- The Axiom of Foundations (completing the axioms of ZF and ZFC), as well as its equivalents such as $\in$-induction.
- The von Neumann hierarchy and the class of well-founded sets.
- The Reflection theorem.
- Gödel's constructible universe, L.

11.1 Axiom of Foundation

11.1.1 The formal language of set theory

The language of set theory is simply $\{\in\}$ where $\in$ is a binary relation symbol. The axioms of set theory govern what kind of properties $\in$ satisfies and models of this theory are models of set theory, and their elements are called sets. This may seem circular, and to a naive extent it almost is. But foundations of mathematics need to be taken from somewhere, and we can do that in many different ways, but one way is to simply presuppose the existence of a universe of sets and work within it. There, we may want to study set theory as a mathematical theory, and that means that we need to formulate it in a particular language and study its structures. There are other philosophical ways out of this problem, but at the end of the day, any of them that "resolve the problem" will rely on the dogmatic (rather than axiomatic) assumption that something exists. We can now understand the term "property" which we used before to mean "first-order formula in the language of set theory".

As a form of syntactic sugar we add to the language of set theory constant symbols such as $\varnothing$ and ω, operations such as $\bigcup$ or $\cap$ or $\mathcal{P}$, or relations such as $\subseteq$ and Ord. The axioms we have so far make almost a complete list of the axioms of modern set theory, with only one missing.

11.1.2 The Axiom of Foundation

Axiom: Foundation

For every non-empty set x there is some $y \in x$ such that $x \cap y = \varnothing$.

Remark

In some places, the Axiom of Foundation is called the Axiom of Regularity.

Proposition 11.1. *Assuming the Axiom of Foundation, for every set a, $a \notin a$.*

Proof. Let a be a set and let $x = \{a\}$. By the Axiom of Foundation, there is some $y \in x$ such that $y \cap x = \varnothing$, but since the only possible value for y is a itself is means that $a \cap x = \varnothing$, and therefore $a \notin a$. □

Exercise 11.1. If $x \in y$, then $y \notin x$.

Exercise 11.2. For any $x_0, \dots, x_n$, if $x_{i+1} \in x_i$ for $i < n$ (that is, $x_n \in \cdots \in x_0$), then $x_0 \notin x_n$. (Visit solution)

Proposition 11.2. *Suppose that f is a function whose domain is ω, then there is some n such that $f(n+1) \notin f(n)$.*

Proof. Let $x = \operatorname{rng} f = \{f(n) \mid n < \omega\}$. By the Axiom of Foundation, there is some $y \in x$ such that $y \cap x = \varnothing$. By the way we defined x, $y = f(n)$ for some n, and since $y \cap x = \varnothing$, then in particular $f(n+1) \notin f(n)$. □

Exercise 11.3. Assuming the Axiom of Foundation, show that $\langle x, y\rangle = \{x, \{x, y\}\}$ is a valid interpretation for ordered pairs. (Visit solution)

Definition 11.3. We say that a set A is *well-founded* if $\in$ is a well-founded relation on A. That is, if $\{\langle a, b\rangle \in A \times A \mid a \in b\}$ is a well-founded relation on A.

Exercise 11.4. If A is well-founded and $B \subseteq A$, then B is well-founded.

Exercise 11.5. The following are equivalent:

1. The Axiom of Foundation.
2. Every set is well-founded.
3. Every transitive set is well-founded.
4. Every set is a subset of a well-founded set. (Visit solution)

Proposition 11.4. *Assuming the Axiom of Foundation, x is an ordinal if and only if x is a transitive set of transitive sets.*

Proof. We saw that an ordinal is itself a transitive set of ordinals, so the implication from "x is an ordinal" to "x is a transitive set of transitive sets" is trivial. In the other direction, suppose that x is a transitive set of transitive sets. As the Axiom of Foundation is assumed, $\in$ is a well-founded relation on x, and by Theorem 11.1, $\in$ is irreflexive. Moreover, every element of x is itself transitive, if $a \in b$ and $b \in c$, then $b \subseteq c$, so $a \in c$. Therefore $\langle x, \in\rangle$ is a well-founded strict partial order. It remains to show that it is a total order.

Suppose not. Then $\{a \in x \mid a \text{ is incomparable with some } b \in x\}$ is non-empty, so by the Axiom of Foundation it has a minimal element a, then $\{b \in x \mid b \notin a \wedge a \notin b\}$ is non-empty, so the set has a minimal element as well, b. Since $a, b \in x$, it follows that $a, b \subseteq x$. However, by minimality, if $c \in a$, then $c \in b$ or $b \in c$. It is impossible that $b \in c$, since by transitivity $b \in a$ and the two are incomparable. So for every $c \in a$, $c \in b$. Similarly, if $c \in b$, then by the minimality of b, $c \in a$ or $a \in c$. The latter is impossible, as transitivity implies $a \in b$, so $c \in a$. Therefore we get that $a \subseteq b$ and $b \subseteq a$, and so $a = b$. This is a contradiction, since a and b were incomparable. So, it must be that x is well-ordered by $\in$, so it is an ordinal. □

Remark

It is certainly possible without the Axiom of Foundation that $x = \{x\}$, in that case it follows that x is a transitive set, and that indeed all of its elements are transitive sets, but x is not an ordinal, since it is not well-ordered by $\in$. We can even have a situation where $x = \{x_n \mid n < \omega\}$ where $x_n = \{x_k \in x \mid k > n\}$, so each of the x_n is a transitive set, but x does not have a minimal element, so it is, again, not an ordinal.

Proposition 11.5. *If T is a well-founded transitive set, then $T = \varnothing$ or $\varnothing \in T$.*

Proof. Suppose that $T \neq \varnothing$. Then there is some $y \in T$ such that $y \cap T = \varnothing$. Since T is transitive, $y \subseteq T$, so $y = y \cap T = \varnothing$. □

11.1.3 Zermelo–Fraenkel axioms

We now have a complete list of the Zermelo–Fraenkel axioms of set theory, or ZF:

1. Axiom of Extensionality
2. Axiom of Union
3. Axiom of Power Set
4. Axiom of Infinity
5. Axiom of Replacement
6. Axiom of Foundation

If we add the Axiom of Choice we get ZFC, which is the de-facto "background theory of mathematics". Note that Replacement, now that we understand the formal language of set theory, is in fact an infinite collection of axioms: for every formula φ we add an axiom that states that if φ defines a function, then the image of a set is a set. The same goes for Separation. If we use second-order logic, however, we can formulate these as a single second-order axiom instead.

Exercise 11.6. Prove the Axiom of Separation and Axiom of Pairing from ZF.

11.2 The universe as we know it

In this section we will assume all the axioms of ZF except the Axiom of Foundation. Note that this means that we are also not going to rely on the Axiom of Choice.

Definition 11.6. The *von Neumann hierarchy* is defined by recursion on Ord:

1. $V_0 = \varnothing$.
2. $V_{\alpha+1} = \mathcal{P}(V_\alpha)$.
3. $V_\alpha = \bigcup\{V_\beta \mid \beta < \alpha\}$ whenever α is a limit ordinal.

We denote by V the class $\bigcup\{V_\alpha \mid \alpha \in \mathrm{Ord}\}$, or equivalently $\{x \mid \exists\alpha \in \mathrm{Ord}, x \in V_\alpha\}$.

Exercise 11.7. For all α, V_α is a transitive set, and if $x \in V_\alpha$ and $y \subseteq x$, then $y \in V_\alpha$.

Exercise 11.8. For all α, $\alpha \subseteq V_\alpha$.

Proposition 11.7. *For all α, V_α is well-founded.*

Proof. Suppose that V_β is well-founded for all $\beta < \alpha$. Let $x \subseteq V_\alpha$ be a non-empty set, we want to find an $\in$-minimal element in x. Namely, some $y \in x$ such that $x \cap y = \varnothing$. Note that by definition, if $y \in V_\alpha$, then there is some β such that $y \subseteq V_\beta$. Pick $y \in x$ such that the least β for which $y \subseteq V_\beta$ is minimal. We claim that $y \cap x = \varnothing$. Otherwise, pick any $z \in y \cap x$, then $z \in V_\beta$, so there is some $\gamma < \beta$ for which $z \subseteq V_\gamma$, in contradiction to the minimality of β. □

Theorem 11.8. *For all x, $x \in V_\alpha$ for some α, if and only if there is a well-founded transitive set which contains x.*

Proof. One direction is trivial, since if $x \in V_\alpha$, then $x \subseteq V_\alpha$, which is a well-founded transitive set. In the other direction, let us assume without loss of generality that x is already transitive. Let $x_\alpha = x \cap V_\alpha$, for all α. Then there is a formula $\theta(y, \alpha)$ which holds for $y \in x$ and the least α such that $y \in x_\alpha$, or with $\alpha = 0$ if no such α exists. By the Axiom of Replacement, $A = \{\beta \mid \exists y \in X, \theta(y, \beta)\}$ is a set, and so we can take $\alpha = \sup A$. If $x \neq x_\alpha$, then $x \setminus x_\alpha \neq \varnothing$, so by well-foundedness, there is a $y \in x \setminus x_\alpha$ which is $\in$-minimal, that is $y \cap (x \setminus x_\alpha) = \varnothing$. Since x is a transitive set, $y \subseteq x$, and therefore we get that $y \subseteq x_\alpha$, as all of its elements are not in $x \setminus x_\alpha$. In particular, $y \subseteq V_\alpha$, so $y \in V_{\alpha+1}$. This means that $\theta(y, \beta)$ must hold for some $\beta \leq \alpha$ after all, and so $y \in V_\alpha$. Therefore $y \in x \cap V_\alpha = x_\alpha$, in contradiction to the assumption that $y \in x \setminus x_\alpha$. Therefore $x = x_\alpha$, so $x \subseteq V_\alpha$ and therefore $x \in V_{\alpha+1}$. □

Theorem 11.9. *The Axiom of Foundation holds if and only if V is the class of all sets.*

Proof. If the Axiom of Foundation holds, then every transitive set is well-found and by Theorem 11.8 we get that every transitive set is in some V_α. By Theorem 4.24, every set is contained in a transitive set, so we get that every set is in V. In the other direction, if V is the class of all sets, then every set is a subset of some V_α, and therefore every set is well-founded, so the Axiom of Foundation holds. □

Remark

We can actually say much more about V itself. We can show that every axiom of ZF holds inside V. This is nearly trivial for all of them except Replacement, where we need to talk about relativisation and consider only formulas which define functions from V to itself. But the result is that if ZF without the Axiom of Foundation has a model, then ZF has a model.

Definition 11.10. Suppose that $x \in V$ the *(von Neumann) rank* of x is the least α such that $x \subseteq V_\alpha$, and we write $\mathrm{rank}(x) = \alpha$ in that case.

Exercise 11.9. Prove that $\mathrm{rank}(x) = \sup\{\mathrm{rank}(y) + 1 \mid y \in x\}$. (Visit solution)

Exercise 11.10. For every α, $\operatorname{rank}(\alpha) = \alpha$.

Exercise 11.11. Suppose that $\operatorname{rank}(x) = \alpha$ and $\operatorname{rank}(x) = \beta$. Compute the rank of $\langle x, y \rangle$. (Visit solution)

Exercise 11.12. Compute $\operatorname{rank}(\omega_1^{\omega_6})$ and $\operatorname{rank}(\omega_6^{\omega_1})$. You may use the fact that if $f : \omega_1 \to \omega_6$, then there is some $\alpha < \omega_6$ such that $\operatorname{rng}(f) \subseteq \alpha$. (Visit solution)

Exercise 11.13. Compute $\operatorname{rank}(V_\omega \times V_\omega)$.

Theorem 11.11 ($\in$-Induction). *The following are equivalent:*

1. *The Axiom of Foundation.*
2. *For every $\varphi(x)$, $\forall x(\forall y(y \in x \to \varphi(y) \to \varphi(x)) \to \forall x \varphi(x))$.*

Proof. Assume the Axiom of Foundation, and suppose that φ is a formula as in the assumption of (2). We prove φ holds by induction on the rank of x. Suppose that $\operatorname{rank}(x) = \alpha$ and whenever $\operatorname{rank}(y) < \alpha$, then $\varphi(y)$ holds. Since $x \subseteq V_\alpha$, if $y \in x$, then there is some $\beta < \alpha$ such that $y \subseteq V_\beta$. Therefore, $\operatorname{rank}(y) < \alpha$, so $\varphi(y)$ holds. Therefore, $\forall y(y \in x \to \varphi(y))$ is a true statement, and therefore $\varphi(x)$ holds as well. So, $\forall x \varphi(x)$ holds.

In the other direction, let us prove that for every x, $x \in V_\alpha$ for some α, by Theorem 4.15 we can indeed express that by a first-order formula, $\varphi(x)$. By assuming (2), what we need to show is that $\forall y(y \in x \to \varphi(y)) \to \varphi(x)$ holds, and the conclusion will be that φ holds for all x, as wanted. And indeed, given x, if for every $y \in x$, $y \in V_\beta$ for some β, by the Axiom of Replacement $\{\beta + 1 \mid \text{For some } y \in x, \operatorname{rank}(y) = \beta\}$ is a set, and letting α be the supremum of this set, we get that $x \subseteq V_\alpha$, so $x \in V_{\alpha+1}$. □

Remark

In weaker foundational settings, e.g. constructive mathematics or weak set theories, often the Axiom of Foundation is given by the $\in$-induction formulation.

The following is an important feature of the von Neumann hierarchy. We will not prove this theorem.

Theorem 11.12 (Reflection). *Suppose that $\bar{x}$ is a tuple of sets and $\varphi(\bar{x})$ holds. Then for every α there is some $\beta > \alpha$ such that $\bar{x} \in V_\beta$ and $V_\beta \models \varphi(\bar{x})$.*

The Reflection Theorem tells us that anything that happens in the universe is reflected in some initial segment. It turns out that the Reflection Theorem is in fact equivalent to the Replacement Axiom. We prove the Reflection theorem for any hierarchy which is similar to the von Neumann hierarchy in that it (1) captures all the sets in the universe; and (2) is defined by unions at limit stages.

Proposition 11.13. *V_ω is countable.*

Proof. We define by recursion $f : V_\omega \to \omega$ by defining its restriction, f_n, to each V_n. We begin with $f_0 : V_0 \to \omega$ being the empty function. If f_n is defined, we let $f_{n+1}(x) = \sum_{y \in x} 2^{f_n(y)}$. This can be also be defined by Theorem 3.16 on the structure $\langle V_\omega, \in \rangle$, as it is well-founded.

First, we claim that if $n < m$, and $x \in V_n$, then $f_n(x) = f_m(x)$. We prove the claim by induction on n. Suppose that the claim holds for all $k < n$, and let $x \in V_n$, note that if $n = 0$

there is no such x and the claim is vacuous. If for some $k < n$, $x \in V_k$, then $f_k(x) = f_n(x)$ by the induction hypothesis, so in particular for all $m > n$, $f_k(x) = f_n(x) = f_m(x)$. Otherwise, $x \subseteq V_{n-1}$, then $f_n(x) = \sum_{y \in x} 2^{f_{n-1}(y)}$. Since $f_{n-1}(y) = f_m(y)$ for all $m \geq n$, we get that for that $m > n$, that $f_m(x) = f_n(x)$.

Therefore the set $\{f_n \mid n < \omega\}$ is a $\subseteq$-chain of functions, and therefore $f = \bigcup\{f_n \mid n < \omega\}$ is a well-defined function from V_ω to ω which satisfies $f(x) = \sum_{y \in x} 2^{f(y)}$. Let us check that it is a bijection, and let us use the fact that given $n < \omega$, there is a unique sequence $\langle n_0, \ldots, n_{k-1}\rangle$ such that $n = \sum_{i<k} 2^{n_i}$, this can be proved by induction.

By induction on $n < \omega$, let us assume that for all $k < n$ there is at most a single $z \in V_\omega$ such that $f(z) = k$. If $f(x) = f(y) = n$, then by the uniqueness of the sum, $n = \sum_{i<k} 2^{n_i}$ for some $\langle n_0, \ldots, n_{k-1}\rangle$. However, by the injectivity below n and the definition of $f(x)$ and $f(y)$ it must be that $z \in x$ if and only if $z \in y$. So $x = y$ and n has at most a single x for which $f(x) = n$.

