B1.2 Set Theory

Martin Bays

HT23 Oxford

Contents

0.1 Acknowledgements

These notes follow the outline of Jonathan Pila's notes for this course, which were themselves originally based on notes of Robin Knight. Other sources include the books of Hils-Loeser, Goldrei, Jech, and Kunen, (see references below), and lecture notes of Itay Kaplan.

1 Introduction

What is a set? One standard informal answer might be: "A set is an unordered collection of objects, called its elements". We can formalise this intuitive idea of the data given by a set as a test for equality:

Principle of Extensionality: Two sets A and B are equal if and only if they have the same elements,

$$A = B \Leftrightarrow \forall x\,(x \in A \leftrightarrow x \in B).$$

But this leaves a trickier question: What sets are there? If we want to say something holds "for all sets A", which A must we consider? It is tempting to give a broad definition, such as:

> *By a "set" we understand any collection to a whole M of specific well-separated objects m (called the "elements" of M) of our intuition or thought.*[1] – (Cantor 1895)

But too broad conceptions of set lead to **paradoxes**.

1.1 Russell's paradox

We might expect:

Unrestricted Comprehension: For any well-defined property $P(x)$, there is a set whose elements are precisely those x which satisfy $P(x)$.

But take $P(x)$ to be the property: x is a set which is not an element of itself. Suppose R is the set whose elements are precisely the x satisfying $P(x)$. Then for any set x,

$$x \in R \Leftrightarrow x \notin x.$$

In particular,

$$R \in R \Leftrightarrow R \notin R.$$

This contradiction shows that Unrestricted Comprehension is **inconsistent**.

1.2 Zermelo-Fraenkel Set Theory

Around the start of the 20th century, Zermelo and (later) Fraenkel developed a version of set theory which avoids Russell's paradox and similar paradoxes.

This **Zermelo-Fraenkel set theory** is the subject of this course. One key feature of this theory is its adoption of the *axiomatic method*: it consists of **axioms**, statements about the universe of sets, intended to suffice to derive all the theorems of ordinary mathematics.

This axiom system is denoted by "ZFC", where C denotes one axiom known as the Axiom of Choice; we also sometimes consider the system "ZF" which omits this axiom.

[1] "Unter einer ‚Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterscheidenen Objekten m uns[e]rer Anschauung oder unseres Denkens (welche die ‚Elementen' von M genannt werden) zu einem Ganzen."

1.3 Foundations

It has become common (though not universal) practice to consider ZFC as forming the *foundations* of mathematics.

Such a position gives in particular a clear answer to the question: what is a proof of a mathematical statement? This proceeds as follows.

Given a mathematical statement S, we first encode it into a formal (first-order) statement σ about sets – remarkably, such encoding appears to always be possible. We then identify the notion of a "proof of S" with the notion of a formal proof of σ from ZFC; the latter has a clear unambiguous definition (presented in B1.1 Logic).

1.4 Consistency

Is ZFC consistent? Could there be some paradox which it fails to avoid?

Sadly, Gödel's 2nd Incompleteness Theorem shows that if ZFC is consistent, then we can not prove (in the above sense, i.e. from ZFC) that it is consistent. Mathematicians and set theorists vary in the extent to which they "believe" that it is consistent.

What we can say for sure is that the mathematics humanity has developed so far has not revealed any inconsistency in ZFC.

1.5 Why study set theory?

(1) Sets are very natural primitive mathematical structures, so are of intrinsic mathematical interest.

(2) As natural mathematical objects, sets arise in many areas of mathematics, so we need to be ready to deal with them.

(3) Since set theory can form a foundation for mathematics, studying foundational issues (e.g. relative consistency of axioms) in set theory suffices for addressing such questions in mathematics as a whole.

This course concentrates on (1) and (2), but also introduces the necessary preliminaries for the Part C course Axiomatic Set Theory, which concentrates on (3).

1.6 Cardinality

One key concern in the study of sets is the size ("cardinality") of a set. Sets can be finite, countable, or uncountable – but there is much more to say than that.

We will define sets X and Y to have the same cardinality if there is a bijection $X \to Y$, and we will see that the ZFC axioms suffice to make this a rich and useful concept, answering in particular questions like:

- Is there a complex number which is not the zero of any integer polynomial?
- How many lines does it take to cover the real plane?
- Can the subsets of $\mathbb{R}$ be exhaustively indexed by real numbers?

1.7 Structure of the course

We will study:

- The axioms of ZFC, introducing them gradually throughout the course.
- Formalisation of mathematics in set theory (concentrating on $\mathbb{N}$).
- Cardinalities.
- Ordinals: These measures of the "length of an infinite process" are important in particular for "transfinite" inductive arguments.
- Axiom of Choice: We study this important axiom in detail, giving a number of equivalent formulations.

2 The first axioms

2.1 Extensionality

ZF1 (Extensionality): For all x and y, x is equal to y if and only if x and y have the same elements:

$$\forall x\, \forall y\, (x = y \leftrightarrow \forall z\, (z \in x \leftrightarrow z \in y)).$$

To make sense of such axioms, we adopt the following way of thinking.

We work in a mathematical *universe* $\mathcal{V}$ consisting of objects, which we call **sets**. When we say "for all x" (written $\forall x$) we mean "for all sets x in the universe $\mathcal{V}$"; similarly $\exists x$ refers to existence in $\mathcal{V}$. Given two sets x and y of $\mathcal{V}$, it may or may not be that $x \in y$ holds; we say that x is an **element of** y when it does. Our axioms are statements using these concepts, and we assume that $\mathcal{V}$ is such that the axioms are true in $\mathcal{V}$. This gives us information about the universe $\mathcal{V}$.[2]

This is how we will discuss the axioms. At first, we know nothing about the universe. As we introduce the axioms, we find out more and more about it.

We are *not* making a philosophical claim that this universe $\mathcal{V}$ we work in is the "real universe" of sets.

The Extensionality axiom ZF1 says that an object a of $\mathcal{V}$ is determined by the information of which objects b of $\mathcal{V}$ satisfy $b \in a$. So we may identify a with the set of such b – this justifies us calling the objects of $\mathcal{V}$ "sets" and reading $\in$ as "is an element of".

Now note that the elements of a set in $\mathcal{V}$ are also sets, as are the elements of its elements, and so on. Such sets are called **hereditary sets**. So the objects of $\mathcal{V}$ are hereditary sets.

There are no cows in our universe $\mathcal{V}$, there are only sets. Nor are there sets of cows in $\mathcal{V}$, there are only sets of sets – which are actually sets of sets of sets, and so on.

From now on, we reserve the word **set** to mean: an object in $\mathcal{V}$.

[2] Those with B1.1 Logic will recognise this as the notion we formalised there of $\mathcal{V}$ being a *model* of the axioms.

2.2 Empty set

We now begin to "populate" our universe $\mathcal{V}$ by giving axioms which guarantee the existence of sets. First off:

ZF2 (Empty Set): There is a set with no elements:

$$\exists x \, \forall y \; y \notin x.$$

We can already deduce that this empty set is unique:

Theorem 2.1. *There is a unique set with no elements.*

Proof. Existence is by the axiom Empty Set. Suppose x and y each have no elements, i.e. $\forall z \; z \notin x$ and $\forall z \; z \notin y$. Then x and y have the same elements, i.e. $\forall z \; (z \in x \leftrightarrow z \in y)$, since both sides of "$\leftrightarrow$" are false for all z. So $x = y$ by Extensionality. □

We denote this unique empty set by $\emptyset$, as usual.

2.3 Pairing

ZF3 (Pairing): For any x and y (not necessarily distinct), there is a set whose elements are precisely x and y:

$$\forall x \, \forall y \, \exists z \, \forall w \; (w \in z \leftrightarrow (w = x \vee w = y)).$$

(Here, $\vee$ means "(inclusive) or"; $P \vee Q$ is true iff at least one of P and Q is.)

Again, we have uniqueness by Extensionality:

Theorem 2.2. *For any x and y there is a unique set whose elements are precisely x and y.*

Proof. Exercise (Sheet 1). □

We denote this unique set by $\{x, y\}$, and by $\{x\}$ in the case that $y = x$.

This is already enough to build up a rich collection of sets: we have the existence of $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset, \{\emptyset\}\}\}$, and so on. But we don't yet know that any sets with more than two elements exist!

End of lecture 1

2.4 Union

ZF4 (Union): For any set x, there is a set whose elements are precisely the elements of the elements of x:

$$\forall x \, \exists y \, \forall z \; (z \in y \leftrightarrow \exists w \; (z \in w \wedge w \in x)).$$

(Here, $\wedge$ means "and".)

By Extensionality again, this set is unique, and we denote it by $\bigcup x$. Note $\bigcup \emptyset = \emptyset$.

We can now define some familiar notation.

Given sets x and y, define $x \cup y := \bigcup \{x, y\}$. This notation is justified by the following easy exercise.

Exercise 2.3. *For any sets x, y, z, we have $z \in x \cup y$ iff $z \in x$ or $z \in y$.*

Given sets x, y, z, define $\{x, y, z\} := \{x, y\} \cup \{z\}$. More generally, recursively define for $n \geq 2$:

$$\{x_1, \ldots, x_{n+1}\} := \{x_1, \ldots, x_n\} \cup \{x_{n+1}\}.$$

Then this behaves as the notation indicates, namely:

$$\forall y \left(y \in \{x_1, \ldots, x_n\} \leftrightarrow \bigvee_{i=1}^{n} y = x_i \right).$$

2.5 Powerset

A set x is a **subset** of a set y, written $x \subseteq y$, if every element of x is an element of y.

$$x \subseteq y \;\Leftrightarrow\; \forall z\, (z \in x \rightarrow z \in y).$$

ZF5 (Powerset): For any set x, there exists a set whose elements are precisely the subsets of x:

$$\forall x\, \exists y\, \forall z\, (z \in y \leftrightarrow z \subseteq x).$$

By Extensionality, this set of all subsets of x is unique. We denote it by $\mathcal{P}(x)$, and call it the **power set** of x.

Example 2.4.

- $\mathcal{P}(\emptyset) = \{\emptyset\}$.
- $\mathcal{P}(\mathcal{P}(\emptyset)) = \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$.
- $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) = \mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

3 Formulas and comprehension

Russell's paradox showed that the unrestricted comprehension principle, the existence of $\{x : P(x)\}$ for any property $P(x)$, is inconsistent.

In ZF, we weaken this principle to comprehension restricted to a set: given a set y, we want the existence of $\{x \in y : P(x)\}$.

However, before we can state this as an axiom, we must formalise a notion of a "property" $P(x)$. To see the problem, consider trying to make sense of:

$$\{n \in \mathbb{N} : n \text{ has no English definition less than a thousand letters long }\}.$$

This would be a non-empty set, since there are only finitely many ways to arrange a thousand letters, so it would have a least element, which would be *the least natural number which has no English definition less than a thousand letters long* – but that is an English definition less than a thousand letters long of this natural number, which is meant to have no such definition; contradiction[3].

To avoid such paradoxes, and to facilitate reasoning about the axioms, we will require P to be expressible in the following formal language of set theory, which B1.1 students will recognise as *first-order logic with a binary relation* $\in$.

[3] This is a version of the "Berry paradox".

3.1 The formal language $\mathcal{L}$

The **formulas** of $\mathcal{L}$ are built up recursively as follows:

- An expression of the form $x = y$ or $x \in y$ is a formula (with x, y any variables).
- If ϕ and ψ are formulas then so are $\neg\phi$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \to \psi)$, and $(\phi \leftrightarrow \psi)$.
- If ϕ is a formula then so are $\forall x\ \phi$ and $\exists x\ \phi$.

Here, $\neg\phi$ is the logical negation of ϕ, true iff ϕ is false, and can be read as "not ϕ". Also, $(\phi \to \psi)$ means "ψ holds if ϕ does", so it is equivalent to $\neg(\phi \wedge \neg\psi)$.

We define some useful abbreviations:

- $x \notin y$ abbreviates $\neg x \in y$.
- $x \subseteq y$ abbreviates $\forall z\ (z \in x \to z \in y)$.
- $\forall x \in y\ \phi$ abbreviates $\forall x\ (x \in y \to \phi)$.
- $\exists x \in y\ \phi$ abbreviates $\exists x\ (x \in y \wedge \phi)$.

An occurrence of a variable in a formula is **free** if it is not in the scope of a quantifier, like the first x in $\exists y\ (y \in x \wedge \forall x\ x \notin y)$. The **free variables** of a formula ϕ are the variables which occur free in ϕ. We often write e.g. $\phi(x, y)$ to denote a formula whose free variables are x and y.

A **sentence** is a formula with no free variables. So a sentence is either true or false in our universe $\mathcal{V}$ of sets.

A formula $\phi(x)$ with one free variable x can be viewed as a property: for a given value of x, it is either true or false.

We have already seen that each of ZF1-4 can be expressed by a single sentence of $\mathcal{L}$. The whole ZFC axiom system will be expressible by (infinitely many) sentences of $\mathcal{L}$. This isn't actually important in this course, and we won't always spell out how it's done, but it will be crucial in the Part C Axiomatic Set Theory course.

A **formula with parameters** is the result of replacing some of the variables in a formula with sets. For example, if a is a set, $\phi(x) := a \in x$ is a formula with parameter a and free variable x, expressing the property of having the set a as an element.

If $\phi(x)$ is a formula with parameters and b is a set, we write $\phi(b)$ for the result of replacing each free occurrence of x in $\phi(x)$ with b, so $\phi(b)$ asserts that b satisfies the property ϕ. Similarly, if y is a variable not occurring in $\phi(x)$, then $\phi(y)$ is the result of replacing each free occurrence of x with y (which in B1.1 was denoted $\phi[y/x]$.)

3.2 Comprehension

ZF6 (Comprehension): For any formula $\phi(x)$ with parameters and any set y, there is a set z whose elements are precisely those elements x of y which satisfy $\phi(x)$:

$$\forall y\ \exists z\ \forall x\ (x \in z \leftrightarrow (x \in y \wedge \phi(x))).$$

By Extensionality, this set is unique, and we denote it by $\{x \in y : \phi(x)\}$.

Remark 3.1. ZF6 can be expressed by $\mathcal{L}$-sentences, as follows.

For each formula $\phi(x, w_1, \ldots, w_n)$,

$$\forall w_1 \ldots \forall w_n \, \forall y \, \exists z \, \forall x \, (x \in z \leftrightarrow (x \in y \wedge \phi(x)))$$

is an $\mathcal{L}$-sentence expressing the instances of ZF6 for those formulas with parameters which result from substituting sets for the variables $w_1, \ldots, w_n$. So we can express ZF6 by the *axiom scheme* consisting of one such sentence for each such formula.

With this restricted version of comprehension, the argument of Russell's paradox does not lead to inconsistency; instead, it proves something interesting.

Theorem 3.2. *There is no set of all sets: there is no set Ω such that $\forall x \; x \in \Omega$.*

Proof. Suppose Ω is such. By Comprehension, consider $R := \{x \in \Omega : x \notin x\}$. Then $R \in \Omega$, so $R \in R$ iff $R \notin R$, contradiction. □

End of lecture 2

Comprehension allows us to implement some more of the usual constructions of set theory.

Lemma 3.3. *Let a be a non-empty set. Then there is a unique set $\bigcap a$ such that for all x,*

$$x \in \bigcap a \Leftrightarrow \forall y \in a \; x \in y.$$

Proof. Uniqueness is by Extensionality. Let

$$\bigcap a := \{x \in \bigcup a : \forall y \in a \; x \in y\},$$

which exists by Comprehension (and Union).

Then this has the desired property: if x is in every element of a then, since $a \neq \emptyset$, x is in some element of a, so $x \in \bigcup a$; so then $x \in \bigcap a$. The converse is immediate. □

We leave $\bigcap \emptyset$ undefined, since it has no sensible definition.

Definition 3.4. For any sets a and b, define:

$$a \cap b := \bigcap \{a, b\}$$
$$a \setminus b := \{x \in a : x \notin b\}.$$

4 Products and relations

In this section, we start to see how some of the usual notions of mathematics can be handled in the set theory we have established so far.

4.1 Cartesian products

If X and Y are sets, we want to be able to consider their Cartesian product $X \times Y$. Its elements should be *ordered* pairs of elements of X and Y, so the first problem is how to encode this notion when all we have are (unordered) sets. For this, we use the following standard coding trick.

Definition 4.1. Given sets x and y, the **ordered pair** with **first co-ordinate** x and **second co-ordinate** y is the set

$$\langle x, y\rangle := \{\{x\}, \{x, y\}\}.$$

The following theorem justifies the terminology.

Theorem 4.2. *For any x, y, x', y', we have $\langle x, y\rangle = \langle x', y'\rangle$ if and only if $x = x'$ and $y = y'$.*

Proof. $\Leftarrow$: Immediate.

$\Rightarrow$: Suppose

$$\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}.$$

First, suppose $y = x$. Then $\{\{x\}, \{x, y\}\} = \{\{x\}\}$, so $\{x', y'\} = \{x\}$, hence $x' \in \{x\}$ and $y' \in \{x\}$, so $x' = x = y'$, and in particular $x = x'$ and $y = y'$, as required.

So we may assume $y \neq x$, and symmetrically $y' \neq x'$.

In particular, $\{x\} \neq \{x', y'\}$, since otherwise $x' = x = y'$. Similarly, $\{x, y\} \neq \{x'\}$.

Hence $\{x\} = \{x'\}$, so $x = x'$, and $\{x, y\} = \{x', y'\}$, so $y' \in \{x, y\}$. Now $y' \neq x$, since $y' \neq x' = x$, so $y' = y$, as required.

□

We can now define the Cartesian product using Powerset and Comprehension.

Proposition 4.3. *Let X and Y be sets. There is a unique set $X \times Y$, called the* ***Cartesian product*** *of X and Y, with the property that the elements of $X \times Y$ are precisely the ordered pairs $\langle x, y\rangle$ where $x \in X$ and $y \in Y$.*

Proof. Uniqueness is by Extensionality. If $x \in X$ and $y \in Y$, then $\langle x, y\rangle = \{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(X \cup Y)$, so $\langle x, y\rangle \in \mathcal{P}(\mathcal{P}(X \cup Y))$. So

$$X \times Y := \{z \in \mathcal{P}(\mathcal{P}(X \cup Y)) : \exists x \in X\ \exists y \in Y\ z = \langle x, y\rangle\}$$

has the desired property – this set exists by Comprehension, since $z = \langle x, y\rangle$ can be expressed in $\mathcal{L}$, namely by

$$\forall w\ (w \in z \leftrightarrow (\forall u\ (u \in w \leftrightarrow u = x) \vee \forall u\ (u \in w \leftrightarrow (u = x \vee u = y)))).$$

□

Notation 4.4. From now on we will allow ourselves to use the $\langle x, y\rangle$ notation in the formulas used in Comprehension, as well as our other defined terms $\{x, y\}$, $\bigcup x$, $\mathcal{P}(x)$, $\emptyset$, and so on. As in the previous proof, it is always possible to eliminate these expressions to write an equivalent formula directly in $\mathcal{L}$. The following example illustrates the general technique (and how much paper we will save with it!):

$$\begin{aligned}
&\forall x\ \left\{x, \bigcup y\right\} \in z \\
&\qquad \Leftrightarrow\ \forall x\ \exists w\ (w = \left\{x, \bigcup y\right\} \wedge w \in z) \\
&\qquad \Leftrightarrow\ \forall x\ \exists w\ (\forall u\ (u \in w \leftrightarrow (u = x \vee u = \bigcup y)) \wedge w \in z) \\
&\qquad \Leftrightarrow\ \forall x\ \exists w\ (\forall u\ (u \in w \leftrightarrow (u = x \vee \forall v\ (v \in u \leftrightarrow \exists t\ (v \in t \wedge t \in y)))) \wedge w \in z)
\end{aligned}$$

4.2 Relations

Definition 4.5. A **binary relation** is a set R of ordered pairs; we then write xRy to mean $\langle x, y\rangle \in R$ (and we use this as an abbreviation in formulas).