To see that f is indeed surjective, if every $k < n$ is in $\operatorname{rng}(f)$, we can write n as $\sum_{i<k} 2^{n_i}$, then for all $i < k$, $n_i < n$. Therefore there is some z_i such that $f(z_i) = n_i$. Let $x = \{z_i \mid i < k\}$, then $x \in V_\omega$, since each $z_i \in V_{n_i}$ and therefore $x \in V_{n+1}$ where $n = \max\{n_i \mid i < k\}$. Therefore $f(x) = n$, so by induction f is a bijection.

Alternatively, as we are allowed to use the Axiom of Choice, by induction we can show that each V_n is finite: $|V_0| = 0$, and $|V_{n+1}| = 2^{|V_n|}$. Therefore V_ω is a countable union of finite sets and by Theorem 6.2 it is countable. □

Exercise 11.14. If $\alpha > \omega$, then V_α is uncountable.

Proposition 11.14. *If x is a transitive set, then $|x| \geq \operatorname{rank}(x)$.*

Proof. We will show that there is a surjective map from x onto $\operatorname{rank}(x)$. We claim that for all $\alpha < \operatorname{rank}(x)$, $x \cap (V_{\alpha+1} \setminus V_\alpha) \neq \varnothing$. If we can prove this, then the function $f\colon x \to \operatorname{rank}(x)$ given by $f(y) = \operatorname{rank}(y)$ is a surjection as wanted. Let us prove this by induction on the rank of x, and so we assume that if y is a transitive set such that $\operatorname{rank}(y) < \operatorname{rank}(x)$, then y has elements of any rank below its own.

If $\alpha < \operatorname{rank}(x)$, then there is some $y \in x$ such that $\operatorname{rank}(y) \geq \alpha$, if $\operatorname{rank}(y) = \alpha$ we are done, so we may assume that $\operatorname{rank}(y) > \alpha$. Since x is transitive, $x \cap V_{\operatorname{rank}(y)}$ is a transitive set which contains y and with the same rank as y. By the induction hypothesis, $x \cap V_{\operatorname{rank}(y)}$ has an element of rank α, as wanted. □

11.3 The L

Just as we show, using V, that adding the Axiom of Foundation will not add any contradictions to our mathematical foundation, the Axiom of Choice raised similar queries. *Kurt Gödel* in 1938 had proved a similar theorem about the Axiom of Choice by defining a class which we denote by L and call *"the Constructible Universe"*.

The idea in the von Neumann hierarchy is to start with the empty set, and just take all possible subsets at each step. Mind you, a background universe must be established in advance, but what the von Neumann universe shows is that by simply iterating the power set operation continuously, starting from the empty set, we can "generate" every set that is in the universe. But the power set operation is "indiscriminate" and takes all the subsets of a set in one go. Gödel refined this idea in order to generate a possibly-smaller universe, but one where the Axiom of Choice, and much more, holds.

Definition 11.15. If x is a set we define the *definable power set*, $\mathrm{Def}(x)$ as follows:

$$\mathrm{Def}(x) = \{A \subseteq x \mid A \text{ is definable from parameters in the structure } \langle x, \in\rangle\}.$$

For example, $X \in \mathrm{Def}(X)$, since in the structure $\langle X, \in\rangle$ the set defined by $x = x$ is exactly X itself. Similarly, if $a \subseteq X$ is a finite set, i.e. $a = \{x_0, \dots, x_n - 1\}$, then $a \in \mathrm{Def}(X)$ as well as witnessed by the formula $\varphi(x, \bar{x})$ given by $\bigvee_{i<n} x = x_i$.

Definition 11.16. We define the *constructible hierarchy* by recursion on Ord:

1. $L_0 = \varnothing$.
2. $L_{\alpha+1} = \mathrm{Def}(L_\alpha)$.
3. $L_\alpha = \bigcup\{L_\beta \mid \beta < \alpha\}$ whenever α is a limit ordinal.

We denote by L the class $\bigcup\{L_\alpha \mid \alpha \in \mathrm{Ord}\}$, or equivalently $\{x \mid \exists\alpha \in \mathrm{Ord}, x \in L_\alpha\}$.

Similarly to the case of V, we can express with a first-order formula the statement $x \in L$. This allows us to formulate the Axiom of Constructibility, $V = L$, which simply states that $\forall x(x \in L)$. This is an additional axiom which we can add to ZFC, which allows us to prove more things about the structure of the universe.

Exercise 11.15. For each α, L_α is a well-founded transitive set and $\alpha \subseteq L_\alpha$.

Exercise 11.16. For all $n < \omega$, $L_n = V_n$. (Hint: every finite set is definable from its elements.)

Proposition 11.17. *For all $\alpha \geq \omega$, $|L_\alpha| = |\alpha|$.*

Proof. Since $L_\omega = V_\omega$, by Theorem 11.13 it is countable and therefore $|L_\omega| = |\omega|$. Suppose that α is a limit ordinal, then $L_\alpha = \bigcup\{L_\beta \mid \beta < \alpha\}$, since $|L_\beta| = |\beta|$, and $|\alpha| = \sup\{|\beta| \mid \beta < \alpha\}$, we get a surjection from $\alpha \times \alpha$ onto L_α: first for L_β fix a bijection $l_\beta\colon \beta \to L_\beta$, and define $f_\alpha\colon \alpha \times \alpha \to L_\alpha$ by

$$f_\alpha(\beta, \gamma) = \begin{cases} l_\beta(\gamma) & \text{If } \beta < \gamma, \\ \varnothing & \text{otherwise.} \end{cases}$$

This means that $|L_\alpha| \leq |\alpha|$, and since $\alpha \subseteq L_\alpha$, we get a bijection using Theorem 5.13.

If $\alpha = \beta + 1$, we have a surjection from $\mathrm{Form}_\in \times L_\beta^{<\omega}$ onto L_α given by

$$F_\alpha(\varphi, \bar{p}) = \begin{cases} \{x \in L_\beta \mid L_\beta \models \varphi(x, \bar{p})\} & \text{If } \varphi \text{ is } \varphi(x, \bar{x}), \\ \varnothing & \text{otherwise.} \end{cases}$$

Since $\{\in\}$ is finite, $\mathrm{Form}_\in$ is a countable set. By the induction hypothesis, $|L_\beta| = |\beta|$, and therefore $|L_\beta^{<\omega}| = |\beta^{<\omega}|$. It is enough to show that for an infinite β, $|\beta| = |\beta^{<\omega}|$. Utilising Theorem 5.21, we get that $|\beta^n| = |\beta|$ for all $n > 0$. So, we have

$$|\beta| \leq |\beta^{<\omega}| \leq |\omega \times \beta| \leq |\beta \times \beta| = |\beta|.$$

Finally, by the same argument,

$$|\mathrm{Form}_\in \times L_\beta^{<\omega}| = |\omega \times \beta| = |\beta|.$$

And since $\alpha = \beta + 1$ and is an infinite ordinal, there is a bijection between α and β, so α maps onto L_α, and similarly to the limit case we get that $|L_\alpha| = |\alpha|$. $\square$

Exercise 11.17. Assume $V = L$ and fix some α_0. Define $\alpha_{n+1} = \sup\{\beta \mid \exists x \in V_{\alpha_n}, x \in L_{\beta+1} \setminus L_\beta\}$, and let $\alpha = \sup\{\alpha_n \mid n < \omega\}$. Show that $V_\alpha = L_\alpha$. (Visit solution)

Remark

It is not immediate, but we can utilise the recursive nature of L so that L_β we had in the limit step is defined by recursion from F_α, once we have fixed a bijection of $\mathrm{Form}_\in$ with ω, which happens to be definable over L_ω. Luckily, we can certainly do that. This means that the proof above is valid already in L. In other words, $\langle L, \in\rangle \models |L_\alpha| = |\alpha|$. In particular, L satisfies the Axiom of Choice, and indeed a very strong form of the Axiom of Choice.

The remark above tells us that more, it tells us that the axiom $V = L$ holds inside L_α, at least whenever α is a limit ordinal. This, is of course, a consequence of the fact that the constructible hierarchy satisfies the conditions of the Reflection Theorem, and this is just one explicit example of that. Using this we can show that if $M \prec L_\alpha$, then $M \cong L_\beta$ for some $\beta \leq \alpha$.

This allows us to prove that if $A \subseteq \omega$ is in L, then $A \in L_{\omega_1}$. Since $|L_{\omega_1}| = \aleph_1$, it follows that in L the Continuum Hypothesis is true. That is, $L \models 2^{\aleph_0} = \aleph_1$. More general proofs show us that for any infinite κ, $2^\kappa = \kappa^+$ holds in L.

We can try and use stronger logics to define different version of L, by allowing the definability to take place in a logic such as second-order logic, or $\mathcal{L}_{\kappa,\lambda}$, or $\mathcal{L}(Q)$ for some quantifier. The resulting models will often depend on what is true or false in V, where the construction "takes place". We can also add predicates and extend the language with predicates, or start with a different initial set, rather than $\varnothing$. These models play a significant role in the modern set theoretic research.

Unlike the equivalence between the Axiom of Foundation and that every set is in the von Neumann hierarchy, the Axiom of Choice can be true in V even if $V \neq L$.

Chapter 12

Models of fragments of ZFC

Chapter Goals

In this chapter we will learn about

- Models of ZFC without Replacement.
- Models of ZFC without Infinity.
- Models of ZFC without Power Set.
- Elementary equivalence and embeddings between transitive models of these theories, the consequences of their existence, and what can we say about them.

12.1 Setting the stage

Unfortunately, we cannot prove from ZFC that there is any set M and a relation E such that $\langle M, E\rangle \models \mathsf{ZFC}$. This is a consequence of Gödel's Incompleteness Theorem (we can, however, assume additional axioms to prove that). But working in ZFC we can prove the existence of models of "fragments" of ZFC. This makes them, in a technical sense, *weaker theories*.

Lemma 12.1. *Suppose that M is a transitive set, then $\langle M, \in\rangle$ satisfies the Axioms of Extensionality and Foundation.*

Proof. Suppose that $x, y \in M$ and $x \neq y$, then without loss of generality there is some $z \in x$ such that $z \notin y$. Since M is transitive, $x \subseteq M$, so $z \in M$. Therefore $M \models \neg(x \subseteq y \land y \subseteq x)$. In other words, if $M \models x \subseteq y \land y \subseteq x$, then $M \models x = y$, as wanted.

Similarly, for the Axiom of Foundation, suppose that $x \in M$, if $x \neq \varnothing$, then there is some $y \in x$ such that $x \cap y = \varnothing$. Since M is transitive, it follows that $y \in M$ and that $x \cap y \subseteq M$ as well. Since $x \cap y = \varnothing$, it must be that $M \models x \cap y = \varnothing$. □

Lemma 12.2. *For any limit ordinal α, V_α is a model of the Axioms of Extensionality, Union, Power Set, Separation, Foundation, and Choice.*

Proof. Since V_α is a transitive set, by Theorem 12.1 it satisfies Extensionality and Foundation. The Axiom of Union and the Axiom of Separation do not increase the rank of a set. Namely, if $\operatorname{rank}(x) = \delta$, then $\operatorname{rank}(\bigcup x) \leq \delta$, and if $\varphi(y)$ is a formula, then $\operatorname{rank}(\{y \in x \mid \varphi(y)\}) \leq \delta$ as well. Therefore these axioms also hold in V_α.

The rank of $\mathcal{P}(x)$ is exactly $\operatorname{rank}(x) + 1$, and since α is a limit ordinal, $\operatorname{rank}(x) + 1 < \alpha$, so $\mathcal{P}(x) \in V_\alpha$ if and only if $x \in V_\alpha$. Therefore V_α satisfies the Axiom of Power Set as well.

Finally, if $x \in V_\alpha$ is a family of non-empty sets, if $c\colon x \to \bigcup x$ (e.g., c is a choice function), so $\operatorname{rank}(c) \leq \operatorname{rank}(x) + 5$. As α is a limit ordinal, any choice function must exist in V_α. □

Remark

This is one of the places where the formulation of the Axiom of Choice really makes a difference. If we formulate it by "Existence of Representatives", which is to say "If x is a family of pairwise disjoint non-empty sets, then there is c such that $|c \cap z| = 1$ for all $z \in x$", then $c \subseteq \bigcup x$, in which case every V_α satisfies the Axiom of Choice, not just the limit ones.

Definition 12.3. Suppose that κ is a cardinal, the set H_κ is the set of "sets which are hereditarily of size $< \kappa$". That means that x, its elements, their elements, and so on, are all smaller in size than κ. This is formalised by saying that the size of the transitive closure of x is smaller than κ. Namely, $H_\kappa = \{x \mid |\operatorname{tcl}(x)| < \kappa\}$.

So $H_{\aleph_0}$ is the set of hereditarily finite sets, $H_{\aleph_1}$ is the set of hereditarily countable sets.

Exercise 12.1. Show that H_κ is a transitive set. (Hint: Theorem 11.14 shows that $H_\kappa \subseteq V_\kappa$.) (Visit solution)

Exercise 12.2. If $a \in H_\kappa$ and $b \subseteq a$, then $b \in H_\kappa$.

Exercise 12.3. If $a, b \in H_\kappa$, then $a \times b \in H_\kappa$. (Visit solution)

Exercise 12.4. Show that the sets H_κ form a hierarchy. Namely, for every x, there is some κ such that $x \in H_\kappa$ and that if κ is a limit cardinal, then $H_\kappa = \bigcup\{H_\lambda \mid \lambda < \kappa\}$.

Lemma 12.4. *H_κ is a model of the Axioms of Union and Choice.*

Proof. Let $x \in H_\kappa$. Since $\bigcup x \subseteq \operatorname{tcl}(x)$, we have that $\operatorname{tcl}(\bigcup x) \subseteq \operatorname{tcl}(x)$, so $\bigcup x \in H_\kappa$.

To see that the Axiom of Choice holds, the proof is similar to the proof in Theorem 12.2. If $x \in H_\kappa$ is a family of non-empty sets, then $\bigcup x \in H_\kappa$, and so $x \times \bigcup x \in H_\kappa$. Any choice function is a subset of the product, and so must be in H_κ as well. □

Remark

The definition of H_κ implies that x, and perhaps an even larger set, can be well-ordered. If we work in ZF we need to modify the definition to allow it to capture possible failures of the Axiom of Choice. This can be done by defining it as "there is a transitive set y containing x such that $\kappa \not\leq^* y$" or by other more complicated means. However, even in these cases the situation gets very subtle and complicated.

12.2 Powerless Models

Definition 12.5. We say that an infinite cardinal κ is *regular* if whenever $A \subseteq \kappa$ and $|A| < \kappa$, then $\sup A < \kappa$. If κ is not regular we say that it is a *singular* cardinal.

For example, $\aleph_1$ is a regular cardinal: if $A \subseteq \omega_1$ is a countable set of countable ordinals, then $\sup A$ is a countable union of countable sets, and so it is a countable ordinal again. On the other hand, $\aleph_\omega$ is a singular cardinal, since $\{\omega_n \mid n < \omega\}$ is a countable set, but $\omega_\omega = \sup\{\omega_n \mid n < \omega\}$.

Theorem 12.6. *Let κ be an infinite cardinal such that if $A \subseteq \kappa$ and $|A| < \kappa$, then $\sup A < \kappa$. Then H_κ satisfies the Axiom of Replacement.*

Proof. Suppose that $\varphi(x, y)$ is a formula such that $H_\kappa \models \forall x \exists! y \varphi(x, y)$. For a given $a \in H_\kappa$, let $b = \{y \in H_\kappa \mid \exists x \in a, H_\kappa \models \varphi(x, y)\}$. Then $|b| \leq |a|$, since φ defines a surjection from a onto b, and for each $y \in b$, $|\operatorname{tcl}(\{y\})| < \kappa$. The set $\{|\operatorname{tcl}(\{y\})| \mid y \in b\}$ has size $\leq |b| < \kappa$, so by our assumption on κ, there is some cardinal $\lambda < \kappa$ such that $|\operatorname{tcl}(\{y\})| \leq \lambda$ for all $y \in b$. Therefore the size of $Y = \bigcup\{\operatorname{tcl}(\{y\}) \mid y \in b\}$ is at most $\lambda \cdot |b| < \kappa$. Note that Y itself is a transitive set, as is $Y \cup \{b\}$, since $b \subseteq Y$. Therefore $b \in H_\kappa$, so the Axiom of Replacement holds. □

Proposition 12.7. *κ^+ is regular whenever κ is infinite.*

Proof. If $A \subseteq \kappa^+$ and $|A| < \kappa^+$, then $|A| \leq \kappa$, and by definition for every $\alpha \in A$, $|\alpha| \leq \kappa$. Since $\sup A = \bigcup A$, let us show that $|\bigcup\{\{\alpha\} \times \alpha \mid \alpha \in A\}| \leq \kappa$, as the set has a natural projection onto $\bigcup A = \sup A$. For each $\alpha \in A$, let $f_\alpha \colon \alpha \to \kappa$ be an injection and let $f \colon A \to \kappa$ be an injection. Then $F(\alpha, \beta) = \langle f(\alpha), f_\alpha(\beta)\rangle$ is a well-defined function when $\langle \alpha, \beta\rangle \in \bigcup\{\{\alpha\} \times \alpha \mid \alpha \in A\}$. Moreover, F is injective and that $\operatorname{rng}(F) \subseteq \kappa \times \kappa$. Since $|\kappa \times \kappa| = \kappa$, this completes the proof. □

Exercise 12.5. Suppose that κ is an uncountable cardinal, then $|\{\delta < \kappa \mid \delta \text{ is a limit ordinal}\}| = \kappa$.
(Visit solution)

Remark

While ZFC proves that every infinite singular cardinal is a limit cardinal, ZFC cannot prove that there are regular limit cardinals other than $\aleph_0$. The axioms that posit the existence of such cardinals are called *large cardinal axioms* and they form one of the most important aspects of modern set theoretic research.