Remark 4.6. If $\langle x, y\rangle \in R$ then $x, y \in \bigcup\bigcup R$, since $x, y \in \{x, y\} \in \langle x, y\rangle \in R$ (from which we obtain $x, y \in \{x, y\} \in \bigcup R$ and hence $x, y \in \bigcup\bigcup R$).

Definition 4.7. The **domain** of a binary relation R is the set

$$\mathrm{dom}(R) := \{x : \exists y\ xRy\},$$

and the **range** of R is the set

$$\mathrm{ran}(R) := \{y : \exists x\ xRy\};$$

these sets exist by Comprehension and Remark 4.6. So $R \subseteq \mathrm{dom}(R) \times \mathrm{ran}(R)$.

A binary relation **on** a set X is a binary relation R with $R \subseteq X \times X$, i.e. with $\mathrm{dom}(R) \subseteq X$ and $\mathrm{ran}(R) \subseteq X$.

Note that with this definition, $=$ and $\in$ and $\subseteq$ are *not* relations, since the domain would be a set of all sets, though their restrictions to any given set are relations.

End of lecture 3

4.2.1 Functions

Definition 4.8. A **function** is a binary relation f with the property that for all x, there is at most one y such that $\langle x, y\rangle \in f$.

We write $f(x) = y$ to mean $\langle x, y\rangle \in f$ (and we use this as an abbreviation in formulas).

The **restriction** of f to a set $a \subseteq \mathrm{dom}(f)$ is the function

$$f|_a := f \cap (a \times \mathrm{ran}(f)) = \{\langle x, y\rangle \in f : x \in a\}.$$

The **image** of a set $a \subseteq \mathrm{dom}(f)$ under f is the set

$$f[a] := \mathrm{ran}(f|_a) = \{y : \exists x \in a\ f(x) = y\}.$$

Definition 4.9. Given sets X and Y, a **function from X to Y** is a function f with $\mathrm{dom}(f) = X$ and $\mathrm{ran}(f) \subseteq Y$. We write $f : X \to Y$ for such a function.

So any function f is a function from $\mathrm{dom}(f)$ to $\mathrm{ran}(f)$.

Proposition 4.10. *Given sets X and Y, there is a set Y^X whose elements are precisely the functions from X to Y.*

Proof. Any $f : X \to Y$ is an element of $\mathcal{P}(X \times Y)$, so by Comprehension it suffices to see that the property of a set $f \subseteq X \times Y$ being a function $X \to Y$ is expressible in $\mathcal{L}$. We can express it as follows:

$$\phi(f) := \forall x \in X\ (\exists y \in Y\ f(x) = y \wedge \forall y'\ (f(x) = y' \to y' = y)).$$

□

Remark 4.11. To partially explain the notation Y^X: for finite sets X and Y we have $|Y^X| = |Y|^{|X|}$.

Remark 4.12. The empty set is a function, called the *empty function*, $\emptyset : \emptyset \to \emptyset$. So $Y^\emptyset = \{\emptyset\}$ for any set Y.

4.2.2 Order relations

Definition 4.13.

- A **strict partial order** (or just **strict order**) on a set X is a relation $< \subseteq X \times X$ satisfying for all $x, y, z \in X$:

  Irreflexivity: $\neg x < x$;

  Transitivity: if $x < y$ and $y < z$ then $x < z$.

  It is a strict **total** order if also we have for all $x, y \in X$ that $x < y$ or $y < x$ or $x = y$.

- A **(totally) ordered set** is a set X equipped with a (total) order.

- If an order is denoted by $<$, we write $x \leq y$ as an abbreviation for $(x < y \vee x = y)$, and $x > y$ for $y < x$.

- A **least** element of a subset $Y \subseteq X$ of an ordered set is a $y \in Y$ such that $y \leq y'$ for all $y' \in Y$ (so y is unique if it exists).

- A **minimal** element of a subset $Y \subseteq X$ of an ordered set is a $y \in Y$ such that $y' < y$ for no $y' \in Y$.

Total orders are also known as *linear* orders.

4.2.3 Equivalence relations

Definition 4.14. An **equivalence relation** on a set X is a binary relation $\sim \subseteq X \times X$ satisfying for all $x, y, z \in X$:

Reflexivity: $x \sim x$;

Symmetry: if $x \sim y$ then $y \sim x$;

Transitivity: if $x \sim y$ and $y \sim z$ then $x \sim z$.

The set of **equivalence classes** of $\sim$ is then

$$\begin{aligned} X/{\sim} &:= \{S \in \mathcal{P}(X) : \exists x \in X \; S = \{y \in X : y \sim x\}\} \\ &= \{\{y \in X : y \sim x\} : x \in X\}. \end{aligned}$$

5 The axiom of infinity and the natural numbers

One of the main purposes of set theory is to clarify the nature of infinite objects, but the axioms we have introduced so far do not imply the existence of an infinite set. We also claimed that all of mathematics can be formalised in terms of sets, but so far we don't even know how to treat the natural numbers and its structure. We deal with both of these problems at once.

5.1 Natural numbers – discussion

Informally, we can encode each natural number as a set by defining $n := \{0, \ldots, n-1\}$:

$$0 := \emptyset;\ 1 := \{0\};\ 2 := \{0,1\};\ 3 := \{0,1,2\} \ldots$$

$$0 = \emptyset;\ 1 = \{\emptyset\};\ 2 = \{\emptyset, \{\emptyset\}\};\ 3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \ldots$$

(recall that these sets exist by Pairing and Union). So

$$n+1 = \{0, \ldots, n\} = \{0, \ldots, n-1\} \cup \{n\} = n \cup \{n\}.$$

It would be tempting to then define $\mathbb{N} := \{0,1,2,\ldots\}$. However, we can not directly write an axiom which asserts the existence of such a set ($\exists x\ \forall y\ (y \in x \leftrightarrow (y = 0 \vee y = 1 \vee y = 2 \ldots))$ is an infinite expression, so not an $\mathcal{L}$-sentence); a deeper problem, if we're trying to establish foundations and define the natural numbers starting only from set theory, is that we would already need the natural numbers to make sense of this "..." notation.

Instead, we proceed as follows.

5.2 Inductive sets and the axiom of infinity

Definition 5.1. The **successor** of a set x is $x^+ := x \cup \{x\}$.

Definition 5.2. A set x is **inductive** if $\emptyset \in x$ and x is closed under the successor operation, i.e. $\forall y\ (y \in x \to y^+ \in x)$.

ZF7 (Infinity): An inductive set exists:

$$\exists x\ (\emptyset \in x \wedge \forall y\ (y \in x \to y \cup \{y\} \in x)).$$

Proposition 5.3. *There exists a unique least inductive set; we denote it by* $\mathbb{N}$.

That is, there is a unique set $\mathbb{N}$ *which is inductive and which is a subset of every inductive set.*

End of lecture 4

Proof. Uniqueness is immediate: if $\mathbb{N}$ and $\mathbb{N}'$ are least inductive sets, then $\mathbb{N} \subseteq \mathbb{N}'$ and $\mathbb{N}' \subseteq \mathbb{N}$, so $\mathbb{N} = \mathbb{N}'$ (by Extensionality).

By Infinity (ZF7), there is an inductive set I. Consider

$$\mathbb{N} := \bigcap \{I' \subseteq I : I' \text{ is inductive}\};$$

this is a set by Comprehension (inductivity can be expressed in $\mathcal{L}$ as in the statement of ZF7).

Note that the intersection of inductive sets is inductive. So $\mathbb{N}$ is inductive, and if J is inductive then $J \cap I \subseteq I$ is inductive, so $\mathbb{N} \subseteq J \cap I \subseteq J$. So $\mathbb{N}$ is least inductive. □

Later, we will also use ω to denote this set $\mathbb{N}$.

We want to build up mathematics within our universe of sets, so from now on we define[4]:

Definition 5.4. A **natural number** is an element of $\mathbb{N}$. We use numerals to denote elements of $\mathbb{N}$: $0 = \emptyset$, $1 = 0^+$, and so on.

Theorem 5.5 (Induction on $\mathbb{N}$)**.** *Suppose $\phi(x)$ is an $\mathcal{L}$-formula with parameters such that $\phi(0)$ holds, and if $n \in \mathbb{N}$ and $\phi(n)$ holds then $\phi(n^+)$ holds. Then $\phi(n)$ holds for all $n \in \mathbb{N}$.*

In other words,

$$\Big(\big(\phi(\emptyset) \wedge \forall n \in \mathbb{N}\ (\phi(n) \to \phi(n^+)) \big) \to \forall n \in \mathbb{N}\ \phi(n) \Big).$$

Proof. The assumption on ϕ precisely means that the set $X := \{n \in \mathbb{N} : \phi(n)\} \subseteq \mathbb{N}$ (which exists by Comprehension) is inductive, hence $X = \mathbb{N}$ by definition of $\mathbb{N}$. □

5.3 The order on $\mathbb{N}$

We will use this induction principle to show that we can define in set-theoretic terms the structure we expect to find on $\mathbb{N}$. First we consider the ordering and successor, then we move on to defining the arithmetic operations.

Definition 5.6. Define a binary relation $<$ on $\mathbb{N}$ by: $x < y \ \Leftrightarrow\ x \in y$.

Note that this is indeed a relation, i.e. $\{\langle x, y\rangle \in \mathbb{N} \times \mathbb{N} : x \in y\}$ is a set.

Lemma 5.7. *$<$ is a strict partial order on $\mathbb{N}$.*

Proof.

Transitivity: We prove by induction on $\mathbb{N}$ that

$$\forall n \in \mathbb{N}\ \forall x\ \forall y\ ((x \in y \wedge y \in n) \to x \in n).$$

So let $\phi(n) := \forall x\ \forall y\ ((x \in y \wedge y \in n) \to x \in n)$. Then $\phi(0)$ holds trivially (since $y \in 0$ holds for no y).

[4] To avoid potential confusion: we do not redefine the notion of formula at this point.

Suppose $\phi(n)$ holds; we show that $\phi(n^+)$ holds. So suppose $x \in y \in n^+ = n \cup \{n\}$. If $y = n$, then $x \in n \subseteq n^+$ so $x \in n^+$. Otherwise, $y \in n$, so $x \in y \in n$, and by $\phi(n)$ we obtain $x \in n$, so again $x \in n^+$. Hence $\phi(n^+)$ holds.

By induction on $\mathbb{N}$ (Theorem 5.5), we deduce $\forall n \in \mathbb{N}\ \phi(n)$, as required.

Irreflexivity: We prove $\forall n \in \mathbb{N}\ n \notin n$ by induction on $\mathbb{N}$. Clearly $0 \notin 0$.

Suppose $n \notin n$, but $n^+ \in n^+ = n \cup \{n\}$. Since $n \in n^+$ but $n \notin n$, we have $n^+ \neq n$, so $n^+ \in n$. But then $n \in n^+ \in n$, so $n \in n$ by transitivity, contradicting $n \notin n$.

□

Lemma 5.8. *For all $n, m \in \mathbb{N}$:*

(i) $n^+ \neq 0$.

(ii) If $n \in m$ then $n^+ \in m^+$.

(iii) If $n \neq 0$, then $n = k^+$ for a unique $k \in \mathbb{N}$.

Proof. (i) $n \in n \cup \{n\} = n^+$, so $n^+ \neq 0 = \emptyset$.

(ii) We prove by induction on m that $\forall m \in \mathbb{N}\ \forall n \in m\ n^+ \in m^+$. This is trivial for $m = 0$. Suppose for m, and let $n \in m^+ = m \cup \{m\}$; we conclude by showing $n^+ \in m^{++}$.

If $n = m$, then $n^+ = m^+ \in m^{++}$, as required. Otherwise, $n \in m$, so $n^+ \in m^+$ by the induction hypothesis, so again $n^+ \in m^{++}$ as required.

(iii) Existence: $\forall n \in \mathbb{N}\ (n = 0 \vee \exists k \in \mathbb{N}\ n = k^+)$ holds by a trivial induction (at the successor step, just use $n^+ = n^+$).

Uniqueness: Suppose $k, l \in \mathbb{N}$ and $k^+ = l^+$ but $k \neq l$. Then $k \in k^+ = l^+ = l \cup \{l\}$ but $k \neq l$, so $k \in l$. Then $k^+ \in l^+ = k^+$ by (ii), contradicting irreflexivity.

□

Theorem 5.9. *$<$ is a strict total order on $\mathbb{N}$.*

Proof. Given Lemma 5.7, it remains to show totality, i.e. $\forall n \in \mathbb{N}\ \forall m \in \mathbb{N}\ \phi(n, m)$ where

$$\phi(n, m) := (m \in n \vee m = n \vee n \in m).$$

We first prove by induction on n that $\forall n \in \mathbb{N}\ (0 \in n \vee 0 = n)$. This is immediate for $n = 0$, and if it holds for n, then $0 \in n^+ = n \cup \{n\}$ since either $0 = n \in n^+$ or $0 \in n \subseteq n^+$.

Now let $n \in \mathbb{N}$. We conclude by proving $\forall m \in \mathbb{N}\ \phi(n, m)$ by induction on m. We have $\phi(n, 0)$ by what we proved above.

Now suppose $\phi(n, m)$; we conclude by proving $\phi(n, m^+)$. If $m = n$ then $n \in m^+$, and if $n \in m$ then $n \in m \in m^+$, and then $n \in m^+$ by transitivity.

Otherwise, $m \in n$. In particular, $n \neq 0$. By Lemma 5.8(iii), $n = k^+ = k \cup \{k\}$ for some $k \in \mathbb{N}$. So either $m = k$, in which case $m^+ = k^+ = n$, or $m \in k$, in which case $m^+ \in k^+ = n$ by Lemma 5.8(ii). □

We consider $\mathbb{N}$ as a totally ordered set with this ordering.

Lemma 5.10. *For $n, m \in \mathbb{N}$, we have $n \leq m \Leftrightarrow n \subseteq m$.*

Proof. $\Rightarrow$: If $n = m$ then certainly $n \subseteq m$. If $n < m$ and $k \in n$, then $k \in m$ by transitivity, so again $n \subseteq m$.

$\Leftarrow$: Suppose $n \not\leq m$. By totality, $m < n$, so then $m \subseteq n$ by the previous step. Then $n \not\subseteq m$, since otherwise $n = m$ by Extensionality, contrary to assumption.

□

End of lecture 5

Theorem 5.11. *Any non-empty subset X of $\mathbb{N}$ has a unique least element, denoted* $\min X$.

Proof. Uniqueness given existence is immediate: if n and n' are each least, then $n \leq n' \leq n$, so $n = n'$.

For existence, suppose $X \subseteq \mathbb{N}$ has no least element. We show $X = \emptyset$ by proving $\forall n \in \mathbb{N}\, \forall m \in n\; m \notin X$ by induction. For $n = 0$ this is trivial, and if it holds for n then it holds for n^+, since otherwise n would be a least element of X. □

Remark. By an easy induction, any $n \in \mathbb{N}$ is a subset of $\mathbb{N}$ (i.e. $\mathbb{N}$ is *transitive*) (Exercise Sheet 2). So $n = \{m \in \mathbb{N} : m < n\}$, and we can see this as justifying our informal notation $n = \{0, \dots, n-1\}$.

5.4 Recursion on $\mathbb{N}$

Theorem 5.12 (Definition by recursion on $\mathbb{N}$)**.** *Let X be a set and $g : X \to X$ a function, and let $x_0 \in X$. Then there exists a unique function $f : \mathbb{N} \to X$ such that*

- *$f(0) = x_0$;*
- *$f(n^+) = g(f(n))$ for all $n \in \mathbb{N}$.*

Proof. For $n \in \mathbb{N}$, say $h : n^+ \to X$ is an *n-approximation* if $h(0) = x_0$ and $h(m^+) = g(h(m))$ for all $m \in n$.

Claim. For each $n \in \mathbb{N}$, there exists a unique n-approximation.

Proof. **Existence:** By induction on n. $\{\langle 0, x_0\rangle\}$ is a 0-approximation. If h is an n-approximation, set $h' := h \cup \{\langle n^+, g(h(n))\rangle\}$. Then h' is an n^+-approximation.

Uniqueness: By induction on n. For $n = 0$, $\{\langle 0, x_0\rangle\}$ is the unique 0-approximation. Suppose the uniqueness for n, and let h_1 and h_2 be n^+-approximations. Then $h_1|_{n^+} = h_2|_{n^+}$, since these are n-approximations, and then $h_1(n^+) = g(h_1(n)) = g(h_2(n)) = h_2(n^+)$, so $h_1 = h_2$.

$\square_{Claim}$

We conclude by proving that there exists a unique f as in the statement.

Uniqueness: Suppose $f, f' : \mathbb{N} \to X$ are as in the statement. Let $n \in \mathbb{N}$. Then $f|_{n^+}$ and $f'|_{n^+}$ are n-approximations, so $f|_{n^+} = f'|_{n^+}$ by the uniqueness in the Claim, and in particular $f(n) = f'(n)$. Hence $f = f'$.

Existence: Each n-approximation $h : n^+ \to X$ is a subset of $\mathbb{N} \times X$. The property of being an n-approximation is expressible in $\mathcal{L}$, by translating the definition. So by Comprehension, there is a set

$$H := \{h \in \mathcal{P}(\mathbb{N} \times X) : \exists n \in \mathbb{N}\, [h \text{ is an } n\text{-approximation}]\}.$$

Let $f := \bigcup H$.

We show that f is as required.

- $\underline{f : \mathbb{N} \to X}$: Let $n \in \mathbb{N}$. There exists an n-approximation h, so $\langle n, h(n)\rangle \in f$. If $\langle n, x\rangle \in f$, then $\langle n, x\rangle \in h'$ for some $h' \in H$. Then h' is an m-approximation for some m with $n \leq m$, so $h'|_{n^+}$ is an n-approximation, so $h'|_{n^+} = h|_{n^+}$, and so $x = h(n)$.
- $\underline{f(0) = x_0}$: If $h \in H$ is the 0-approximation, we have $\langle 0, x_0\rangle \in h \subseteq f$.
- $\underline{f(n^+) = g(f(n)) \text{ for } n \in \mathbb{N}}$: Let $h \in H$ be the n^+-approximation. Then $f(n^+) = h(n^+) = g(h(n)) = g(f(n))$.