Theorem 12.8. *For any infinite cardinal κ, the Axiom of Power Set fails in H_{κ^+}. So H_{κ^+} is a model of* ZFC *without Power Set.*

Proof. Note that $\kappa \in H_{\kappa^+}$, so $\mathcal{P}(\kappa) \subseteq H_{\kappa^+}$, but $\mathcal{P}(\kappa) \notin H_{\kappa^+}$. Since $\mathcal{P}(\kappa)$ is a transitive set and $|\mathcal{P}(\kappa)| = 2^\kappa \geq \kappa^+$ by Theorem 1.38 and the definition of a successor cardinal. Therefore, for every $P \in H_{\kappa^+}$ there exists some $a \in H_{\kappa^+}$ such that $a \subseteq \kappa$ and $a \notin P$. So in H_{κ^+} it is impossible for κ to have a power set, and so the Axiom of Power Set fails. Since κ^+ is an uncountable regular cardinal, $\omega \in H_{\kappa^+}$ so the Axiom of Infinity, and all the other axioms of ZFC hold there. □

Remark

When omitting the Axiom of Power Set we usually strengthen the Axiom of Replacement to the *Axiom of Collection* (which states that for any $\varphi(x, y)$ and any set A we can find a set B such that for any $a \in A$ there is some $b \in B$ such that $\varphi(a, b)$ holds). These two axioms are equivalent when the Axiom of Power Set holds, but otherwise the Axiom of Collection is stronger.

Interestingly, when working without the Axiom of Power Set, we can no longer prove that the Axiom of Choice implies that every set can be well-ordered or Zorn's Lemma, and many other "reasonable consequences" of the Axiom of Choice may fail as well.

12.3 Not beyond infinity

Theorem 12.9. *$\langle V_\omega, \in\rangle$ satisfies* ZFC *without the Axiom of Infinity.*

Proof. Since ω is a limit ordinal, by Theorem 12.2 it is enough to verify the Axiom of Replacement holds. If we can show that $V_\omega = H_{\aleph_0}$, then by Theorem 12.6 we get that V_ω is a model of the Axiom of Replacement. Since $H_{\aleph_0} \subseteq V_\omega$, it is enough to prove that $V_\omega \subseteq H_{\aleph_0}$. But this is easy, since if $x \in V_\omega$, then there is some $n < \omega$ such that $x \in V_n$. Since V_n is a finite transitive set, x must be in $H_{\aleph_0}$. □

Remark

Note that V_ω satisfies the negation of the Axiom of Infinity.

It is noteworthy to remark that when we replace the Axiom of Infinity (in ZFC) with its negation we get a theory that is equivalent to the Peano Arithmetic, that is the standard axioms of the natural numbers. If we strengthen the theory slightly more by requiring that every set is contained in a transitive set, then the theory has a stronger sense of equivalence with Peano Arithmetic.

12.4 Replacement Killers

Theorem 12.10. *Suppose that δ is a limit ordinal, then the Axiom of Replacement fails in $V_{\delta+\delta}$, and the Axiom of Infinity holds. That is,* ZFC *without the Axiom of Replacement, but with the Axiom of Separation holds in $V_{\delta+\delta}$.*

Proof. Since $\omega \in V_{\delta+\delta}$, the Axiom of Infinity holds. By Theorem 12.2 it remains to show the failure of Replacement. Consider $\varphi(x, y)$ to be the formula: $x < \delta$ and $y = \delta + x$, or $x \notin \delta$ and $x = y$. Note that formally speaking, φ has three variables: x, y, z where z takes the place of δ, and we then use δ as a parameter for the definition.

Let us check that $V_{\delta+\delta} \models \forall x \exists! y \varphi(x, y)$. Let $x \in V_{\delta+\delta}$, if $x < \delta$, then $\varphi(x, \delta + x)$ holds and $\delta + x < \delta + \delta$, so $\delta + x \in V_{\delta+\delta}$, and if $x \notin \delta$, then $\varphi(x, x)$ holds. If $V_{\delta+\delta} \models \varphi(x, y) \wedge \varphi(x, y')$, then either $x \notin \delta$, then $x = y$ and $x = y'$, so $y = y'$; and if $x < \delta$ then both y and y' must be equal to $\delta + x$. So in either case, $V_{\delta+\delta} \models \forall x \exists! y \varphi(x, y)$.

However, $\delta \in V_{\delta+\delta}$ and so if the Axiom of Replacement was true, there should be some $y \in V_{\delta+\delta}$ such that $y = \{x \in V_{\delta+\delta} \mid \exists\alpha < \delta, \varphi(\alpha, x)\}$. Easily, y must be $\{\delta + \alpha \mid \alpha < \delta\}$, whose rank is $\delta + \delta$ and therefore is not an element of $V_{\delta+\delta}$, so the Axiom of Replacement fails. □

Exercise 12.6. The Axiom of Replacement fails in V_{ω_1}. (Hint: $\omega_1 \leq^* \mathcal{P}(\omega\times\omega)$.) (Visit solution)

Exercise 12.7. More generally, show that if for some $\alpha < \kappa$, $|V_\alpha| \geq \kappa$ (or more generally, some $x \in V_\kappa$ satisfies that $|x| \geq \kappa$), then V_κ is not a model of the Axiom of Replacement.

Exercise 12.8. Suppose that κ is an uncountable regular cardinal such that for all $\alpha < \kappa$, $2^\alpha < \kappa$. Show that $V_\kappa \models \mathsf{ZFC}$. (Hint: show that $H_\kappa = V_\kappa$.) (Visit solution)

12.5 Elementary submodels of fragments and their stories

Theorem 12.11. *Suppose that κ is an uncountable regular cardinal such that for all $\alpha < \kappa$, $2^\alpha < \kappa$. Then there is some $\lambda < \kappa$ such that $V_\lambda \prec V_\kappa$.*

Proof. We define by recursion a sequence of elementary submodels of V_κ. Let M_0 be a countable elementary submodel of V_κ and let $\operatorname{rank}(M_0) = \lambda_0 < \kappa$. We define by recursion M_{n+1} and λ_{n+1} as follows, let M_{n+1} be an elementary submodel of V_κ generated by V_{λ_n}. Since $|V_{\lambda_n}| < |V_\kappa|$, $|M_{n+1}| = |V_{\lambda_n}| + \aleph_0 < \kappa$, and therefore $\lambda_{n+1} = \operatorname{rank}(M_{n+1}) < \kappa$.

Finally, let $\lambda = \sup\{\lambda_n \mid n < \omega\}$. Then $\lambda < \kappa$ and $V_\lambda = \bigcup\{M_n \mid n < \omega\} = \bigcup\{V_{\lambda_n} \mid n < \omega\}$. We can use the Tarski–Vaught Criterion to test elementarity. If $V_\kappa \models \exists x\varphi(x, \bar{y})$ for $\bar{y} \in V_\lambda$, then there is some $n < \omega$ such that $\bar{y} \in M_n$, by elementarity $M_n \models \exists x\varphi(x, \bar{y})$, so there is some $x \in M_n \subseteq V_\lambda$ such that $M_n \models \varphi(x, \bar{y})$, so $V_\kappa \models \varphi(x, \bar{y})$ for some $x \in V_\lambda$. □

Theorem 12.12. *Suppose that $\alpha < \beta$ and $V_\alpha \prec V_\beta$, then V_α satisfies the Axiom of Replacement.*

Proof. Suppose that $\varphi(x, y)$ is a formula such that $V_\alpha \models \forall x\exists! y\varphi(x, y)$ and let $A \in V_\alpha$. Then for every $a \in A$ there is a unique $b \in V_\alpha$ such that $V_\alpha \models \varphi(a, b)$. By elementarity, $V_\beta \models \forall x\exists! y\varphi(x, y)$ as well, and for every $a \in A$, if $V_\alpha \models \varphi(a, b)$, then $b \in V_\alpha$ and $V_\beta \models \varphi(a, b)$ as well. Therefore setting $B = \{b \in V_\alpha \mid \exists a \in A, V_\beta \models \varphi(a, b)\}$ we obtain that

$$V_\beta \models \forall x(x \in A \to \exists y(y \in B \land \varphi(x, y))),$$

in particular, $V_\beta \models \exists Y(Y = \{b \mid \exists a \in A, \varphi(a, b)\})$, so by elementarity V_α must satisfy the same, and so the Axiom of Replacement holds in V_α. □

In particular, in the above case, both V_α and V_β are models of ZFC. As we remarked, we cannot prove the existence of such α and β from the axioms of ZFC themselves. But in fact, this assumption requires even more than just the existence of some $V_\alpha \models \mathsf{ZFC}$.

Exercise 12.9. Suppose that κ is a regular cardinal and $V_\kappa \models \mathsf{ZFC}$, show that there are $\alpha < \beta < \kappa$ such that $V_\alpha \prec V_\beta$. That is, the least α and β such that $V_\alpha \prec V_\beta$ must be singular cardinals! (Visit solution)

Proposition 12.13. *Suppose that α is the least such that $V_\alpha \models \mathsf{ZFC}$. If $V_\alpha \equiv V_\beta$, then $\alpha = \beta$.*

Proof. Since α is the least such that $V_\alpha \models \mathsf{ZFC}$, it has to be the case that $\alpha \leq \beta$. If $\alpha < \beta$, then V_β satisfies "There exists an ordinal δ such that $V_\delta \models \mathsf{ZFC}$". However, since α is the least such ordinal, this sentence is false in V_α, so it must be that $\alpha = \beta$. □

Exercise 12.10. Show that if V_α is the third ordinal such that $V_\alpha \models \mathsf{ZFC}$ and $V_\alpha \equiv V_\beta$, then $\alpha = \beta$.

On the other hand, in yet another stark contrast between elementary equivalence and elementary submodels, we can in fact prove the following statement outright in ZFC.

Proposition 12.14. *Let $\kappa = |\mathcal{P}(\omega)|$. Then there are $\alpha < \beta < \kappa^+$ such that $V_\alpha \equiv V_\beta$.*

Proof. Note that $\mathrm{Sent}_\in$ is a countable set, so we can enumerate it as $\{\varphi_n \mid n < \omega\}$. For each V_α for $\alpha < \kappa^+$ let $T_\alpha = \{n < \omega \mid V_\alpha \models \varphi_n\}$. This defines a function from κ^+ to $\mathcal{P}(\omega)$. Since $|\mathcal{P}(\omega)| < \kappa^+$ this function cannot be injective, so there are α and β such that $T_\alpha = T_\beta$ and therefore $V_\alpha \equiv V_\beta$. □

Note that in all of the cases we have seen above, the elementary embedding witnessing that $V_\alpha \prec V_\beta$ was just the identity function. Or at the very least, nothing was put in place to require more than that. Requiring that there is an embedding which not the identity is even harder to prove. Even if we assume that there is a regular cardinal κ for which $V_\kappa \models \mathsf{ZFC}$, we still cannot prove that any such elementary embedding is not the identity.

The reason is that if $j\colon V_\alpha \to V_\beta$ is an elementary embedding, this means that whatever properties x has in V_α, $j(x)$ must have in V_β. If j is the identity function, then we are just saying that the properties of x are preserved between V_α and V_β. But if j is not the identity, we are saying, in effect, that the properties of x are "reflected" in higher and higher sets.

Exercise 12.11. Suppose that $j\colon V_\alpha \to V_\beta$ is an elementary embedding and for some ordinal δ, $\delta < j(\delta)$. Let κ be the least such ordinal and let $\mathcal{U} = \{A \subseteq \kappa \mid \kappa \in j(A)\}$. Then $\mathcal{U}$ is a free ultrafilter on κ.

Theorem 12.15. *Suppose that $\kappa > \omega$ and $M \prec H_\kappa$ is a countable elementary submodel. If $x \in M$ and x is countable, then $x \subseteq M$.*

Proof. Let $x \in M$ be a countable set. Then there is some $f \in H_\kappa$ such that $f\colon \omega \to x$ is a bijection. Note that H_κ satisfies the statement "f is a bijection between ω and x". First we note that ω itself is definable in H_κ as the least limit ordinal, and therefore by elementarity, $\omega \in M$, and the same can be said about each $n < \omega$. Next, since H_κ satisfies the property "there is a bijection between ω and x", then by elementarity, there exists such a bijection in M.

Let $f \in M$ be a bijection between ω and x. Then for every $n < \omega$, by elementarity $f(n) \in M$. However, since f was surjective, that means that for every $y \in x$, $y \in M$, as wanted. □

As an immediate corollary we obtain that if $M \prec H_{\aleph_1}$, then M is a transitive set. The above theorem will work for V_α, when α is a limit, as well.

Exercise 12.12. If $\kappa > \omega_1$ and $M \prec H_\kappa$ is countable, then M is not transitive. (Visit solution)

Appendix A

Selected solutions

Solutions to Chapter 1

Solution to Exercise 1.6 (Return to exercise)

Proof. Let A and B be two collections. Let us show that $A \subseteq B$ if and only if $A \setminus B = \varnothing$ first.

Suppose that $A \subseteq B$, we will show that $A \setminus B = \{a \mid a \in A \text{ and } a \notin B\} = \varnothing$. Since $\varnothing$ is a subset of $A \setminus B$, it is enough to show that $A \setminus B$ does not have any members. Given $a \in A \setminus B$, by definition $a \in A$ and $a \notin B$. By our assumption, $A \subseteq B$, so since $a \in A$ it follows that $a \in B$. Therefore $a \in B$ and $a \notin B$, which is impossible, so no such a can exist and $A \setminus B = \varnothing$.

In the other direction, suppose that $A \setminus B = \varnothing$, we will show that if $a \in A$, then $a \in B$. Indeed, given $a \in A$, since $A \setminus = \varnothing$, it follows that $a \notin A \setminus B$, so either $a \notin A$ or $a \in B$. Since we chose a such that $a \in A$, it must be that $a \in B$, so $A \subseteq B$.

Next, we will show that $A \cap B = \varnothing$ if and only if $A \setminus B = A$. Assume first that $A \cap B = \varnothing$, since $A \setminus B \subseteq A$, it is enough to show that $A \subseteq A \setminus B$. Given $a \in A$, we know by the assumption that $a \notin A \cap B$. Therefore it must be that $a \notin A$ or $a \notin B$. Since $a \in A$, it follows that $a \notin B$, and so we have that $a \in A \setminus B$, and the equality follows.

Finally, suppose that $A \setminus B = A$, we will show that $A \cap B = \varnothing$. Suppose that $a \in A \cap B$, then $a \in A$ and $a \in B$. Since $a \in A$, by the assumption we have that $a \in A \setminus B$. Therefore, $a \in A$ and $a \notin B$. This means that $a \in B$ and $a \notin B$, which is impossible, so no such a can exist, and therefore $A \cap B = \varnothing$. □

Solution to Exercise 1.12 (Return to exercise)

Proof. Recall that $A \triangle B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$.

For the first identity, first recall that $A \cup \varnothing = A$ and $A \cap \varnothing = \varnothing$. Therefore $A \triangle \varnothing = A \setminus \varnothing = A$ as wanted.

For the second identity, note that $A \cap A = A \cup A = A$, so $A \triangle A = A \setminus A = \varnothing$. □

Solution to Exercise 1.14 (Return to exercise)

Proof.

1. $\mathcal{P}(\varnothing) = \{\varnothing\}$,
2. $\mathcal{P}(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}$,
3. $\mathcal{P}(\{\varnothing, \{\varnothing\}\}) = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$. □

Solution to Exercise 1.20 (Return to exercise)

Proof. We saw that if A is a set, then $B = \{a \in A \mid a \notin a\}$ is a subset of A such that $B \notin A$. Therefore $\mathcal{P}(A) \nsubseteq A$. Since $\varnothing$ is a subset of every set, in particular, $\varnothing \subseteq \mathcal{P}(\varnothing)$. □

Solution to Exercise 1.23 (Return to exercise)

Proof. The definition does not give us an ordered pair. Suppose that $\{x, \{\varnothing, y\}\} = \{a, \{\varnothing, b\}\}$. If $x = \{\varnothing, b\}$ and $a = \{\varnothing, y\}$, then both sets are $\{a, x\}$. So by taking $y \neq b$ we have a counterexample, e.g. $y = 4$ and $b = 5$ would give us that $\langle\{\varnothing, 5\}, 4\rangle = \langle\{\varnothing, 4\}, 5\rangle$ had this would define an ordered pair. □

Solution to Exercise 1.28 (Return to exercise)

Proof. Counterexample: $A = B = C = \{5\}$, under the assumption that 5 is not interpreted as the ordered pair $\langle 5, 5\rangle$.

We get that $A \cap (B \times C) = \varnothing$, whereas $(A \cap B) \times (A \cap C) = \{5\} \times \{5\} = \{\langle 5, 5\rangle\}$. □

Solution to Exercise 1.30 (Return to exercise)

Proof. We will show that E is reflexive (on $\mathbb{Z}$), symmetric, and transitive.

Reflexive For any $a \in \mathbb{Z}$, $a - a = 0$ (which is even), so $\langle a, a\rangle \in E$.

Symmetric Suppose that $a, b \in \mathbb{Z}$, then $a - b = -(b - a)$, since the negative of an even number is even, the relation is symmetric.

Transitive Suppose that $a, b, c \in \mathbb{Z}$ and $a - b$ and $b - c$ are both even. Then $(a - b) + (b - c) = a - c$ is the sum of two even integers, and therefore is even as well. □

Solution to Exercise 1.36 (Return to exercise)

Proof. We will show that K_f is reflexive (on A), symmetric and transitive.

Reflexive For any $a \in A$, $f(a) = f(a)$, so trivially, K_f is reflexive.

Symmetric For any $a, b \in A$, if $f(a) = f(b)$, then $f(b) = f(a)$, and therefore K_f is symmetric.

Transitive For any $a, b, c \in A$, if $f(a) = f(b)$ and $f(b) = f(c)$, then $f(a) = f(c)$ since equality is transitive. Therefore K_f is symmetric as well. □

Solution to Exercise 1.37 (Return to exercise)

Proof. Suppose that E is an equivalence relation on A, let $B = A/E$, the quotient set. We define $f\colon A \to B$ by $f(a) = a/E$. Let us show that $E = K_f$.

Suppose that $\langle a, b\rangle \in E$, then $a/E = b/E$ and therefore $f(a) = f(b)$, so $\langle a, b\rangle \in K_f$ as well. In the other direction, suppose that $f(a) = f(b)$, then $a/E = b/E$ and as we saw, this means that $a \mathrel{E} b$, or $\langle a, b\rangle \in E$, and so equality holds. □

Solution to Exercise 1.42 (Return to exercise)

Proof. The function f is well-defined since if $\{a\} = \{a'\}$, by Extensionality it means that $a \in \{a'\}$ and since $\{a'\}$ has exactly one member, a', it means that $a = a'$.

The function is injective by the same argument: if $\{a\} = \{b\}$, that is $f(a) = f(b)$, then $a = b$. □

Solutions to Chapter 2

Solution to Exercise 2.5 (Return to exercise)

Proof. $\langle \mathbb{Q}, \leq \rangle$ has no minimal or maximal elements: for any rational number q, $q-1$ and $q+1$ are also rational numbers and $q - 1 < q < q + 1$. □

Solution to Exercise 2.7 (Return to exercise)

Proof. We claim that if A has at least two distinct elements, then $\mathcal{P}(A)$ is not totally ordered by $\subseteq$. To see that, note that if $a, b \in A$ are distinct elements, then $\{a\}$ and $\{b\}$ are incomparable in $\subseteq$.

On the other hand, if $A = \{a\}$, then $\mathcal{P}(A) = \mathcal{P}(\{a\}) = \{\varnothing, \{a\}\}$ is easily totally ordered by $\subseteq$. Moreover, if $A = \varnothing$, then $\mathcal{P}(A) = \{\varnothing\}$ is also trivially a total order.

Therefore the class we are looking for is

$$C = \{A \mid A \text{ is empty or there is some } a \text{ such that } A = \{a\}\}.$$

This is a proper class. If it was a set, then by the Axiom of Union $\bigcup C$ was a set as well. However, if a is any set, then $\{a\} \in C$, so $\bigcup C$ will include the class of all sets, which we saw is a proper class. Therefore $\bigcup C$ cannot be a set, so C cannot be a set either. □

Solution to Exercise 2.11 (Return to exercise)

Proof. Consider the order $\preccurlyeq$ defined on $\mathbb{Z}$ given by $n \preccurlyeq m$ if and only if $n = 0$ or ($nm > 0$ and $|n| \leq |m|$). In other words, 0 is a minimum, then we order the positive and negative integers in two chains. We need to show that this is a partial order and that it satisfies the structural requirements.

To see that it is reflexive, note that $0 \preccurlyeq 0$ by the first clause of the definition, and if $n \neq 0$, then $nn = n^2 > 0$ and $|n| \leq |n|$, so indeed this relation is reflexive. It is antisymmetric, since if $n \preccurlyeq m$ and $m \preccurlyeq n$, then either $n = m = 0$ or else $|n| \leq |m| \leq |n|$ which means that $|n| = |m|$, and since $nm > 0$ we must have that they both have the same sign (either negative or positive), so $n = m$. Finally, if $k \preccurlyeq n \preccurlyeq m$, if $k = 0$ we are done. Since $n = 0$ implies $k = 0$ as well, we may assume this is not the case either. In the remaining case, $nmnk = n^2mk > 0$, but since n^2 is positive, which means that mk must be positive too. So all three integers have the same sign, and $|k| \leq |n| \leq |m|$, so $|k| \leq |m|$ as well, as wanted.

Note that 0 is the minimum element, since given any $m \in \mathbb{Z}$, $0 \preccurlyeq m$ by definition. Also, there are no maximal elements, since if $n \geq 0$, then $n \preccurlyeq n + 1$, and if $n \leq 0$, then $n \preccurlyeq n - 1$.

Finally, $P = \{k \in \mathbb{Z} \mid 0 \leq k\}$ and $N = \{k \in \mathbb{Z} \mid k \leq 0\}$ are easily chains, and they are also maximal, in the case of P, any negative integer in incomparable with its absolute value, e.g. -3 is incomparable with 3, since $-3 \cdot 3 = -9 < 0$; and in the case of N the opposite holds. So these chains are maximal. Suppose that C is a maximal chain, if $k \in C$ is positive, then all the elements of C must be non-negative, and so $C \subseteq P$, and by maximality $C = P$. Similarly, if $k \in C$ is negative, then all the elements in C must be non-positive, so $C \subseteq N$ and by maximality of C we get $C = N$. □

Solution to Exercise 2.14 (Return to exercise)

Proof. Suppose that $A = \{a\}$, then both directions of the equivalence are vacuously true.

Let us deal with the case where A has at least two elements. If A is dense, whenever $a < b$ then there is some c such that $a < c < b$, therefore it is impossible that b is a successor of a, and therefore a cannot have a successor.

In the other direction, we need to show that A is dense. If $a < b$, since b is not the successor of a, there must be some c such that $a < c < b$, so A must be dense. □

Solution to Exercise 2.15 (Return to exercise)

Proof. Let $\langle A, <_A \rangle$ and $\langle B, <_B \rangle$ be two strict orders. We say that $F\colon A \to B$ is an embedding (of strict orders) if F is injective and $a <_A a'$ if and only if $F(a) <_B F(a')$. □

Solution to Exercise 2.17 (Return to exercise)

Proof. Suppose that C is a chain in A, and $x, y \in F[C]$ are elements in B. Then by definition, there are $u, v \in A$ such that $F(u) = x$ and $F(v) = y$. Since C was a chain, $u \leq_A v$ or $v \leq_A u$. Since F was an embedding, $F(u) = x \leq_B y = F(v)$ or $F(v)y = \leq_B x = F(u)$, so x and y are comparable. If C is an antichain, the proof is the same, noting that if x and y are comparable in $\leq_B$, then u and v must be comparable in $\leq_A$ and therefore must be equal.

Suppose that F is an isomorphism. If $a \in A$ is a maximal element, we claim that $F(a)$ is a maximal element in B. Suppose that $b \in B$ and $F(a) \leq b$. Since F is surjective, there is some $a' \in A$ such that $F(a') = b$. Therefore, $F(a) \leq_B F(a')$, so by the definition of embedding, $a \leq_A a'$. Since a is maximal, it must be that $a = a'$ and therefore $F(a) = F(a') = b$, so $F(a)$ is maximal in B. The rest of the proofs are similar. □

Solution to Exercise 2.18 (Return to exercise)

Proof. Let $\mathrm{id} = \{\langle n, n \rangle \mid n \in \mathbb{N}\}$, then it is certainly reflexive on $\mathbb{N}$. The relation is transitive and antisymmetric for trivial reasons: if $n \,\mathrm{id}\, m$ and $m \,\mathrm{id}\, n$, then $n = m$ by definition, and the same holds in the case of transitivity. To see that this is not isomorphic to the standard linear order $\leq$, note that this partial order is not linear: $1 \not\mathrm{id}\, 2$ whereas $\leq$ is linear. □

Solution to Exercise 2.23 (Return to exercise)

Proof. We define the function $F\colon \mathcal{P}(A \cup B) \to \mathcal{P}(A) \times \mathcal{P}(B)$ by $F(X) = \langle X \cap A, X \cap B \rangle$. Let us verify first that this is an embedding. We have that $X \subseteq Y$ if and only if $X \cap A \subseteq Y \cap A$ and $X \cap B \subseteq Y \cap B$, so this is indeed an order embedding of $\subseteq$ into the pointwise product of the two power sets.

To see that this is surjective, if $\langle X, Y \rangle \in \mathcal{P}(A) \times \mathcal{P}(B)$, then $X \cup Y \in \mathcal{P}(A \cup B)$, and since $A \cap B = \varnothing$, $X = (X \cup Y) \cap A$ and $Y = (X \cup Y) \cap B$. So $F(X \cup Y) = \langle X, Y \rangle$. □

Solution to Exercise 2.24 (Return to exercise)

Proof. We have seen that the lexicographic product of partial orders is a partial order, so it remains to check the linearity. Suppose that $\langle a_0, b_0\rangle \not\leq_{\text{Lex}} \langle a_1, b_1\rangle$, then either $a_0 \not\leq_A a_1$, in which case by the linearity of A, $a_1 <_A a_0$, and therefore $\langle a_1, b_1\rangle \leq_{\text{Lex}} \langle a_0, b_0\rangle$, or else $a_0 = a_1$ but $b_0 \not\leq_B b_1$, which then by the linearity of B we have $b_1 <_B b_0$, so again $\langle a_1, b_1\rangle \leq_{\text{Lex}} \langle a_0, b_0\rangle$ as wanted. □

Solution to Exercise 2.26 (Return to exercise)

Proof. If $\langle z, q\rangle <_{\text{Lex}} \langle z', q'\rangle$ in $\mathbb{Z}\times\mathbb{Q}$, then either $z < z'$ and then $\langle z, q\rangle <_{\text{Lex}} \langle z, q+1\rangle <_{\text{Lex}} \langle z', q'\rangle$ or $z = z'$ in which case $\langle z, \frac{q+q'}{2}\rangle$ is witnessing the density of $\mathbb{Z}\times\mathbb{Q}$. On the other hand, there is no pair strictly between $\langle 0, 0\rangle$ and $\langle 0, 1\rangle$ in $\mathbb{Q}\times\mathbb{Z}$. □

Solutions to Chapter 3

Solution to Exercise 3.4 (Return to exercise)

Proof. Let $\mathcal{A} = \{\mathbb{N}\setminus\{0, \dots, n-1\} \mid n \in \mathbb{N}\}$. We claim that $\mathcal{A}$ is non-empty and does not have a minimal element, and therefore $\mathcal{P}(\mathbb{N})$ is not well-founded.

First, setting $n = 0$ shows that $\mathbb{N} \in \mathcal{A}$, so it is non-empty. Suppose that $X \in \mathcal{A}$, then there is some n such that $X = \mathbb{N}\setminus\{0, \dots, n-1\}$. Therefore $X \setminus \{n\} = \mathbb{N}\setminus\{0, \dots, n\}$ is also in $\mathcal{A}$, but $X \setminus \{n\}$ is a proper subset of X, and therefore X was not minimal.

Alternatively, since $\mathbb{N}$ does not have a maximal element, $\langle \mathbb{N}, \geq\rangle$ does not have a minimal element, so it is not well-founded. □

Solution to Exercise 3.6 (Return to exercise)

Proof. We define a function $F\colon \mathcal{P}(B) \to \mathcal{P}(A)$ given by $F(X) = \{a \in A \mid f(a) \in X\}$. We claim that F is an embedding of $\mathcal{P}(B)$ into $\mathcal{P}(A)$. To see that F is order preserving note that $X \subseteq Y$ implies that if $a \in A$ and $f(a) \in X$, then $f(a) \in Y$, so $F(X) \subseteq F(Y)$. In the other direction, if $F(X) \subseteq F(Y)$, then for every $a \in A$, if $f(a) \in X$, then $f(a) \in Y$. Since f was surjective, if $x \in X$, then there is some $a \in A$ such that $f(a) = x$, therefore $a \in F(X)$, and therefore $a \in F(Y)$, and therefore $f(a) \in Y$, so $x \in Y$ as wanted.

Since $\mathcal{P}(A)$ is well-founded and $\mathcal{P}(B)$ embeds into it, $\mathcal{P}(B)$ is well-founded, and therefore B is finite. □

Solution to Exercise 3.8 (Return to exercise)

Proof. If $B \subseteq A$ is a subset without a minimal element, let $S = A \setminus B$. We claim that S satisfies the property in question. Let $a \in A$ be such that for all $b < a$, $b \in S$. If $a \notin S$, then by definition $a \in B$. However, in that case it must be a minimal element of B, since any $b < a$ is in S, i.e. not in B. As B does not have minimal elements, that case is impossible, so $a \notin B$ and therefore $a \in S$. By our assumption on $\leq$ it must be that $S = A$, so $B = \varnothing$, and therefore any non-empty subset has a minimal element, which is to say that $\leq$ is a well-founded relation on A. □

Solution to Exercise 3.10 (Return to exercise)

Proof. We prove this by finite induction, with $\varphi(x)$ being "$\mathcal{P}(x)$ is finite". $\mathcal{P}(\varnothing) = \{\varnothing\}$, which is finite since its power set is well-founded (it is $\{\varnothing, \{\varnothing\}\}$ and we can check by hand that it is well-founded). So, $\varphi(\varnothing)$ holds.

Suppose that $\varphi(A)$ holds, that is $\mathcal{P}(A)$ is finite, and let $x \notin A$. We will see that $\mathcal{P}(A \cup \{x\})$ is finite as well. Since $x \notin A$, $A \cap \{x\} = \varnothing$, so $\mathcal{P}(A \cup \{x\}) \cong \mathcal{P}(A) \times \mathcal{P}(\{x\})$ as a pointwise product. In particular, it is well-founded, since both $\mathcal{P}(A)$ is finite, and thus well-founded, and $\mathcal{P}(\{x\})$ is finite (checking by hand that it is well-founded is not hard). By finite induction, φ holds for all finite sets. That is, if A is finite, then $\mathcal{P}(A)$ is finite. □

Solution to Exercise 3.15 (Return to exercise)

Proof. Let us prove this by finite induction. This is trivial for $\varnothing$ and for singletons. Suppose this holds for B and let $A = B \cup \{a\}$ for some $a \notin B$, let $\leq$ be a partial ordering of A. Let us first assume, without loss of generality, that a is a maximal element in this partial order. In this case, $\leq \cap B \times B = \leq_B$ is a partial ordering of B, so by the induction hypothesis, there is some linear order $\preceq_B$ which extends it. We simply add a as a maximum to this order. Namely, $\preceq_A = \preceq_B \cup \{\langle b, a\rangle \mid b \in A\}$, or spelled explicitly, $x \preceq y$ if and only if $x, y \in B$ and $x \preceq y$ or $x \in B$ and $y = a$. It is not hard to see that this is indeed a linear ordering as wanted.