□

We deduce a version which allows us to treat parameters uniformly:

Corollary 5.13. *Suppose A and X are sets, and $g_0 : A \to X$ and $g_+ : A \times X \to X$ are functions. Then there exists a unique $f : A \times \mathbb{N} \to X$ such that:*

- *$f(a, 0) = g_0(a)$;*
- *$f(a, n^+) = g_+(a, f(a, n))$ for all $n \in \mathbb{N}$.*

Proof. For each a, by Theorem 5.12 there is a unique $f_a : \mathbb{N} \to X$ such that $f_a(0) = g_0(a)$ and $f_a(n^+) = g_+(a, f_a(n))$ for all $n \in \mathbb{N}$. So if we can define $f(a, n) := f_a(n)$, this will be unique with the desired property.

To see that there is such a function f, observe that we can express in a formula that f_a is defined by recursion using $g_0(a)$ and $g_+(a, x)$, so there is a formula $\phi(x, y)$ such that for all $a \in A$, f_a is unique such that $\phi(a, f_a)$ holds. Then ϕ defines by Comprehension the function $F : A \to X^{\mathbb{N}}$; $a \mapsto f_a$, i.e. $F = \{\langle x, y\rangle \in A \times X^{\mathbb{N}} : \phi(x, y)\}$.

Then the function defined by $f(a, n) := F(a)(n)$, i.e.

$$f = \{\langle\langle a, n\rangle, x\rangle \in (A \times \mathbb{N}) \times X : \langle a, \langle n, x\rangle\rangle \in F\},$$

is as required.

More concisely:

$$f = \{\langle\langle a, n\rangle, x\rangle \in (A \times \mathbb{N}) \times X : \exists h \in X^{\mathbb{N}}\ ((h(0) = g_0(a) \wedge \forall m \in \mathbb{N}\ h(m^+) = g_+(a, h(m))) \\ \wedge h(n) = x)\}.$$

□

5.5 Arithmetic on $\mathbb{N}$

Definition 5.14. Define functions $+, \cdot, \widehat{} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ using Corollary 5.13 such that for all $n, m \in \mathbb{N}$

- - $n + 0 = n$
  - $n + m^+ = (n + m)^+$
- - $n \cdot 0 = 0$
  - $n \cdot m^+ = n \cdot m + n$
- - $n\widehat{}0 = 1$
  - $n\widehat{}m^+ = (n\widehat{}m) \cdot n$

We normally write n^m for $n\widehat{}m$. We use the usual operator precedence rules, so $n + m \cdot k$ means $n + (m \cdot k)$ rather than $(n + m) \cdot k$.

Explicitly, to obtain $+$ we apply Corollary 5.13 with $g_0(n) = n$ and $g_+(n, k) = k^+$; then for $*$ we apply it with $g_0(n) = 0$ and $g_+(n, k) = k + n$; then for $\widehat{}$ we apply it with $g_0(n) = 1$ and $g_+(n, k) = k \cdot n$.

Now we use induction to confirm the usual arithmetic properties of these operations.

Theorem 5.15. *For all $n, m, k \in \mathbb{N}$:*

(i) $n + 1 = n^+$.

(ii) $(n + m) + k = n + (m + k)$ *(Associativity of $+$).*

(iii) $0 + k = k$

(iv) $n + k^+ = n^+ + k$

(v) $n + k = k + n$ *(Commutativity of $+$).*

(vi) $n \cdot 1 = n$.

(vii) $n \cdot (m + k) = n \cdot m + n \cdot k$ *(Distributivity of $\cdot$ over $+$).*

(viii) $(n \cdot m) \cdot k = n \cdot (m \cdot k)$ *(Associativity of $\cdot$).*

(ix) $n \cdot k = k \cdot n$ *(Commutativity of $\cdot$).*

(x) $m^{n+k} = m^n \cdot m^k$.

(xi) $m^{n \cdot k} = (m^n)^k$.

Proof. We will be brief. Where we use induction below, it is on k, with $m, n \in \mathbb{N}$ arbitrary.

(i) $n + 1 = n + 0^+ = (n + 0)^+ = n^+$.

(ii)
- Base case: $(n + m) + 0 = n + m = n + (m + 0)$.
- Inductive step: $(n + m) + k^+ = ((n + m) + k)^+ = (n + (m + k))^+ = n + (m + k)^+ = n + (m + k^+)$.

(iii)
- Base case: $0 + 0 = 0$.
- Inductive step: $0 + k^+ = (0 + k)^+ = k^+$.

(iv)
- Base case: $n + 0^+ = (n + 0)^+ = n^+ = n^+ + 0$.
- Inductive step: $n + k^{++} = (n + k^+)^+ = (n^+ + k)^+ = n^+ + k^+$.

(v)
- Base case: $n + 0 = n = 0 + n$, by (iii).
- Inductive step: $n + k^+ = (n + k)^+ = (k + n)^+ = k + n^+ = k^+ + n$, by (iv).

(vi) $n \cdot 1 = n \cdot 0^+ = (n \cdot 0) + n = 0 + n = n$, by (iii).

(vii)
- Base case: $n \cdot (m + 0) = n \cdot m = n \cdot m + 0 = n \cdot m + n \cdot 0$.
- Inductive step: $n \cdot (m + k^+) = n \cdot (m + k)^+ = n \cdot (m + k) + n = (n \cdot m + n \cdot k) + n = n \cdot m + (n \cdot k + n) = n \cdot m + n \cdot k^+$ (using associativity of $+$).

(viii)
- Base case: $(n \cdot m) \cdot 0 = 0 = n \cdot 0 = n \cdot (m \cdot 0)$.
- Inductive step: $(n \cdot m) \cdot k^+ = (n \cdot m) \cdot k + n \cdot m = n \cdot (m \cdot k) + n \cdot m = n \cdot (m \cdot k + m) = n \cdot (m \cdot k^+)$ (using distributivity).

(ix) Exercise.

(x)
- Base case: $m^{n+0} = m^n = m^n \cdot 1 = m^n \cdot m^0$.
- Inductive step: $m^{n+k^+} = m^{(n+k)^+} = m^{n+k} \cdot m = (m^n \cdot m^k) \cdot m = m^n \cdot (m^k \cdot m) = m^n \cdot m^{k^+}$.

(xi)
- Base case: $m^{n \cdot 0} = m^0 = 1 = (m^n)^0$
- Inductive step: $m^{n \cdot k^+} = m^{n \cdot k + n} = m^{n \cdot k} \cdot m^n = (m^n)^k \cdot m^n = (m^n)^{k^+}$.

□

5.6 Defining $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ (not on syllabus)

Although it is not on the course syllabus, we briefly indicate how we can use $\mathbb{N}$ and its arithmetic structure to construct within our universe of sets more of the familiar structures of mathematics.

First, we define $\mathbb{Z}$ as $(\mathbb{N} \times \mathbb{N})/\sim$ where $(n, m) \sim (n', m') \Leftrightarrow n + m' = m + n'$; we identify $(n, m)/\sim$ with $n - m$ and define addition and multiplication correspondingly.

Then we can define $\mathbb{Q}$ as $(\mathbb{Z} \times (\mathbb{N} \setminus \{0\}))/\sim'$ where $(n, m) \sim' (n', m') \Leftrightarrow n \cdot m' = m \cdot n'$; then identify $(n, m)/\sim'$ with $\frac{n}{m} \in \mathbb{Q}$ and define addition and multiplication accordingly.

Now $\mathbb{R}$ can be defined as the set of *Dedekind cuts* in $\mathbb{Q}$: that is, we identify $r \in \mathbb{R}$ with $\{q \in \mathbb{Q} : q < r\} \subseteq \mathbb{Q}$ – the point being that we can define the set of such subsets of $\mathbb{Q}$ as the downwards-closed proper non-empty subsets with no greatest element (so this is a subset of $\mathcal{P}(\mathbb{Q})$ by Comprehension). We define addition and multiplication accordingly.

We can then proceed to develop real and complex analysis based on this definition of $\mathbb{R}$, defining in particular $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, identifying (a, b) with $a + ib$

and defining addition and multiplication accordingly. You may like to think how we could continue in this vein to define your favourite objects of mathematics (including those of logic!).

End of lecture 6

6 Cardinality

One key contribution of set theory is to give a rigorous mathematical development of the notion of the "size" of an infinite object, which we call its *cardinality*. We first explore what we can understand of this with the axioms we have so far. Then we add another axiom, Replacement, which will let us reach larger cardinalities. Later, we will add the Axiom of Choice, and see that this significantly clarifies the structure of cardinalities (while still leaving some natural questions undecided).

6.1 Classes

If $\phi(x)$ is a formula, it may or may not be that there is a set $\{x : \phi(x)\}$. We have $\{x : x \neq x\} = \emptyset$, but there is no set $\{x : x = x\}$.

Nonetheless, it is convenient to reuse some of the notation and terminology we use for sets to talk about $\{x : \phi(x)\}$.

Definition 6.1.

- If $\phi(x)$ is a formula with parameters, we call $\{x : \phi(x)\}$ a **class**. We denote classes with boldface characters.
- If $\mathbf{X} = \{x : \phi(x)\}$ and $\mathbf{Y} = \{x : \psi(x)\}$ are classes:
  - $a \in \mathbf{X}$ means $\phi(a)$;
  - $\mathbf{X}$ and $\mathbf{Y}$ are equal if $\forall x\ (\psi(x) \leftrightarrow \phi(x))$.
  - $\mathbf{X}$ is a **subclass** of $\mathbf{Y}$, denoted $\mathbf{X} \subseteq \mathbf{Y}$, if $\forall x\ (\phi(x) \to \psi(x))$.
- $\mathbf{V} := \{x : x = x\}$, the class of all sets.
- Sets are classes: a set a is identified with the class $\{x : x \in a\}$.
- A **proper class** is a class which is not a set.