Suppose now that a was not a maximal element of $\langle A, \leq\rangle$. Let $a' \in A$ be a maximal element, which exists since A is finite, and consider the function $f\colon A \to A$ defined by

$$f(x) = \begin{cases} a & x = a' \\ a' & x = a \\ x & \text{Otherwise.} \end{cases}$$

, using this f we can define $\leq^*$ on A given by $x \leq^* y$ if and only if $f(x) \leq f(y)$. It is not hard to verify that this is a partial order, and indeed f is an isomorphism between the two partial orders. Now a is indeed a maximal element of $\leq^*$, so by the previous case there is a linear ordering $\preceq^*$ which extends it, and defining $x \preceq y$ if and only if $f(x) \preceq^* f(y)$ defines a linear ordering which extends $\leq$ as wanted. □

Remark

This is a case where induction "on the size of A" would be significantly easier, where we can just remove the maximal element and then put it back into the top as we did in the first case. We can simplify the above proof by splitting B into three parts: "above a", "below a", and "incomparable with a". Then extend ordering of B, and argue that the "above" part must be entirely above the "below" part, so we can insert a between them anywhere we want and completely disregard the part incomparable with a.

Solution to Exercise 3.18 (Return to exercise)

Proof. We claim that $\langle n, 0\rangle$, for $n > 0$, is a limit point of $\mathbb{N} \times \mathbb{N}$. Suppose that $\langle i, j\rangle <_{\text{Lex}} \langle n, 0\rangle$, then $i < n$ or $i = n$ and $j < 0$. Since $0 = \min \mathbb{N}$, it must be that $i < n$. Therefore, the successor of $\langle i, j\rangle$, which is $\langle i, j + 1\rangle$, is also below $\langle n, 0\rangle$ in the lexicographic order. So $\langle n, 0\rangle$ is a limit point for all $n > 0$. On the other hand, every $n \in \mathbb{N}$ is either 0 or a successor. Since isomorphism preserves the property of being a successor and a limit, the two cannot be isomorphic. □

Solution to Exercise 3.24 (Return to exercise)

Proof. Let us define $G\colon B^{\subseteq A} \to B$ simply as $G(f) = \min B \setminus \operatorname{rng} f$ if f is not surjective and $\min B$ otherwise. Then $F\colon A \to B$ given by the recursion theorem satisfies that $F(\min A) = \min B$, and so on. The argument in the proof of the Comparison Theorem shows that F is the wanted function.

In the case of definition by cases, we define $G_s\colon B \to B$ simply as the successor function. Namely, $G_s(b) = b'$ if it exists, or b itself if $b = \max B$. For G_l we take the same G as above, and for $F(\min A)$ we take $\min B$. □

Solution to Exercise 3.25 (Return to exercise)

Proof. Let f, g be two isomorphisms $A \to B$. Let $a \in A$ be the least such that $f(a) \neq g(a)$. Without loss of generality, $b_0 = f(a) <_B g(a) = b_1$. Since f and g are isomorphisms, then there is some a_0 such that $g(a_0) = b_0$ and since g is an isomorphism, $a_0 <_A a$, by minimality of a, $f(a_0) = g(a_0) = b_0$, which is impossible since $f(a_0) < f(a) = b_0$. So no such a can exist, and therefore $f = g$. □

Solution to Exercise 3.27 (Return to exercise)

Proof. By the comparison theorem, either $\mathbb{N}$ is isomorphic to an initial segment of A (proper or the entire order), in which case we are done, or else A is isomorphic to a proper initial segment of $\mathbb{N}$. Let $T \subseteq \mathbb{N}$ be the set $\{n \mid I_{\mathbb{N}}(n) \text{ embeds into } A\}$, since A is infinite, $T = \mathbb{N}$, and so it is impossible that A is isomorphic to a proper initial segment of $\mathbb{N}$. □

Solutions to Chapter 4

Solution to Exercise 4.2 (Return to exercise)

Proof. Let $X \in \mathcal{P}(A)$ and let $x \in X$. Since $X \in \mathcal{P}(A)$, $X \subseteq A$, and therefore $x \in A$. Since A is transitive, $x \subseteq A$ and therefore $x \in \mathcal{P}(A)$. □

Solution to Exercise 4.5 (Return to exercise)

Proof. If $\alpha \in \alpha$, then $\in$ is not a strict linear order on α, and so not a well-order of α. □

Solution to Exercise 4.10 (Return to exercise)

Proof. If A is an ordinal, then it is a transitive set by definition. We will show that if $\xi \in A$, then ξ is an ordinal. Suppose that this is not the case, then $\{\xi \in A \mid \xi \text{ is not an ordinal}\}$ is non-empty and therefore has a minimal element, ξ. Since ξ is the smallest counterexample, if $\xi' \in \xi$, then ξ' is an ordinal. Since $\xi' \in A$, and A is transitive, $\xi' \subseteq A$ and if $\xi'' \in \xi'$, then it must be that ξ'' is an ordinal, otherwise it would be the smallest counterexample. This is because $\xi'' \in \xi$ as well. However, this means that ξ is a transitive set which is well-ordered by $\in$, as $\xi \subseteq A$ by the transitivity of A. In other words, ξ is an ordinal after all, so it was not a counterexample.

In the other direction, if A is a set of ordinals, then A is naturally well-ordered by $\in$. Therefore, if A is transitive it is a transitive set which is well-ordered by $\in$. In other words, A is an ordinal. □

Solution to Exercise 4.16 (Return to exercise)

Proof. Since every $X \in \mathcal{F}$ is inductive, then $\varnothing \in X$ for all $X \in \mathcal{F}$. Therefore $\varnothing \in \bigcap \mathcal{F}$. Suppose that $a \in \bigcap \mathcal{F}$, then for every $X \in \mathcal{F}$, $a \in X$. Since every such X is inductive, $a \cup \{a\} \in X$ as well. In other words, for all $X \in \mathcal{F}$, $a \cup \{a\} \in X$, so $a \cup \{a\} \in \bigcap \mathcal{F}$. □

Solution to Exercise 4.17 (Return to exercise)

Proof. Since ω is the smallest inductive set, it is enough to show that if δ is a limit ordinal, then δ is an inductive set. If we show that, then $\omega \subseteq \delta$ which means that $\omega \leq \delta$ as wanted. Let δ be a limit ordinal.

Since δ is a limit ordinal, $\delta > 0$ and so $\varnothing \in \delta$. Next, suppose that $\alpha \in \delta$, then $\alpha \cup \{\alpha\}$ is its ordinal successor. Since $\alpha < \delta$, then by the definition of a successor $\alpha \cup \{\alpha\} \leq \delta$, but since $\alpha \cup \{\alpha\}$ is a successor ordinals, and therefore not a limit, it must be that $\alpha \cup \{\alpha\} \neq \delta$, so $\alpha \cup \{\alpha\} \in \delta$, and therefore δ is an inductive set as wanted. □

Solution to Exercise 4.20 (Return to exercise)

Proof. Let $\alpha = \omega$ and let $\beta = 1$, then $1 + \omega = \sup\{1 + n \mid n < \omega\}$ by the recursive definition of ordinal addition. Since $\{1 + n \mid n < \omega\} = \omega \setminus \{0\}$, it means that $1 + \omega = \omega$. On the other hand, $\omega + 1 \neq \omega$, since $\omega \in \omega + 1$. □

Solution to Exercise 4.21 (Return to exercise)

Proof. Let $\delta > 0$ be a limit ordinal, then δ does not have a maximal element as a linearly ordered set. In particular, if $\alpha < \delta$, then α is not maximal below δ and therefore $\alpha + 1 < \delta$ as well.

In the other direction, if $\delta > 0$ is not a limit ordinal, then δ is a successor ordinal, namely $\delta = \alpha + 1$ for some $\alpha < \delta$. In particular, $\alpha + 1 \geq \delta$. □

Solution to Exercise 4.24 (Return to exercise)

Proof. Define $n \prec m$ if and only if n is even and m is odd or if $n \equiv m \pmod 2$ and $n < m$. In other words, we "move" the odd numbers to be on top of the even numbers. We define $f \colon \mathbb{N} \to \omega + \omega$ to be the function

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even,} \\ \omega + \frac{n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Note that f is well-defined, since if n is odd, $n - 1 \in \mathbb{N}$ and is even, so $\frac{n-1}{2}$ is a natural number.

We claim that f is an order preserving bijection between $\langle \mathbb{N}, \prec \rangle$ and $\omega + \omega$. First, suppose that $n \prec m$, we will show that $f(n) < f(m)$. If n is even and m is odd, then $f(n) < \omega$ and $f(m) \geq \omega$, so $f(n) < f(m)$. In the case where both have the same parity, this is even more immediate: if both are even, then $f(n) = \frac{n}{2} < \frac{m}{2} = f(m)$, and if both are odd, then $\frac{n-1}{2} < \frac{m-1}{2}$, in which case $f(n)$ embeds as a proper initial segment into $f(m)$ and therefore $f(n) < f(m)$. Next, assume that $f(n) < f(m)$, we will show that $n \prec m$. If $f(n) < \omega \leq f(m)$, then by the definition of f we have that n is even and m is odd, in which case $n \prec m$. If $f(n)$ and $f(m)$ are both finite ordinals, then $n = 2f(n) < 2f(m) = m$, and therefore $n \prec m$. Similarly, if both are infinite, then $f(n) = \omega + \frac{n-1}{2}$ and similarly for m, and so $n < m$ and therefore $n \prec m$ as both are odd.

Now that we have shown that f is an order embedding, it remains to show that it is a surjective function. Indeed, if $\alpha < \omega+\omega$ then either $\alpha < \omega$ in which case $f(\alpha+\alpha) = \alpha$, as $\alpha+\alpha$ is an even natural number; if $\omega \leq \alpha$, then $\alpha = \omega + k$ for some $k < \omega$, and by letting $n = 2k+1$ we have that $f(n) = \alpha$ as wanted. □

Solutions to Chapter 5

Solution to Exercise 5.6 (Return to exercise)

Proof. For each $B \subseteq A$ let $\chi_B \colon A \to 2$ be the function given by $\chi_B(a) = \begin{cases} 1 & \text{if } a \in B, \\ 0 & \text{if } a \notin B. \end{cases}$

We will show that $F(B) = \chi_B$ is a bijection between $\mathcal{P}(A)$ and 2^A, and therefore $\mathcal{P}(A) \sim 2^A$.

Suppose that $B, C \subseteq A$, if $B \neq C$, then $B \triangle C \neq \varnothing$. In this case, let $a \in A$ be such that $a \in B \triangle C$. Then $\chi_B(a) \neq \chi_C(a)$, so $\chi_B \neq \chi_C$, therefore F is an injective function.

Suppose that $f \colon A \to 2$ is a function, let $B = f^{-1}(1) = \{a \in A \mid f(a) = 1\}$. We claim that $\chi_B = f$. Indeed, for all $a \in A$, $\chi_B(a) = 1$ if and only if $f(a) = 1$, and therefore $\chi_B(a) = 0$ if and only if $f(a) = 0$. So $\chi_B = f$ as wanted, and therefore F is a surjection. □

Solution to Exercise 5.10 (Return to exercise)

Proof. We define surjections from ω onto $\mathbb{Z}$ and $\mathbb{Q}$.

We define $f \colon \omega \to \mathbb{Z}$ given by $f(n) = (-1)^{p(n)} \frac{n-p(n)}{2}$, where $p \colon \omega \to 2$ is the parity function, namely $p(n) = 0$ when n is an even integer and $p(n) = 1$ when n is odd. Note that $n - p(n)$ is always an even integer, so $\frac{n-p(n)}{2}$ is always a natural number. Given any $k \in \mathbb{Z}$, if $k < 0$, then letting $n = |k|2 + 1$, where $|k|$ is the absolute value of k, we have that n is odd, so $p(n) = 1$, therefore $f(n) = -\frac{n-1}{2} = k$. If $k \geq 0$, then taking $n = 2k$ we get that $f(n) = k$ again. So f is a surjection, and therefore $\mathbb{Z}$ is countable. (It is not hard to check that f is in fact a bijection, too!)

We define $g \colon \omega \to \mathbb{Q}$ by first defining $g_0 \colon \omega \times (\mathbb{Z} \setminus \{0\}) \to \mathbb{Q}$, and then composing it with a surjection $g_1 \colon \omega \to \omega \times (\mathbb{Z} \setminus \{0\})$. Then function $g_0(n, m) = \frac{n}{m}$, and easily any rational number can be expressed as the ratio of a natural number and a non-zero integer (positive or negative), so g_0 is surjective. Since $\mathbb{Z} \setminus \{0\}$ is countable and since the product of two countable sets is countable, $\omega \times (\mathbb{Z} \setminus \{0\})$ is countable, so there is a surjection $g_1 \colon \omega \to \omega \times (\mathbb{Z} \setminus \{0\})$. Setting $g = g_1 \circ g_0$ we get that $g \colon \omega \to \mathbb{Q}$ is a surjective function as wanted. □

Solution to Exercise 5.11 (Return to exercise)

Proof. If A is finite, we already saw that there is an embedding of A into $\mathbb{Q}$. So we may assume that A is countably infinite. Let $A = \{a_n \mid n < \omega\}$ and $\mathbb{Q} = \{q_n \mid n < \omega\}$ be enumerations of the two sets.

We define by recursion $f \colon A \to \mathbb{Q}$. Let $A_n = \{a_i \mid i < n\}$ and suppose that $f \colon A_n \to \mathbb{Q}$ was defined and it was an embedding of orders, where A_n is ordered by $\prec$. Using Exercise 3.10, there is an embedding of $A_{n+1} = A_n \cup \{a_n\}$ into $\mathbb{Q}$ which extends the embedding defined on A_n. Let us define $f(a_n) = q_k$ where k is the least index of a rational number for which the definition gives an embedding.

Let f be the function given by the recursion, then f is an embedding since if $a_n, a_m \in A$, then both are in $A_{\max\{n,m\}+1}$ and f was defined there so that $a_n < a_m \iff f(a_n) < f(a_m)$. $\square$

Solution to Exercise 5.12 (Return to exercise)

Proof. Note that if $f \in \omega^{<\omega}$, then $f \subseteq n \times \omega$ for some $n < \omega$, and moreover $|f| = n$, by the obvious bijection: $g(i) = \langle i, f(i)\rangle$. Therefore $f \in \operatorname{fin}(\omega \times \omega)$. In other words, $\omega^{<\omega} \subseteq \operatorname{fin}(\omega \times \omega)$. Since $\omega \times \omega$ is countable, $\operatorname{fin}(\omega \times \omega)$ is countable as well, so $\omega^{<\omega}$ is countable.

Alternatively, we can let $\{p_n \mid n < \omega\}$ be the set of prime numbers and we can consider $F\colon \omega^{<\omega} \to \omega$ given by $F(f) = \prod_{i \in \operatorname{dom} f} p_i^{f(i)+1}$, then this F is injective, since given any two natural numbers they have a unique decomposition into prime numbers.

Alternatively alternatively, by induction each ω^n is countable, and this is a countable union of countable sets. This, of course, uses the Axiom of Choice. However, we can avoid that by also defining injections $e_n\colon \omega^n \to \omega$ by recursion given by a fixed $g\colon \omega \times \omega \to \omega$, and noting that $\omega^{n+1} \sim \omega^n \times \omega$. This provides us with a sequence of enumerations of each ω_n so we can now inject $\omega^{<\omega}$ into $\omega \times \omega$ by mapping $f\colon n \to \omega$ to $\langle n, e_n(f)\rangle$. $\square$

Solution to Exercise 5.16 (Return to exercise)

Proof. Recall that $\mathbb{R} \sim \mathcal{P}(\omega) \sim 2^\omega$. Therefore $\mathbb{R}^\omega \sim (2^\omega)^\omega \sim 2^{\omega\times\omega} \sim 2^\omega \sim \mathbb{R}$. $\square$

Solution to Exercise 5.17 (Return to exercise)

Proof. Suppose that α is not an initial ordinal, then there is some $\beta < \alpha$ and an injective function $f\colon \alpha \to \beta$. Since $\beta \subsetneq \alpha$, this means that $f\colon \alpha \to \alpha$ is an injective function that is not a bijection. Therefore, by Theorem 3.10 it cannot be that α is a finite set, so it is not a finite ordinal. Therefore, every finite ordinal is an initial ordinal.

In the case of ω, there is no injective function from an infinite set into a finite set, and since ω is the least infinite ordinal, it must be an initial ordinal. $\square$

Solution to Exercise 5.20 (Return to exercise)

Proof. We saw that if A is a finite set and $f\colon A \to A$ is injective, then f is a bijection. Therefore, if $\alpha < \omega$ and $\beta \leq \alpha$, then a bijection $f\colon \alpha \to \beta$ would be an injection $f\colon \alpha \to \alpha$, but then f is onto α, so $\alpha = \beta$.

Next, let us show that if $\alpha > \omega$ and α is a successor ordinal, then α is not an initial ordinal. Let β be such that $\alpha = \beta + 1$, we define the following function:

$$g(\xi) = \begin{cases} 0 & \text{if } \xi = \beta, \\ \xi + 1 & \text{if } \xi < \omega, \\ \xi & \text{otherwise.} \end{cases}$$

We claim that $g\colon \alpha \to \beta$ is a bijection. If $g(\xi_0) = g(\xi_1) = \zeta$, then if $\zeta = 0$, it must be that $\xi_0 = \xi_1 = \beta$; if $0 < \zeta < \omega$, then $\xi_0 = \xi_1 = \zeta - 1$; and otherwise $\xi_0 = \zeta = \xi_1$. And given any $\zeta < \beta$, if $\zeta = 0$, then $g(\beta) = \zeta$; if $0 < \zeta < \omega$, then $g(\zeta - 1) = \zeta$; and if $\zeta \geq \omega$, then $g(\zeta) = \zeta$.

Finally, $\omega + \omega$ is the order type of $\omega \times 2$ and we saw that this is a countable set, so $\omega + \omega$ is a limit ordinal but it is not an initial ordinal. $\square$

Solution to Exercise 5.23 (Return to exercise)

Proof. First we can get simplify the expression and get rid of $\aleph_2$.

$$\aleph_3^{\aleph_5} \leq \aleph_3^{\aleph_5} \cdot \aleph_2 \leq \aleph_3^{\aleph_5} \cdot \aleph_3 = \aleph_3^{\aleph_5+1} = \aleph_3^{\aleph_5}.$$

Next, we can simplify better. Since $2 < \aleph_3 < \aleph_5$, we get that $2^{\aleph_5} \leq \aleph_3^{\aleph_5} \leq \aleph_5^{\aleph_5} = 2^{\aleph_5}$. Therefore, the simplest form is $2^{\aleph_5}$. □

Solution to Exercise 5.24 (Return to exercise)

Proof. If $f\colon \lambda \to \kappa$, then $f \subseteq \lambda \times \kappa$ and $|f| = \lambda$. Since $\lambda \leq \kappa$ we have that $|\lambda \times \kappa| = \kappa$. Let $c\colon \lambda \times \kappa \to \kappa$ be some bijection, then $C\colon \kappa^\lambda \to [\kappa]^\lambda$ given by $C(f) = c[f]$ is an injective function.

Next, for $f \in \kappa^\lambda$, let $R(f) = \operatorname{rng}(f)$ if f was injective, otherwise $\operatorname{rng}(f) \cup \lambda$. Then every subset of κ of size λ is the image of an injective function whose domain is λ, and therefore $R\colon \kappa^\lambda \to [\kappa]^\lambda$ is surjective as wanted. □

(In order to conclude there is a bijection we need to be able to choose an injective function for each set, and this requires the Axiom of Choice.)

Solutions to Chapter 6

Solution to Exercise 6.1 (Return to exercise)

Proof. We prove this by Finite Induction. If $\mathcal{F} = \varnothing$, then $\varnothing$ is a choice function by vacuity. Suppose that $\mathcal{F}$ admits a choice function and A is a non-empty set such that $A \notin \mathcal{F}$. By the induction hypothesis, $\mathcal{F}$ admits a choice function, f, and let $a \in A$ be some element. Then $f \cup \{\langle A, a\rangle\}$ is a choice function from $\mathcal{F} \cup \{A\}$. Therefore, by Finite Induction every finite family of non-empty sets admits a choice function. □

Solution to Exercise 6.2 (Return to exercise)

Proof. Since the real numbers are linearly ordered, every finite subset has a minimum element. Given a family of finite subsets of $\mathbb{R}$, simply choose $\min A$ for each non-empty, finite $A \subseteq \mathbb{R}$. □

Solution to Exercise 6.5 (Return to exercise)

Proof. Let us use the Teichmüller–Tukey Lemma to prove that a maximal chain exists. Let $\mathcal{F} = \{C \subseteq P \mid C \text{ is a chain in } \langle P, \leq\rangle\}$. If we can show that $\mathcal{F}$ has a finite character, then by the Teichmüller–Tukey Lemma it has a $\subseteq$-maximal element, which is by definition and maximal chain.

Indeed, suppose that C is a chain, then every subset of C is also a chain, in particular every finite subset of C. In the other direction, if C is not a chain, then there are $p, q \in C$ which are not comparable, in which case $\{p, q\} \in \operatorname{fin}(C)$ and $\{p, q\}$ is not a chain, and so not in $\mathcal{F}$. Therefore $\langle P, \leq\rangle$ has a maximal chain.

To show that the Axiom of Choice follows from this principle (also known as *Hausdorff's Maximal Principle*), let us show that Zorn's Lemma follows from it. Suppose that $\langle P, \leq\rangle$ is a partial order where every chain has an upper bound. By the principle, there is a maximal chain, $C \subseteq P$, and by the assumption on P, C has an upper bound c. Let us show that c is a maximal element.

Firstly, note that c is comparable with all the elements in C and therefore by the maximality of C, $c \in C$. Suppose that for $p \in P$ we have $c \leq p$, then p is comparable with all the elements of C as well, so $p \in C$. But since c is an upper bound, $p \leq c$, so by antisymmetry, $c = p$. Therefore c is maximal. $\square$

Solution to Exercise 6.6 (Return to exercise)

Proof. If $\aleph_1 \not\leq |A|$, then $|A| < \aleph_1$. By definition, if $|A| < \aleph_1$, then A injects into ω, as $\aleph_1$ is the cardinal of ω_1, the set of all countable ordinals. Therefore, if A is uncountable, $|A| \not< \aleph_1$, and therefore $\aleph_1 \leq |A|$. $\square$

Solution to Exercise 6.10 (Return to exercise)

Proof. Suppose that $\mathcal{F} \cup \{X\} \subseteq \mathcal{U}$, where $\mathcal{U}$ is some filter. Then for any $Y, Z \in \mathcal{U}$, $Y \cap Z \in \mathcal{U}$, so $Y \cap Z \neq \varnothing$. In particular, any $Y \in \mathcal{F}$ and X must satisfy that $X \cap Y \neq \varnothing$.

In the other direction, suppose that $X \cap Y \neq \varnothing$ for all $Y \in \mathcal{F}$. We define

$$\mathcal{U} = \{Y \subseteq A \mid \text{There exists } Y' \in \mathcal{F} \text{ such that } X \cap Y' \subseteq Y\},$$

and we claim that $\mathcal{U}$ is a filter which extends $\mathcal{F} \cup \{X\}$. Clearly, if $Y \in \mathcal{F}$, then $Y \cap X \subseteq Y$, as well as $Y \cap X \subseteq X$, so $X \in \mathcal{U}$ as well. It remains to show that $\mathcal{U}$ is a filter on A.

If $Y \in \mathcal{U}$ and $Y \subseteq Z$, then there is some $Y' \in \mathcal{F}$ such that $X \cap Y' \subseteq Y$ and therefore $X \cap Y' \subseteq Z$, so $Z \in \mathcal{F}$. If $Y, Z \in \mathcal{U}$, then there is some $Y', Z' \in \mathcal{F}$ such that $Y' \cap X \subseteq Y$ and $Z' \cap X \subseteq Z$, therefore $Y' \cap Z' \cap X \subseteq Y \cap Z$ and since $\mathcal{F}$ is a filter, $Y' \cap Z' \in \mathcal{F}$ as well, so $Y \cap Z \in \mathcal{U}$. Finally, Clearly, $A \in \mathcal{U}$, but if $\varnothing \in \mathcal{U}$ then for some $Y \in \mathcal{F}$, $X \cap Y \subseteq \varnothing$, but no such Y exists by our assumption, so $\varnothing \notin \mathcal{U}$. $\square$

Solution to Exercise 6.12 (Return to exercise)

Proof. Suppose that $\bigcap \mathcal{F} \in \mathcal{F}$, then $\bigcap \mathcal{F} \neq \varnothing$, so let $a \in \bigcap \mathcal{F}$ be some element. By the definition of $\bigcap \mathcal{F}$, for all $X \in \mathcal{F}$, $a \in X$. For every $b \neq a$, if $\{b\} \in \mathcal{F}$, then $a \in \{b\}$ which is impossible. So $\bigcap \mathcal{F} = \{a\}$ and therefore $\mathcal{F} = \mathcal{F}_a$. $\square$

Solution to Exercise 6.15 (Return to exercise)

Proof. Let $\mathcal{F}$ be an ultrafilter extending $\mathcal{F}_{\text{fin}}$. Such ultrafilter exists by the previous theorem. By a previous exercise, $\mathcal{F}$ must be free, since it is an ultrafilter and $\mathcal{F}_{\text{fin}} \subseteq \mathcal{F}$. $\square$

Solutions to Chapter 7

Solution to Exercise 7.1 (Return to exercise)

Proof. Every atomic term has the form x_i for a free variable or c, the constant symbol. Therefore the set of atomic terms is $\{x_n \mid n < \omega\} \cup \{c\}$ which is a countably infinite set. The atomic formulas are of the form $t = t'$ where t and t' are terms, as well as $R(t_0, t_1, t_2)$ where t_i are terms. Let us show that there are countably many terms, and therefore $\{t = t' \mid t, t' \in \text{Term}\}$ is in bijection with $\text{Term} \times \text{Term}$, which is a countable set; $\{R(t_0, t_1, t_2) \mid t_0, t_1, t_2 \in \text{Term}\}$ is in bijection with $\text{Term} \times \text{Term} \times \text{Term}$, which is also a countable set, so the set of atomic formulas is the union of two countable sets and is therefore countable. And so, both sets have cardinality $\aleph_0$.

To see that $|\text{Term}| = \aleph_0$, it is enough to show that $|\text{Term}| \leq \aleph_0$, as a lower bound was establish by the set of atomic terms. Note that every term is a finite sequence of symbols, and therefore $\text{Term} \subseteq (\{F, c\} \cup \{x_n \mid n < \omega\})^{<\omega}$. We have seen that $\omega^{<\omega}$ is a countable set, and therefore Term is a countable set as wanted. □

Solution to Exercise 7.3 (Return to exercise)

Proof. We prove that by induction on the structure of the formula. So, first we need to prove that if t is a term, then $\sigma_0(t) = \sigma_1(t)$ assuming that σ_i agree on the value of all the free variables appearing in t.

If t is an atomic term, then either t is a constant symbol, c, in which case $\sigma_i(t) = c^M$ for $i < 2$, or else t is a free variable so by the assumption $\sigma_0(t) = \sigma_1(t)$.

Next, if $\sigma_0(t_i) = \sigma_1(t_i)$ for some terms $t_0, \ldots, t_{n-1}$ and F is an n-ary function symbol, then

$$\begin{aligned}\sigma_0(F(t_0, \ldots, t_{n-1})) =& F^M(\sigma_0(t_0), \ldots, \sigma_0(t_{n-1})) \\ &= F^M(\sigma_1(t_0), \ldots, \sigma_1(t_{n-1})) = \sigma_1(F(t_0, \ldots, t_{n-1})).\end{aligned}$$

We can now prove the same for formulas. For atomic formula, $M \models_{\sigma_0} t = t'$ if and only if $\sigma_0(t) = \sigma_0(t')$, by the above we get that $\sigma_0(t) = \sigma_1(t)$ and $\sigma_0(t') = \sigma_1(t')$, so $M \models_{\sigma_0} t = t'$ if and only if $M \models_{\sigma_1} t = t'$. Similarly for atomic formulas of the form $R(t_0, \ldots, t_{n-1})$ where R is an n-ary relation symbol. Suppose that $M \models_{\sigma_0} \varphi \wedge \psi$, then $M \models_{\sigma_0} \varphi$ and $M \models_{\sigma_0} \psi$, by the induction hypothesis $M \models_{\sigma_1} \varphi$ and $M \models_{\sigma_1} \psi$, so $M \models_{\sigma_1} \varphi \wedge \psi$. The proof for the rest of the connectives and negation is similar.

Finally, suppose that $M \models_{\sigma_0} \exists x \varphi$, then there is some m such that $M \models_{\sigma_0[x/m]} \varphi$. By the induction hypothesis, $M \models_{\sigma_1[x/m]} \varphi$, since any free variable in φ (except for x) has the same value as σ_0 and therefore the same as σ_1, and $\sigma_i[x/m]$ both assign x the value m. So $M \models_{\sigma_1} \exists \varphi$. The proof for $\forall x \varphi$ is similar. □

Solution to Exercise 7.5 (Return to exercise)

Proof. The axioms for a partial order are:

Reflexive $\forall x(x < x)$.

Anti-symmetric $\forall x \forall y(x < y \wedge y < x \to x = y)$.

Transitive $\forall x \forall y \forall z(x < y \wedge y < z \to x < z)$.

We can replace Reflexivity by $\forall x \neg(x < x)$ to obtain the strict version. To obtain a linear order we add:

Linearity $\forall x \forall y(x < y \vee y < x \vee x = y)$.

We will write the properties in the version that works for strict partial orders.

1. $\exists x \forall y(x \neq y \to x < y)$.
2. $\forall x \exists y(x < y)$.
3. $x < y \wedge \forall z \neg(x < z \wedge z < y)$. □

Solution to Exercise 7.6 (Return to exercise)

Proof. Let φ_n be $\exists x_0 \dots \exists x_{n-1}(\bigwedge_{i<j<n} x_i \neq x_j)$. If $M \models \varphi$ then there are $m_0, \dots, m_{n-1} \in M$ such that $M \models \bigwedge_{i<j<n} m_i \neq m_j$, and therefore the function $f(i) = m_i$ is an injective function from n to M. □

Solution to Exercise 7.8 (Return to exercise)

Proof. Let us define the interpretation function on N:

1. If c is a constant symbol, $c^N = f(c^M)$.
2. If F is an n-ary function symbol, $F^N(f(x_0), \dots, f(x_{n-1})) = f(F^M(x_0, \dots, x_{n-1}))$.
3. If R is an n-ary relation symbol, $R^N = \{\langle f(x_0), \dots, f(x_{n-1})\rangle \mid \langle x_0, \dots, x_{n-1}\rangle \in R^M\}$.

Note that F^N is well-defined since f is a bijection so given any $y_0, \dots, y_{n-1} \in N$, there are $x_0, \dots, x_{n-1} \in M$ such that $f(x_i) = y_i$ for all $i < n$. Now, by the very definition of the interpretation given above, f is an isomorphism of $\mathcal{L}$-structures. □

Solution to Exercise 7.10 (Return to exercise)

Proof. We let $\exp\colon \mathbb{R} \to \mathbb{R}^+$ be the function $\exp(x) = e^x$. Then exp is a bijective function and by the exponentiation laws, $\exp(x+y) = e^{x+y} = e^x \cdot e^y = \exp(x) \cdot \exp(y)$, so it is indeed an embedding and in fact an isomorphism, as exp is a bijection.

To find an embedding that is not an isomorphism we have to use the Axiom of Choice. Note that $\langle \mathbb{R}, +\rangle$ is in fact a vector space over $\mathbb{Q}$. Fix some basis $B \subseteq \mathbb{R}$ for this vector space, then $|B| = 2^{\aleph_0}$. Therefore, taking any subset $B' \subseteq B$ such that $|B'| = |B|$ the $\operatorname{span}_{\mathbb{Q}}(B')$ is isomorphic to $\langle \mathbb{R}, +\rangle$. Fixing such $B' \subseteq B$ and an isomorphism $f\colon \mathbb{R} \to \operatorname{span}_{\mathbb{Q}}(B')$ we can then consider $\exp \circ f\colon \mathbb{R} \to \mathbb{R}^+$ which is a composition of two embeddings. However, since f is not a bijection, $\exp \circ f$ is not a bijection either. □

Solution to Exercise 7.13 (Return to exercise)

Proof.