Remark 6.2.

- By Theorem 3.2, $\mathbf{V}$ is a proper class.
- The elements of a class are always sets, not proper classes.
- Comprehension says that a subclass of a set is a set.

6.2 Cardinalities

Definition 6.3. Sets X and Y **have the same cardinality** (or are **equinumerous**), written $X \sim Y$, if there exists a bijection $X \to Y$.

We think of $\sim$ as a *class relation* on $\mathbf{V}$, defined by

$$\phi(X, Y) := \exists f [f \text{ is a bijection } X \to Y].$$

Then actually $\sim$ is a class equivalence relation:

Lemma 6.4. *For any sets X, Y, Z:*

- *$X \sim X$.*
- *If $X \sim Y$ then $Y \sim X$.*
- *If $X \sim Y$ and $Y \sim Z$ then $X \sim Z$.*

Proof. Straightforward by considering identity functions, inverses, and compositions respectively. □

Provisional Definition 6.5. The **cardinality** $|X|$ of a set X is the equivalence class of X under $\sim$:

$$|X| := \{Y : Y \sim X\}.$$

(This is a proper class, unless $X = \emptyset$.)

So $|X| = |Y| \Leftrightarrow X \sim Y$.

Later, using the axiom of choice, we will redefine $|X|$ to be a particular canonical element of this class, which we will call a *cardinal number*.

6.3 Comparing cardinalities

Definition 6.6.

- $|X| \leq |Y|$ if there exists an injection $X \to Y$.
- $|X| < |Y|$ if $|X| \leq |Y|$ and $|X| \neq |Y|$.

Lemma 6.7. *These are well-defined: if $|X| = |X'|$ and $|Y| = |Y'|$ and there exists an injection $X \to Y$, then there exists an injection $X' \to Y'$.*

Proof. Immediate by composing with bijections. □

Lemma 6.8 (Tarski's Fixed Point Theorem)**.** *Let X be a set. Then any monotone function $H : \mathcal{P}(X) \to \mathcal{P}(X)$ has a fixed point, where:*

- *$H : \mathcal{P}(X) \to \mathcal{P}(X)$ is* monotone *if $A \subseteq B$ implies $H(A) \subseteq H(B)$ (for $A, B \subseteq X$).*
- *A* fixed point *of H is a $C \subseteq X$ with $H(C) = C$.*

Proof. Let $\mathcal{D} := \{A \subseteq X : A \subseteq H(A)\}$, and let $C := \bigcup \mathcal{D}$.

- $C \subseteq H(C)$: Let $c \in C$. Then $c \in A \in \mathcal{D}$ say. Then $A \subseteq H(A)$, so $c \in H(A)$. But $A \subseteq C$, so $H(A) \subseteq H(C)$ by monotonicity, so $c \in H(C)$.
- $H(C) \subseteq C$: Since $C \subseteq H(C)$, by monotonicity $H(C) \subseteq H(H(C))$, so $H(C) \in \mathcal{D}$. Hence $H(C) \subseteq C$.

So $H(C) = C$, as required. □

Theorem 6.9 (Schröder-Bernstein Theorem[5])**.** *If* $|X| \leq |Y| \leq |X|$ *then* $|X| = |Y|$.

Proof. Say $f : X \to Y$ and $g : Y \to X$ are injections.

Define $H : \mathcal{P}(X) \to \mathcal{P}(X)$ by

$$H(A) := g[f[A]^{\mathsf{c}'}]^{\mathsf{c}} := X \setminus g[Y \setminus f[A]]$$

where we define for this proof the complements $A^{\mathsf{c}} := X \setminus A$ and $D^{\mathsf{c}'} := Y \setminus D$.

Then H is monotone, since $f[\cdot]$ and $g[\cdot]$ are inclusion-preserving while complement is inclusion-reversing; explicitly:

$$\begin{aligned} A \subseteq B \subseteq X &\Rightarrow f[A] \subseteq f[B] \\ &\Rightarrow f[A]^{\mathsf{c}'} \supseteq f[B]^{\mathsf{c}'} \\ &\Rightarrow g[f[A]^{\mathsf{c}'}] \supseteq g[f[B]^{\mathsf{c}'}] \\ &\Rightarrow H(A) = g[f[A]^{\mathsf{c}'}]^{\mathsf{c}} \subseteq g[f[B]^{\mathsf{c}'}]^{\mathsf{c}} = H(B). \end{aligned}$$

By Lemma 6.8, there is $A \subseteq X$ with $H(A) = A$. Then $A^{\mathsf{c}} = H(A)^{\mathsf{c}} = g[f[A]^{\mathsf{c}'}]$.

So we have bijections $f|_A : A \to f[A]$ and $g|_{f[A]^{\mathsf{c}'}} : f[A]^{\mathsf{c}'} \to A^{\mathsf{c}}$, and putting them together yields a bijection $f|_A \cup (g|_{f[A]^{\mathsf{c}'}})^{-1} : X \to Y$. □

Corollary 6.10. *$<$ is a strict partial (class) order on $\mathbf{V}/\sim$, i.e. for all X, Y, Z:*

(i) $|X| \not< |X|$.

(ii) If $|X| < |Y|$ *and* $|Y| < |Z|$ *then* $|X| < |Z|$.

Proof. (i) Immediate from the definition.

(ii) We have $|X| \leq |Z|$ by composing injections witnessing $|X| \leq |Y| \leq |Z|$. If $|X| = |Z|$, then $|X| \leq |Y| \leq |X|$ so $|X| = |Y|$ by Schröder-Bernstein, contrary to assumption.

□

It is perhaps natural to expect this order to be total, so that we can really think of cardinality as a linear scale of largeness. However, this does not follow from ZF, and we will see later that, modulo ZF, this order is total if and only if the Axiom of Choice holds.

End of lecture 7

[5] This is also known variously as the Cantor/Schröder-Bernstein Theorem, or the Cantor-Bernstein Theorem.

6.4 Finite sets

Definition 6.11. A set X is **finite** if $|X| = |n|$ for some $n \in \mathbb{N}$. Otherwise, X is **infinite**.

Lemma 6.12. *Let X be finite.*

(i) Any subset of X is finite.

(ii) ("Pigeonhole principle") Any injective function $f : X \to X$ is surjective.

(iii) $\mathbb{N}$ is infinite.

Proof. Exercise (sheet 2). *Hint: For (i) and (ii), prove the result by induction when $X = n \in \mathbb{N}$, and then compose with bijections when X is arbitrary. For the inductive step in (ii), if $f : n^+ \to n^+$ is an injection but $k \in n^+ \setminus \operatorname{ran}(f)$, then $f' := \sigma \circ f : n^+ \to n$ is injective where $\sigma = (n\ k) : n^+ \to n^+$ is the transposition, hence $f'|_n : n \to n$ is also injective, but $f'(n) \in n \setminus \operatorname{ran}(f'|_n)$, contradicting the induction hypothesis.* □

Lemma 6.13. *Let $n, m \in \mathbb{N}$. Then $n < m \Leftrightarrow |n| < |m|$.*

Proof. First note that if $n \leq m$, then $n \subseteq m$ by Lemma 5.10, so $|n| \leq |m|$ since the inclusion $n \to m$ is injective.

Suppose $n < m$, so in particular $n \leq m$ and so $|n| \leq |m|$. If $|n| = |m|$, then there is a bijection $f : m \to n$, but then f is also a function $f : m \to m$ which is injective but not surjective, contradicting Lemma 6.12(ii). So $|n| < |m|$.

Conversely, if $|n| < |m|$ then $|n| \not\geq |m|$ so $n \not\geq m$, so $n < m$. □

So the order on the natural numbers agrees with the order on their cardinalities, which partially justifies:

Notation 6.14. If $n \in \mathbb{N}$, we usually write the cardinality $|n|$ as n.[6] e.g. $|\emptyset| = 0$, $|\{3\}| = 1$.

6.5 Countable sets

Notation 6.15. We write $\aleph_0$ ("aleph null") for the cardinality $|\mathbb{N}|$.[7]

Definition 6.16. A set X is

- **countable** if $|X| \leq \aleph_0$.
- **countably infinite** if it is countable and infinite.
- **uncountable** if it is not countable.

Theorem 6.17. *A set X is countably infinite if and only if $|X| = \aleph_0$.*

In other words, there is no infinite cardinality below $\aleph_0$.

[6] For now this is an abuse of notation, since the set n is not actually equal to the proper class $|n|$, but this will be fixed when we eventually redefine $|\cdot|$.

[7] Later we will redefine $\aleph_0$ along with $|\mathbb{N}|$, such that $\aleph_0 = |\mathbb{N}|$ will remain true.

Proof. If $|X| = \aleph_0$ then X is infinite since $\mathbb{N}$ is (by Lemma 6.12(iii)), so X is countably infinite.

Conversely, suppose $|X| \leq |\mathbb{N}|$ and X is infinite; we show that $|X| = |\mathbb{N}|$. We may assume $X \subseteq \mathbb{N}$, since X is in bijection with its image under an injection $X \to \mathbb{N}$.

Recall that by Theorem 5.11, any non-empty subset $\emptyset \neq Y \subseteq \mathbb{N}$ has a unique least element $\min Y$. Since X is infinite and subsets of finite sets are finite, $X \setminus n \neq \emptyset$ for any $n \in \mathbb{N}$.

So define by recursion $f : \mathbb{N} \to X$ by $f(0) := \min X$ and $f(n^+) := \min(X \setminus f(n)^+)$ (the "first element of X after $f(n)$").