1. $\forall x \forall y \forall z(x < y \wedge y < z \to x < z) \wedge \forall x(\neg(x < x)) \wedge \forall x \forall y(x = y \vee x < y \vee y < x)$.
2. $\forall x \forall y(x < y \to \exists z(x < z \wedge z < y))$.
3. $\forall x \exists y \exists z(y < x \wedge x < z)$. □

Solution to Exercise 7.14 (Return to exercise)

Proof. We extend the two tuples into enumerations of the rational numbers, and we proceed with the back-and-forth argument. Note that by the way we defined it, up to step n, $F(p_i) = q_i$, as wanted. □

Solution to Exercise 7.15 (Return to exercise)

Proof. To see that T is consistent, it is enough to show that given any structure N and φ, it is impossible that $N \models \varphi$ and $N \models \neg\varphi$. If we can show that, then, since $M \models T$, if T was inconsistent, then $M \models \varphi \wedge \neg\varphi$, but that would be impossible. However, note that by the definition of $N \models \varphi$ we get that $N \models \varphi$ if and only if $N \not\models \neg\varphi$. So if $N \models \varphi$ and $N \models \neg\varphi$, then $N \models \varphi$ and also $N \not\models \varphi$ which is impossible.

To see that T is complete, suppose that φ is any $\mathcal{L}$-sentence, then either $M \models \varphi$ or $M \models \neg\varphi$. Without loss of generality, let us assume that $M \models \varphi$, then $\varphi \in T$, so if $N \models T$, it must be that $N \models \varphi$. Therefore $T \models \varphi$. □

Solution to Exercise 7.16 (Return to exercise)

Proof. Suppose that T is a consistent and complete theory and $M \models T$. If $T \models \varphi$, then $M \models \varphi$ by the assumption that $T \models \varphi$ and $M \models T$. If $T \not\models \varphi$, then $T \models \neg\varphi$, and therefore $M \models \neg\varphi$, so $M \not\models \varphi$. □

Solution to Exercise 7.18 (Return to exercise)

Proof. Suppose that $M \models \exists x \forall y(x = y)$, then $M = \{m\}$. Therefore, if $N \models \exists x \forall y(x = y)$ it must be that $N \models M$, as both are singletons. Therefore, if φ is any sentence, if $\{m\} \models \varphi$, then every model of $\exists x \forall y(x = y)$ must satisfy φ, so the theory is complete.

In contrast, $\exists x(x = x)$ simply suggests that M is not the empty set, so it holds in both $\{m\}$ and $\{a, b\}$ where $a \neq b$. However φ_2 from Problem 3 above holds in exactly one of them. □

Solutions to Chapter 8

Solution to Exercise 8.2 (Return to exercise)

Proof. Assume that T is complete and let M, N be two models of T. If $M \models \varphi$, then it must be that $T \models \varphi$, otherwise by the completeness of T, $T \models \neg\varphi$ in which case $M \models \neg\varphi$. Therefore $N \models \varphi$, since $N \models T$. Similarly, if $N \models \varphi$ the same argument shows that $M \models \varphi$. Therefore $M \models \varphi$ if and only if $N \models \varphi$ so the two models are elementarily equivalent.

Assume that any two models of T are elementarily equivalent, if $T \not\models \varphi$, then there is some $M \models T$ such that $M \models \neg\varphi$. If $N \models T$, then by the assumption, $M \equiv N$, so $N \models \neg\varphi$, and so $T \models \neg\varphi$, and is therefore complete. □

Solution to Exercise 8.7 (Return to exercise)

Proof. There are many possible solutions. For example, consider $\mathcal{L} = \{<\}$ where $<$ is a binary relation symbol. Let T be the following theory:

ψ_1: $\forall x \neg(x < x) \land \forall x \forall y \forall z(x < y \land y < z \to x < z)$

ψ_2: φ_1

ψ_3: $\varphi_1 \to \varphi_2$

ψ_4: $\varphi_3 \to \varphi_4$

ψ_5: $\varphi_5 \to \forall x \exists y(x < y)$

Here φ_n are the sentences which state that there are at least n distinct objects in the model. The theory, therefore, states that $<$ is a strict partial order, that there is at least one element in the universe; but if there is one, then there are at least two; and if there are three, then there are at least four; and if there are five distinct objects, then $<$ has no maximal elements, which we saw is a property that cannot hold for finite models.

In other words, a model of T must have exactly two elements or exactly four elements. Finally, T has infinite models, e.g. ω. □

Solution to Exercise 8.8 (Return to exercise)

Proof. Recall that c_m is the function such that $c_m(i) = m$ for all $i \in I$. First, let us see that j is an embedding. For readability, let us denote by N the structure $M^I/\mathcal{U}$.

If c is a constant symbol, then $c^N = [c_{c^M}]_{\mathcal{U}} = j(c^M)$. If $F^M(\bar{m}) = y$, then by definition $F^N(j(\bar{x})) = j(y)$ as $\{i \mid M \models F^M(\bar{x}) = y\} = I$, and similarly for any relation symbol.

The embedding is elementary, because $M \models \varphi(\bar{m})$ if and only if $\{i \in I \mid M \models \varphi(\bar{c}_{\bar{m}}(i))\} = I$ if and only if $N \models \varphi(j(\bar{m}))$. □

Solution to Exercise 8.10 (Return to exercise)

Proof. Suppose that $[f]_{\mathcal{U}}$ is an upper bound for $j[\mathbb{R}]$. We will find some g such that

$$\mathbb{R}^\omega/\mathcal{U} \models [g]_{\mathcal{U}} < [f]_{\mathcal{U}} \text{ and } j(r) < [g]_{\mathcal{U}} \text{ for all } r \in \mathbb{R}.$$

Let $g\colon \omega \to \mathbb{R}$ be the function $g(n) = f(n) - 1$. Easily, we get that $\mathbb{R}^\omega/\mathcal{U} \models [g]_{\mathcal{U}} < [f]_{\mathcal{U}}$, since $\{n < \omega \mid g(n) < f(n)\} = \omega \in \mathcal{U}$.

On the other hand, if for every $r \in \mathbb{R}$, $j(r) < [f]_{\mathcal{U}}$, then in particular, $j(r+1) < [f]_{\mathcal{U}}$. Therefore $\{n < \omega \mid r+1 < f(n)\} \in \mathcal{U}$, however, $r+1 < f(n)$ if and only if $r < f(n) - 1 = g(n)$. So $\{n < \omega \mid r < g(n)\} \in \mathcal{U}$, and so $j(r) < [g]_{\mathcal{U}}$. Therefore there is no least upper bound for $j[\mathbb{R}]$, despite the set being bounded.

Note that we only used the order of $\mathbb{R}$ in the properties of the structure (we did not rely on $+$ and $-$ being defined in $\mathbb{R}^\omega/\mathcal{U}$ in the definition of g or in the proof that $[g]_{\mathcal{U}}$ is an upper bound for $j[\mathbb{R}]$). So, indeed, the *completeness property* of $\langle \mathbb{R}, < \rangle$ which states that every non-empty set which is bounded from above has a supremum must fail in $\mathbb{R}^\omega/\mathcal{U}$, and so it is not first-order. □

Solution to Exercise 8.11 (Return to exercise)

Proof. Let $\mathcal{L}$ be the language with $\{c_\alpha \mid \alpha < \kappa\}$ as constant symbols. Let T be the theory $\{c_\alpha \neq c_\beta \mid \alpha < \beta < \kappa\}$. Note that T has a model by compactness, since any finite subtheory only mentions finitely many constant symbols and can therefore be interpreted in a suitable finite set. If $M \models T$, then $i(\alpha) = c_\alpha^M$ is an injective function from κ into M. □

Solutions to Chapter 9

Solution to Exercise 9.2 (Return to exercise)

Proof. We need to check that M_0 is a substructure of M_1. Since $M_i \prec N$, it must be that any constant symbol c would satisfy $c^N = c^{M_i}$, and similarly $F^N(\bar{m}) = F^{M_i}(\bar{m})$ whenever $m_i \in M_0$, as well as $R^N \cap M_0^n = R^N \cap M_1^n \cap M_0^n$ when R is an n-ary relation symbol.

To see that M_0 is an elementary substructure, given any $\bar{m} \in M_0$ we have that $M_0 \models \varphi(\bar{m})$ if and only if $N \models \varphi(\bar{m})$ since $M_0 \prec N$, but that holds if and only if $M_1 \models \varphi(\bar{m})$, since $M_1 \prec N$ and $m_i \in M_1$ for $i < n$. So $M_0 \prec M_1$.

The proof that if $M_0 \prec M_1$ and $M_1 \prec N$, then $M_0 \prec N$ is similar. If $\bar{m} \in M_0$, then $M_0 \models \varphi(\bar{m})$ if and only if $M_1 \models \varphi(\bar{m})$ if and only if $N \models \varphi(\bar{m})$. □

Solution to Exercise 9.4 (Return to exercise)

Proof. Let $\mathcal{L}$ be the language with a single constant symbol c, and let $c^{\omega_2} = \omega_1$. Then if $\alpha \prec \omega_2$ it must be that $\omega_1 < \alpha$. Note, by the way, that by the Tarski–Vaught Criterion, the elementary substructures of ω_2 are of the form $A \cup \{\omega_1\}$ where A is an infinite set. □

Solution to Exercise 9.6 (Return to exercise)

Proof. The theory of graphs is simply $\forall x \forall y (x\ E\ y \to y\ E\ x \wedge \neg(x\ E\ x))$. Let us write the randomness, which requires a schema of axioms: For every $n, m < \omega$ we write the following sentence where $\bar{x}$ is an n-tuple and $\bar{y}$ is an m-tuple,

$$\forall \bar{x} \forall \bar{y} \exists z \left(\bigwedge_{i<n, j<m} (x_i \neq y_j \wedge z \neq x_i \wedge z \neq x_j) \wedge \bigwedge_{i<n} z\ E\ x_i \wedge \bigwedge_{j<m} \neg(z\ E\ y_j) \right).$$

□

Solution to Exercise 9.8 (Return to exercise)

Proof. Let $\varphi(x)$ be the formula defining a in M and let $\psi(x, y)$ be the formula defining A when setting $y = a$ as a parameter. Namely, $A = \{m \in M \mid M \models \psi(m, a)\}$.

We define $\psi'(x) = \exists y (\varphi(y) \wedge \psi(x, y))$. Then $M \models \psi'(m)$ if and only if there exists y such that $M \models \varphi(y)$ and $\psi(m, y)$ holds. However, since φ defines a in M, it must be that $y = a$, so $\psi'(x)$ holds if and only if $\psi(x, a)$ holds, which happens if and only if $x \in A$. □

Solution to Exercise 9.10 (Return to exercise)

Proof. Let φ and ψ be the formulas which define A and B respectively. Then $m \in A \cup B$ if and only if $M \models \varphi(m)$ or $M \models \psi(m)$, and therefore $\varphi \vee \psi$ defines $A \cup B$. Similarly, $A \cap B$ is defined by $\varphi \wedge \psi$ and $M \setminus A$ is defined by $\neg\varphi$. □

Solution to Exercise 9.11 (Return to exercise)

Proof. Note that $n < m$ if and only if there is some k such that $n + k = m$ and $n \neq m$. In other words, $\varphi(x, y)$ given by $\exists z (x + z = y) \wedge x \neq y$ defines the standard order. □

Solution to Exercise 9.13 (Return to exercise)

Proof. Consider the automorphism of $\langle \mathbb{Z}, < \rangle$ given by $f(k) = k + 1$. This is indeed an automorphism since f is a bijection, and $n < m$ if and only if $n + 1 < m + 1$. Suppose that $\varphi(x, y, z)$ defined addition, namely $\langle \mathbb{Z}, < \rangle \models \varphi(x, y, z)$ if and only if $x + y = z$. Then $\langle \mathbb{Z}, < \rangle \models \varphi(f(x), f(y), f(z))$, since f is an automorphism. However, in that case we have that $z + 1 = (x + 1) + (y + 1) = x + y + 2 = z + 2$, which is impossible. □

Solution to Exercise 9.14 (Return to exercise)

Proof. First, let us see that every $n < \omega$ is definable. But this is easy since n satisfies the formula $c_n = n$. The structure itself cannot have any automorphism which is not the identity, since each c_n must be fixed, and the only element that is not interpreted as a constant is ω itself. So using automorphisms of this structure is not going to be helpful in showing that ω is undefinable.

Let $\mathcal{L}_n$ be the language where we only have the first n constant symbols. Suppose that $\varphi(x)$ was a formula that defined ω. Since φ is a formula, it must be finite, therefore it can only mention finitely many constant symbols, so for some n, $\varphi(x)$ is really a formula in $\mathcal{L}_n$. In this case, the formula would define ω in $\omega + 1$ as a $\mathcal{L}_n$-structure as well. However, in this case we can find an automorphism moving ω to $n + 1$, so φ did not define ω after all.

Another proof goes by using the Upward Löwenheim–Skolem to obtain some M such that $\omega + 1 \prec M$ and $|M| > \aleph_0$. Pick some $m \in M$ which is not in $\omega + 1$, then consider the automorphism which exchanges ω with m. Since it preserves all the constant symbols, it is indeed an automorphism, but since ω was moved, it is not definable in M, and therefore not definable in $\omega + 1$ by elementarity. □

Solutions to Chapter 10

Solution to Exercise 10.2 (Return to exercise)

Proof. Both of these are well-ordered sets without a maximum. Let T be the second-order theory of well-orders without a maximum element, that is:

φ_1: $\forall x \neg(x < x) \land \forall x \forall y \forall z(x < y \land y < z \to x < z)$

φ_2: $\forall x \exists y(x < y)$

φ_3: $\forall A(\exists x(x \in A) \to \exists x(x \in A \land \forall z(z \in A \to z = x \lor x < z)))$

To characterise $\omega + \omega$, note that it is the unique well-ordered set without a maximum which has exactly one limit point. Let $\varphi_{\text{lim}}(x)$ be the formula stating that x is a limit point, which in the case of well-ordered sets can be phrased in multiple equivalent ways. For example,

$$\exists z(z < x) \land \forall y(y < x \to \exists z(y < z \land z < x))$$

which states that x is not the minimum element and if $y < x$, then there is some z between them, so x is not a successor. So, to characterise $\omega + \omega$ we can add to T the axiom

$$\exists x(\varphi_{\text{lim}}(x) \land \forall y(\varphi_{\text{lim}}(y) \to x = y)).$$

Similarly, $\omega \cdot \omega$ is the unique well-ordered set without a maximum whose limit points have order type ω. This, again, can be described in many different ways. For example we can add these two axioms:

$$\forall x \exists y(x < y \land \varphi_{\text{lim}}(y)),$$

and

$$\forall x(\forall y(\varphi_{\text{lim}}(y) \to x < y) \lor \exists y((y < x \lor x = y) \land \varphi_{\text{lim}}(y) \land \forall z(z < x \land \varphi_{\text{lim}}(z) \to z < y \lor z = y))).$$

The first simply states that there is no largest limit point. The second axiom states that any point is either below all the limit points (i.e., it is a finite ordinal) or else it has a last limit point below it. □

Solution to Exercise 10.4 (Return to exercise)

Proof. We saw that there is a second-order formula which states that a set is finite, φ_{fin}, and there is a second-order formula which states that a set is countably infinite, φ_ω. Therefore, a set is of size $\aleph_1$ if and only if

$$\forall A(\varphi_{\text{fin}}(A) \vee \varphi_\omega(A) \vee \exists F(\operatorname{dom} F = A \wedge \forall x \forall y \forall z(F(x,y) \wedge F(x,z) \rightarrow y = z) \wedge \forall y \exists x F(x,y))).$$

In other words, every set A is finite, countable, or has a surjection onto the entire structure. Alternatively, we could have posited the existence of a well-ordered set which is not countable but every proper initial segment is countable. □

Solution to Exercise 10.5 (Return to exercise)

Proof. We can use the language $\{<, c\}$ where $<$ is a binary relation symbol and c is a constant. Building on the proof of Theorem 10.6, we can simply consider the axioms $\neg\varphi_n(c)$ along with the theory which characterises ω. Every finite subtheory of this theory has a model, since any finitely many axioms only have finitely many axioms of the form $\neg\varphi_n(c)$, so by interpreting c as a large enough natural number provides a model of the finite subtheory. However the theory itself does not have a model, as the model would have to be isomorphic to ω, but also c cannot be interpreted as any of the elements.

We can, however, replace the infinite list of axioms, $\{\neg\varphi_n(c) \mid n < \omega\}$ by its conjunction, $\bigwedge_{n<\omega} \neg\varphi_n(c)$. This is a single $\mathcal{L}_{\omega_1,\omega}$ sentence. Now, on the other hand, the theory is finite and does not have any models. □

Solution to Exercise 10.9 (Return to exercise)

Proof. Note that $M \models \forall^\infty x\varphi(x)$ if and only if $\{m \in M \mid M \models \neg\varphi(m)\}$ is a finite set. Therefore this holds if and only if for some $n < \omega$ there are exactly n members of M for which φ is false.

Given $n < \omega$ we can write the sentence

$$\varphi_n := \exists x_0 \ldots \exists x_{n-1} \forall y(\bigwedge_{i<n} y \neq x_i \rightarrow \varphi(y)),$$

note that in the case that $n = 0$ this might be ill-formed, so we just let $\varphi_0 = \forall y\varphi(y)$. Now, using the infinitary part of $\mathcal{L}_{\omega_1,\omega}$ we have that $\forall^\infty x\varphi(x)$ holds if and only if $\bigvee_{n<\omega} \varphi_n$ holds. □

Solution to Exercise 10.10 (Return to exercise)

Proof. Note that φ_{lim} from the solution of Problem 1 is a first-order formula in the language $\{<\}$. We can therefore write the following theory:

φ_1: $\forall x\neg(x < x) \wedge \forall x \forall y \forall z(x < y \wedge y < z \rightarrow x < z) \wedge \forall x \forall y(x < y \vee y < x \vee x = y)$

φ_2: $\forall x \exists y(x < y \wedge \varphi_{\text{lim}}(y))$

φ_3: $\forall x \forall^\infty y(\varphi_{\text{lim}}(y) \rightarrow x < y)$

φ_4: $\forall x \exists y((y < x \vee y = x) \wedge (\varphi_{\text{lim}}(y) \vee \forall z(y < z \vee y = z)) \wedge \forall^\infty z(z < x \rightarrow z < y))$

The theory states that $<$ is a strict linear order, every point in the order has a limit point above it, given any x, all but finitely many limit points lie above it (so below x only finitely many limit points exist), and given any x there is a $y \leq x$ which is a limit point or the minimum element, such that only finitely many points lie between x and y.

If M is a model of this theory, if M is not well-ordered, then there is an infinite sequence which is descending in the order (so, a function $f\colon \omega \to M$ such that $f(n+1) < f(n)$). If there is infinitely many points which do not have a limit point between them, then φ_4 is false. If every point in the sequence has a limit point with just finitely many points of between them, then φ_3 is false, since there are infinitely many limit points below any given element of the decreasing sequence of elements.

Since M is well-ordered, by φ_2 it must be a limit of limit ordinals, so $\operatorname{otp}(M) \geq \omega \cdot \omega$, but if $\operatorname{otp}(M) > \omega \cdot \omega$, then there is a point which has infinitely many limit points below it, so φ_3 must be false in that case. Therefore $M \cong \omega \cdot \omega$. □

Solution to Exercise 10.12 (Return to exercise)

Proof. Let $\mathcal{F}_M = \{M\}$ be the trivial filter on a set M, then it is easy to see that $\forall^{\mathcal{F}_M}$ is $\forall$.

For $\forall^\infty$ we take the filter $\{A \subseteq M \mid M \setminus A \text{ is finite}\}$, i.e. $\mathcal{F}_{\text{fin}}$. □

Solutions to Chapter 11

Solution to Exercise 11.2 (Return to exercise)

Proof. Take the set $X = \{x_0, \dots, x_n\}$, then X is non-empty, so there is some $x \in X$ such that $x \cap X = \varnothing$. For all $i < n$, $x_{i+1} \in x_i \cap X$, so it must be that $x_n \cap X = \varnothing$. In particular, $x_0 \notin x_n$. □

Solution to Exercise 11.3 (Return to exercise)

Proof. Suppose that $\{x, \{x, y\}\} = \{a, \{a, b\}\}$, we want to show that $x = a$ and $b = y$. Since $x \in \{x, y\}$, it must be the case that $x \neq \{x, y\}$, otherwise $x \in x$, so if $x = a$ it must be that $\{x, y\} = \{a, b\}$ and therefore $y = b$. If $x \neq a$, then it means that $x = \{a, b\}$ and $a = \{x, y\}$. However, this means that $x \in a \in x$ which is a contradiction to the previous exercise. □

Solution to Exercise 11.5 (Return to exercise)

Proof. It is easy to see that (2) implies (3) which implies (4) by using Theorem 4.24, which implies (2) by Theorem 3.3 (noting that the embedding here is the identity).

Assume the Axiom of Foundation, and let x be a non-empty set. If $y \subseteq x$ is a non-empty set, then by the Axiom of Foundation, there is some $z \in y$ such that $z \cap y = \varnothing$. That is, z is an $\in$-minimal element in y. Therefore x is a well-founded set.

In the other direction, assuming (2), let x be a non-empty set, then x is well-founded, and therefore x has an $\in$-minimal element, which is some $y \in x$ such that for any $z \in x$, $z \notin y$. In other words, $y \cap x = \varnothing$. So the Axiom of Foundation holds. □

Solution to Exercise 11.9 (Return to exercise)

Proof. Note that $\operatorname{rank}(u) < \alpha$ if and only if $u \in V_\alpha$. Therefore, if $\operatorname{rank}(x) = \alpha$ and $y \in x$, then $\operatorname{rank}(y) < \alpha$. Therefore $\operatorname{rank}(x) \geq \sup\{\operatorname{rank}(y) + 1 \mid y \in x\}$. In the other direction, since α is the least such that $x \subseteq V_\alpha$, it means that α is the least such that $\operatorname{rank}(y) + 1 \leq \alpha$ for all $y \in x$, and therefore $\operatorname{rank}(x) \leq \sup\{\operatorname{rank}(y) + 1 \mid y \in x\}$. □

Solution to Exercise 11.11 (Return to exercise)

Proof. We use the previous exercise to compute the rank.

$$\begin{aligned}
\operatorname{rank}(\langle x, y\rangle) &= \operatorname{rank}(\{\{x\}, \{x, y\}\}) \\
&= \max\{\operatorname{rank}(\{x\}) + 1, \operatorname{rank}(\{x, y\}) + 1\} \\
&= \max\{\alpha + 2, \max\{\alpha + 1, \beta + 1\} + 1\} \\
&= \max\{\alpha + 2, \beta + 2\} \\
&= \max\{\alpha, \beta\} + 2.
\end{aligned}$$

□

Solution to Exercise 11.12 (Return to exercise)

Proof. Let us compute $\operatorname{rank}(\omega_1^{\omega_6})$ first. Given any $f\colon \omega_6 \to \omega_1$, the elements of f are ordered pairs of the form $\langle \alpha, f(\alpha)\rangle$ where $\alpha < \omega_6$ and $f(\alpha) < \omega_1$. Since $\operatorname{rank}(\alpha) = \alpha$, by the previous exercise, $\operatorname{rank}(\langle \alpha, f(\alpha)\rangle) = \alpha + 2$ for all $\alpha \geq \omega_1$. Therefore, $\operatorname{rank}(f) = \sup\{\alpha + 3 \mid \alpha < \omega_6\} = \omega_6$, and therefore $\operatorname{rank}(\omega_1^{\omega_6}) = \sup\{\operatorname{rank}(f) + 1 \mid f\colon \omega_6 \to \omega_1\} = \sup\{\omega_6 + 1\} = \omega_6 + 1$.

Next, computing $\operatorname{rank}(\omega_6^{\omega_1})$, if $f\colon \omega_1 \to \omega_6$, then there is some $\omega_1 < \alpha < \omega_6$ such that $f\colon \omega_1 \to \alpha$, and by a similar calculation as before, $\operatorname{rank}(f) \leq \alpha + 3$. Of course, for each $\alpha < \omega_6$, the constant function $c_\alpha\colon \omega_1 \to \{\alpha\}$ has rank of at least α. Therefore $\operatorname{rank}(\omega_6^{\omega_1}) = \sup\{\operatorname{rank}(f) + 1 \mid f\colon \omega_1 \to \omega_6\} = \sup\{\alpha + 4 \mid \alpha < \omega_6\} = \omega_6$. □

Solution to Exercise 11.17 (Return to exercise)

Proof. Note that $V_\alpha = \bigcup\{V_\beta \mid \beta < \alpha\} = \bigcup\{V_{\alpha_n} \mid n < \omega\}$, and similarly $L_\alpha = \bigcup\{L_{\alpha_n} \mid n < \omega\}$. Since $L_\alpha \subseteq V_\alpha$ by the way we define the two sets, it is enough to show inclusion in the other direction. Note that $\alpha_{n+1} = \min\{\beta \mid V_{\alpha_n} \subseteq L_\beta\}$, so $V_{\alpha_n} \subseteq L_{\alpha_{n+1}} \subseteq V_{\alpha_{n+1}}$. So $V_\alpha = L_\alpha$. □

Solutions to Chapter 12

Solution to Exercise 12.1 (Return to exercise)

Proof. Suppose that $x \in H_\kappa$, then $|\operatorname{tcl}(x)| < \kappa$. We know that $\operatorname{tcl}(x)$ can be mapped onto its rank, and therefore $\operatorname{rank}(x) = \operatorname{rank}(\operatorname{tcl}(x)) < \kappa$, so $x \in V_\kappa$. Therefore $H_\kappa \subseteq V_\kappa$, so it is a set. To see that it is transitive, if $x \in H_\kappa$ and $y \in x$, then $\operatorname{tcl}(y) \subseteq \operatorname{tcl}(x)$, so $y \in H_\kappa$. □

Solution to Exercise 12.3 (Return to exercise)

Proof. Note that $t = \operatorname{tcl}(a) \cup \operatorname{tcl}(b)$ is a transitive set, as the union of two transitive sets, and $|t| < \kappa$. Next, $t' = t \cup \{\{x, y\} \mid x, y \in t\}$ is a transitive set, since if $u \in t'$, then either $u \in t$ and so $u \subseteq t$, or else $u = \{x, y\}$ for some $x, y \in t$. By cardinal arithmetic, $|t'| < \kappa$ as well. Next, consider $t'' = t' \cup \{\{x, y\} \mid x, y \in t'\}$, then by the same argument as before t'' is transitive and $|t''| < \kappa$. Finally, $a \times b \subseteq t''$ and therefore $\operatorname{tcl}(a \times b) \subseteq t''$, so $a \times b \in H_\kappa$. □

Solution to Exercise 12.5 (Return to exercise)

Proof. If κ is regular, this is easy since $|\alpha+\omega| = |\alpha| < \kappa$ for every infinite ordinal below κ, the set of limit ordinals must be unbounded below κ and therefore has size κ. If κ is a singular cardinal, then $\sup\{\lambda^+ \mid \lambda < \kappa\} = \kappa$, and therefore

$$\{\delta < \kappa \mid \delta \text{ is a limit ordinal}\} = \bigcup_{\lambda<\kappa} \{\delta < \lambda^+ \mid \lambda \text{ is a limit ordinal}\}.$$

Therefore the cardinality of the set is $\sup\{\lambda^+ \mid \lambda < \kappa\} = \kappa$ as wanted. □

Remark

While the proof relies on the Axiom of Choice (in proving that κ^+ is a regular cardinal for all κ), the statement is in fact true without appealing to the Axiom of Choice. The proof is slightly more intricate and technical.

Solution to Exercise 12.6 (Return to exercise)

Proof. Let $\varphi(x, y)$ denote the following formula: if $x \subseteq \omega\times\omega$ and x is a well-ordering of $\operatorname{dom} x$, then $y = \operatorname{otp}(x)$; otherwise $x = y$. Then $V_{\omega_1} \models \forall x \exists! y \varphi(x, y)$. If x is not a well-ordering of some subset of ω, then $x = y$, so y exists and is unique. If x is in fact a well-ordering of a subset of ω, then its order type is uniquely determined and it is a countable ordinal, so it is in V_{ω_1}.

If $\alpha < \omega_1$, then there is an injection $f\colon \alpha \to \omega$, and so $x = \{\langle f(\beta), f(\gamma)\rangle \mid \beta < \gamma < \alpha\}$ is a well-ordering of a subset of ω (specifically, $\operatorname{rng}(f)$) that has order type α. This holds except for the case of $\alpha = 1$ where $x = \varnothing$ has order type 0. Regardless, $\{y \mid \exists x \subseteq \omega\times\omega, V_{\omega_1} \models \varphi(x, y)\}$ is a set of rank ω_1, so the Axiom of Replacement fails in V_{ω_1}, since $\mathcal{P}(\omega\times\omega) \in V_{\omega_1}$, but its image under the function defined by φ is not. □

Solution to Exercise 12.8 (Return to exercise)

Proof. It is enough to show that $V_\kappa \subseteq H_\kappa$, as the latter is a model of the Axiom of Replacement, and so $V_\kappa \models \mathsf{ZFC}$. We will show by induction that $|V_\alpha| < \kappa$ for all $\alpha < \kappa$, therefore these are witnesses that if $x \in V_\kappa$, $x \in H_\kappa$.

We saw that V_ω is a countable set. Suppose that $|V_\alpha| = \lambda < \kappa$, then $|V_{\alpha+1}| = 2^\lambda < \kappa$ by the assumption on κ. It remains to check for a limit ordinal α. Suppose that $\alpha < \kappa$ is a limit ordinal and for all $\beta < \alpha$, $|V_\beta| < \kappa$. Since κ is a cardinal, $|\alpha| < \kappa$, and therefore $\{|V_\beta| \mid \beta < \alpha\}$ is a subset of κ of small size, by regularity, there is some $\lambda < \kappa$ such that $|V_\beta| \leq \lambda$ for all $\beta < \alpha$. This allows us to define a injection from V_α into $\lambda\times\alpha$ by choosing $f_\beta\colon V_{\beta+1} \to \lambda$ to be an injection, and then mapping $x \in V_{\beta+1} \setminus V_\beta$ to $\langle f_\beta(x), \beta\rangle$, which is easily an injective function, and therefore $|V_\alpha| \leq \max\{\lambda, |\alpha|\} < \kappa$. □

Solution to Exercise 12.9 (Return to exercise)

Proof. We saw that given such κ, there is some $\alpha < \kappa$ such that $V_\alpha \prec V_\kappa$. Repeating the same proof, this time defining $\beta_0 = \alpha + 1$ and letting $M_0 \prec V_\kappa$ be an elementary submodel generated by V_{β_0}, and continuing recursively by defining $M_{n+1} \prec V_\kappa$ to be an elementary submodel generated by $V_{\beta_{n+1}}$, where $\beta_{n+1} = \operatorname{rank}(M_n)$, we have that for $\beta = \sup\{\beta_n \mid n < \omega\}$ we have that $V_\beta = \bigcup\{M_n \mid n < \omega\} \prec V_\kappa$. Since $V_\alpha \subseteq V_\beta$, it follows that $V_\alpha \prec V_\beta$.

To see that the least α and β for which $V_\alpha \prec V_\beta$ must be singular note that if κ is the least regular cardinal for which $V_\kappa \models \mathsf{ZFC}$, then there are such $\alpha < \beta < \kappa$. By elementarity, V_α and V_β are both models of ZFC, so it is impossible for them to be regular by the minimality of κ. □

Solution to Exercise 12.12 (Return to exercise)

Proof. Note that if $\alpha < \omega_1$, then in H_κ there is an injection from α into ω, since $\alpha \times \omega \in H_\kappa$, and therefore any $f\colon \alpha \to \omega$, injective or otherwise, is also in H_κ.

Therefore, in H_κ, ω_1 is the least ordinal for which there is no injection into ω. Since ω is definable, as the least infinite ordinal (for example), ω_1 is also definable. Therefore, if $M \prec H_\kappa$ it follows that $\omega_1 \in M$. So, if M is countable, it cannot be that $\omega_1 \subseteq M$ and therefore M is not transitive. □