Then f is injective because $n > m \Rightarrow f(n) > f(m)$ by induction on n. (This is trivial for $n = 0$. Suppose for n and suppose $m < n^+$; then $f(n) \geq f(m)$ by the IH, and $f(n^+) = \min(X \setminus f(n)^+) > f(n)$, so $f(n^+) > f(m)$ as required.)

This shows that $|\mathbb{N}| \leq |X|$, and we conclude $|X| = |\mathbb{N}|$ (by Schröder-Bernstein) as required.

□

Corollary 6.18. *A non-empty set X is countable if and only if there exists a surjection $\mathbb{N} \to X$.*

Proof. $\underline{\Leftarrow}$: If $f : \mathbb{N} \to X$ is surjective, then $g(x) := \min\{n \in \mathbb{N} : f(n) = x\}$ defines an injection $g : X \to \mathbb{N}$ (which exists by Comprehension within $X \times \mathbb{N}$).

$\underline{\Rightarrow}$: By Theorem 6.17, by composing with a bijection we may suppose that either $X = \mathbb{N}$, in which case the result is immediate, or $X = n$ for some $n \in \mathbb{N}$. Then $n > 0$ since $X \neq \emptyset$, and we can define a surjection $f : \mathbb{N} \to n$ by

$$f(i) = \begin{cases} i & \text{if } i < n \\ 0 & \text{else.} \end{cases}$$

□

Remark 6.19. A natural generalisation would be: $|X| \leq |Y|$ whenever a surjection $Y \to X$ exists. We can not yet prove this for uncountable Y, but it will follow from the Axiom of Choice.

6.6 Cardinal arithmetic

Definition 6.20. Define addition, multiplication, and exponentiation of cardinalities by:

- $|X| + |Y| := |X \cup Y|$ if $X \cap Y = \emptyset$.
- $|X| \cdot |Y| := |X \times Y|$.
- $|X|^{|Y|} := |X^Y|$.

Exercise 6.21. *These do define well-defined operations on cardinalities (consider bijections). (To see that $|X| + |Y|$ is always defined, note that $|X| = |\{0\} \times X|$ and $|Y| = |\{1\} \times Y|$, and $(\{0\} \times X) \cap (\{1\} \times Y) = \emptyset$.)*

End of lecture 8

Proposition 6.22.

(a) For all cardinalities κ, λ, μ:

(i) $\kappa + \lambda = \lambda + \kappa$

(ii) $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$

(iii) $\kappa + 0 = \kappa$

(iv) $\kappa \cdot \lambda = \lambda \cdot \kappa$

(v) $\kappa \cdot (\lambda \cdot \mu) = (\kappa \cdot \lambda) \cdot \mu$

(vi) $\kappa \cdot 1 = \kappa$

(vii) $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$

(viii) $\kappa^{\lambda+\mu} = \kappa^{\lambda} \cdot \kappa^{\mu}$

(ix) $\kappa^{\lambda \cdot \mu} = (\kappa^{\lambda})^{\mu}$.

(b) These operations agree on finite cardinalities with the operations on $\mathbb{N}$ defined by recursion in Definition 5.14.

(c) If $\kappa \leq \kappa'$ and $\lambda \leq \lambda'$ then:

- $\kappa + \lambda \leq \kappa' + \lambda'$
- $\kappa \cdot \lambda \leq \kappa' \cdot \lambda'$
- $\kappa^{\lambda} \leq \kappa'^{\lambda'}$ *if* $\kappa \neq 0$.

Proof.

(a) (i) Say $\kappa = |X|$ and $\lambda = |Y|$ and $X \cap Y = \emptyset$. Then $X \cup Y = Y \cup X$, so $|X| + |Y| = |Y| + |X|$.

(ii),(iii) Similar equalities of sets show these.

(iv) $\langle x, y \rangle \mapsto \langle y, x \rangle$ defines a bijection $X \times Y \to Y \times X$, so $|X| \cdot |Y| = |X \times Y| = |Y \times X| = |Y| \cdot |X|$.

(v)-(ix) Similar bijections show these (see Sheet 3).

(b) Exercise (Sheet 3).

(c) Exercise (Sheet 3).

□

Proposition 6.23. $|\mathcal{P}(X)| = 2^{|X|}$ *for any set X.*

Proof. The function $F : \mathcal{P}(X) \to 2^X$ defined by

$$F(Y)(x) = \begin{cases} 0 & \text{if } x \notin Y \\ 1 & \text{if } x \in Y \end{cases}$$

(i.e. $F(Y)$ is the *indicator function* of Y in X) is a bijection. □

Theorem 6.24 (Cantor)**.** *Let X be a set. Then there is no surjection $X \to \mathcal{P}(X)$.*

Proof. Let $f : X \to \mathcal{P}(X)$. Let

$$D := \{x \in X : x \notin f(x)\} \subseteq X.$$

Suppose $a \in X$ and $D = f(a)$. Then $a \in D$ iff $a \notin f(a) = D$, contradiction. So $D \in \mathcal{P}(X) \setminus ran(f)$, so f is not surjective. □

Corollary 6.25. *$\kappa < 2^\kappa$ for any cardinality κ.*

Proof. Say $\kappa = |X|$, so $2^\kappa = |\mathcal{P}(X)|$ (by Proposition 6.23). Then $x \mapsto \{x\}$ is an injection $X \to \mathcal{P}(X)$, so $\kappa \leq 2^\kappa$. If $\kappa = 2^\kappa$, then there is a bijection $X \to \mathcal{P}(X)$, contradicting Theorem 6.24. □

Lemma 6.26. *$\aleph_0 \cdot \aleph_0 = \aleph_0$.*

Proof.

- $\underline{\aleph_0 \leq \aleph_0 \cdot \aleph_0}$: $n \mapsto \langle n, 0\rangle$ is an injection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$.
- $\underline{\aleph_0 \cdot \aleph_0 \leq \aleph_0}$: $\langle n, m\rangle \mapsto 2^n \cdot 3^m$ is an injection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$. This follows from the Fundamental Theorem of Arithmetic (unique prime factorisation), whose proof we omit.

Alternatively, $\langle n, m\rangle \mapsto \frac{1}{2}(n+m)(n+m+1)+m = \binom{n+m+1}{2} + m$ is a bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ (again, we omit details). □

Theorem 6.27.

(i) $|\mathbb{Q}| = \aleph_0$.

(ii) $|\mathbb{R}| = 2^{\aleph_0}$.

Proof.

(i) $|\mathbb{Q}| \geq \aleph_0$ since $n \mapsto \frac{n}{1}$ is an injection, and $\langle\langle n, m\rangle, k\rangle \mapsto \frac{n-m}{k+1}$ is a surjection $(\mathbb{N} \times \mathbb{N}) \times \mathbb{N} \to \mathbb{Q}$, and $|(\mathbb{N} \times \mathbb{N}) \times \mathbb{N}| = \aleph_0$ by Lemma 6.26 (twice), so $|\mathbb{Q}| \leq \aleph_0$ by Corollary 6.18.

(ii) The map which associates a real with its Dedekind cut, $x \mapsto \{q \in \mathbb{Q} : q < x\}$, is an injection $\mathbb{R} \to \mathcal{P}(\mathbb{Q})$ since $\mathbb{Q}$ is dense in $\mathbb{R}$. So $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{Q})| = 2^{|\mathbb{Q}|} = 2^{\aleph_0}$.

For the converse, we can use ternary expansions. Define $\Phi : 2^{\mathbb{N}} \to \mathbb{R}$ by

$$\Phi(f) := \sum_{n=0}^{\infty} \frac{f(n)}{3^n}.$$

Then Φ is injective (using $\sum_{n=1}^{\infty} 3^{-n} = \frac{1}{2} < 1$; binary expansions would not work), so $2^{\aleph_0} \leq |\mathbb{R}|$.

□

7 Replacement and Foundation

The axiom system we have established so far, ZF1-7, is more or less the system originally proposed by Zermelo. To obtain ZF, we add two further axioms which were proposed a little later. Replacement radically increases the power of the system, while Foundation trims the universe by denying the existence of "pathological" sets like $u = \{u\}$.

7.1 Replacement

Definition 7.1. If $\mathbf{X}$ and $\mathbf{Y}$ are classes, a formula with parameters $\phi(x, y)$ defines a **class function** $\mathbf{F} : \mathbf{X} \to \mathbf{Y}$ if:

- $\phi(x, y)$ implies $x \in \mathbf{X}$ and $y \in \mathbf{Y}$, and
- for all $x \in \mathbf{X}$ there is a unique y such that $\phi(x, y)$ holds.

We then write $\mathbf{F}(x) = y$ to mean $\phi(x, y)$.

End of lecture 9

Example 7.2. $\mathcal{P} : \mathbf{V} \to \mathbf{V}$ is the class function defined by

$$\psi(x, y) := \forall w\ (w \in y \leftrightarrow w \subseteq x).$$

ZF8 (Replacement): If a is a set and $\mathbf{F} : a \to \mathbf{V}$ is a class function, then its range $\mathbf{F}[a] := \{\mathbf{F}(x) : x \in a\}$ is a set.

Remark 7.3. Then $\mathbf{F} : a \to \mathbf{F}[a]$ is actually a function, by Comprehension. So we could equivalently state Replacement as: a class function on a set is a function.

Remark 7.4. As with Comprehension, we can formalise Replacement by an axiom scheme consisting of, for each $\mathcal{L}$-formula $\phi(x, y, z_1, \ldots, z_n)$, the sentence

$$\begin{aligned}\forall z_1 \ldots \forall z_n\ \forall w\ (\forall x \in w\ \exists y\ (\phi(x, y, z_1, \ldots, z_n) \wedge \forall y'\ (\phi(x, y', z_1, \ldots, z_n) \to y' = y))\\ \to \exists v\ \forall u\ (u \in v \leftrightarrow \exists x \in w\ \phi(x, u, z_1, \ldots, z_n))).\end{aligned}$$

One immediate application of Replacement is to strengthen our recursion principle on $\mathbb{N}$.

Theorem 7.5 (Recursion on $\mathbb{N}$, class form). *If x_0 is a set and $\mathbf{G} : \mathbf{V} \to \mathbf{V}$ is a class function, then there exists a unique function f with* $\mathrm{dom}(f) = \mathbb{N}$ *such that*

- *$f(0) = x_0$;*
- *$f(n^+) = G(f(n))$ for all $n \in \mathbb{N}$.*

Proof. Exactly as in the proof of Theorem 5.12, for each n there is a unique n-approximation. Then $\mathbf{F}(n) :=$ [the unique n-approximation] is a class function $\mathbf{F} : \mathbb{N} \to \mathbf{V}$, so by Replacement, $H := \mathbf{F}[\mathbb{N}] = \{h : \exists n \in \mathbb{N}\ [h \text{ is an } n\text{-approximation}]\}$ is a set. Set $f := \bigcup H$, and conclude exactly as in Theorem 5.12. □

Example 7.6. There is a cardinality greater than any of $\aleph_0, 2^{\aleph_0}, 2^{2^{\aleph_0}}, \ldots$, in the following sense.

Applying recursion with $x_0 := \mathbb{N}$ and $G := \mathcal{P}$, we obtain a function f with $\mathrm{dom}(f) = \mathbb{N}$ and $f(0) = \mathbb{N}$, $f(1) = \mathcal{P}(\mathbb{N})$, $f(2) = \mathcal{P}(\mathcal{P}(\mathbb{N}))$, ...

Then $\mathrm{ran}(f) = f[\mathbb{N}]$ is a set which we could write as $\{\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}(\mathcal{P}(\mathbb{N})), \ldots\}$.

Let $X := \bigcup f[\mathbb{N}]$. Then for any $n \in \mathbb{N}$, we have $f(n) \subseteq X$ and hence $|X| \geq f(n)$.

One can show that ZF1-7 do not suffice to prove the existence of such a cardinality.

7.2 Foundation

ZF9 (Foundation): Every non-empty set x has an $\in$-minimal element, i.e. an element $y \in x$ such that no element of x is an element of y:

$$\forall x\ (x \neq \emptyset \to \exists y \in x\ y \cap x = \emptyset).$$

This axiom forbids certain "pathological" behaviour of sets:

Theorem 7.7. *(i) There is no x with $x \in x$.*

(ii) There are no x and y with $x \in y \in x$.

(iii) More generally, there is no infinite descending $\in$-chain, i.e. no function f with $\mathrm{dom}(f) = \mathbb{N}$ *and $f(n^+) \in f(n)$ for all $n \in \mathbb{N}$ (so $f(0) \ni f(1) \ni f(2) \ni \ldots$.*

Proof. (i) If $x \in x$, then $\{x\}$ violates Foundation: the only element of $\{x\}$ is x, but $x \cap \{x\} = \{x\} \neq \emptyset$.

(ii) If $x \in y \in x$, then $\{x, y\}$ violates Foundation since $y \in x \cap \{x, y\}$ and $x \in y \cap \{x, y\}$.

(iii) Exercise (Sheet 3).

□

One can show that if ZF1-8 are consistent, then so are ZF1-9: adding Foundation can not introduce a contradiction. In particular, any set we prove to exist using ZF1-8 (such as $\mathbb{N}$ and $\mathbb{R}$) does not violate Foundation. So adding Foundation is "harmless", and substantially simplifies the set theoretic universe. However, Foundation will not actually be used in the remainder of these notes.

ZF1-9 form the axiom system ZF. Later we will add the Axiom of Choice (AC) to form ZFC, but we delay this until we need it.

End of lecture 10

8 Well-ordered sets and ordinals

One way the natural numbers arise is as a measure of size, and we have now generalised this to infinite cardinalities. Another way natural numbers arise is in enumerating elements of an ordered set in which "the nth element" makes sense.

We now generalise this *ordinal* sense of a natural number to infinite (*transfinite*) ordinal numbers. First we define the orders which can be enumerated in this sense – those in which the "next" element always exists, even if not every element is of this form – then we define and study a notion of *ordinal number* with which we can enumerate such orders, so that each element is "the αth element" for some ordinal number α.

We won't use Foundation in this section; it would slightly simplify some of the proofs, but it isn't necessary.

8.1 Well-ordered sets

Definition 8.1. A **well-ordered** set (or **well-order**) is a totally ordered set $(X, <)$ which is **well-founded**, meaning:

- Every non-empty subset $\emptyset \neq Y \subseteq X$ has a least element.

We denote this least element $\min Y$ (or $\min_X Y$).

Example 8.2.

- $\mathbb{N}$ is well-ordered by $< = \in$, by Theorem 5.11.
- $\mathbb{Z}$ is not well-ordered by its usual order $<$, since $\mathbb{Z}$ has no least element. Same for $\mathbb{R}$.
- $[0, 1] \subseteq \mathbb{R}$ is not well-ordered by $<$, since $(0, 1]$ lacks a least element.
- $\{-\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\} \cup \mathbb{N} \subseteq \mathbb{R}$ is well-ordered by $<$.
- Any subset Y of a well-ordered set $(X, <)$ is well-ordered by the restriction of $<$, and we write this well-order as $(Y, <)$.

Definition 8.3. Let $(X, <)$ be a well-ordered set.

- An **initial segment** of X is a subset $S \subseteq X$ which is downwards closed in X, i.e. $\forall y \in S\ \forall x \in X\ (x < y \to x \in S)$. It is a **proper** initial segment if $S \neq X$.
- We consider initial segments as well-ordered sets, with the restriction of $<$.
- For $a \in X$, define $X_{<a} := \{x \in X : x < a\}$.

Remark 8.4. The proper initial segments of X are precisely the sets $X_{<a}$ for $a \in X$: indeed, if $S \subsetneq X$ is a proper initial segment then $S = X_{<\min(X \setminus S)}$.

Definition 8.5. An **embedding** of a totally ordered set $(X, <)$ in a totally ordered set $(Y, <')$ is a function $\theta : X \to Y$ which is strictly monotone, i.e. $x < x' \Rightarrow \theta(x) <' \theta(x')$ for all $x, x' \in X$.

An **isomorphism** is a surjective embedding, and we write $(X, <) \cong (Y, <')$ and say the ordered sets are **isomorphic** if an isomorphism exists.

Well-orders are highly rigid:

Lemma 8.6. *If $(X,<)$ is a well-order and $\theta : (X,<) \to (X,<)$ is an embedding, then $\theta(x) \geq x$ for all $x \in X$.*

Proof. Suppose not. Then $a := \min\{x \in X : \theta(x) < x\}$ exists. But then $\theta(a) < a$, so $\theta(\theta(a)) < \theta(a)$ since θ is an embedding, contradicting minimality of a. □

Lemma 8.7. *A well-order is not isomorphic to any of its proper initial segments.*

Proof. If $\sigma : X \to X_{<x}$ is an isomorphism, then $\sigma(x) < x$, contradicting Lemma 8.6. □

Lemma 8.8. *Let $(X,<)$ be a well-order.*

(i) The only isomorphism $X \to X$ is the identity.

(ii) If $(X,<) \cong (Y,<')$, then there is a unique isomorphism $X \to Y$.

Proof. (i) If $\sigma : X \to X$ is an isomorphism, then so is σ^{-1}, so by Lemma 8.6, for all $x \in X$ we have $\sigma(x) \leq x$ and $\sigma^{-1}(x) \leq x$, hence $x \leq \sigma(x) \leq x$, hence $\sigma(x) = x$.

(ii) If $\sigma, \tau : X \to Y$ are isomorphisms then $\tau^{-1}(\sigma(x)) = x$ for all $x \in X$ by (i), so $\sigma = \tau$.

□

Any two well-orders are comparable:

Theorem 8.9. *Let $(X,<)$ and $(Y,<')$ be well-orders. Then either $(X,<)$ is isomorphic to an initial segment of $(Y,<')$, or $(Y,<')$ is isomorphic to an initial segment of $(X,<)$.*

Proof. Define

$$\sigma := \{\langle x,y\rangle \in X \times Y : (X_{<x},<) \cong (Y_{<'y},<')\}.$$

Then σ is a function, since if $\langle x,y\rangle, \langle x,y'\rangle \in \sigma$, then $(Y_{<y},<) \cong (Y_{<y'},<)$, so $y = y'$ by Lemma 8.7. Symmetrically, σ is injective.

Let $\langle x,y\rangle \in \sigma$, so say $\tau : X_{<x} \to Y_{<'y}$ is an isomorphism. Then if $x' < x$, then $\tau|_{X_{<x'}} : X_{<x'} \to Y_{<'\tau(x')}$ is an isomorphism, so $\langle x', \tau(x')\rangle \in \sigma$.

Hence $X' := \mathrm{dom}(\sigma) \subseteq X$ is an initial segment of X, and symmetrically $Y' := \mathrm{ran}(\sigma) \subseteq Y$ is an initial segment of Y, and $\sigma : X' \to Y'$ is an isomorphism, since $\sigma(x') = \tau(x') <' y = \sigma(x)$.

So $X' \cong Y'$. If X' and Y' are proper initial segments, say $X' = X_{<x}$ and $Y' = Y_{<'y}$, then $\langle x,y\rangle \in \sigma$, contradicting $X' = \mathrm{dom}(\sigma)$. So either $X' = X$ or $Y' = Y$, as required. □

End of lecture 11

8.2 Ordinals

Definition 8.10. A set a is **transitive** if every element of a is a subset of a, i.e. $x \in y \in a \Rightarrow x \in a$.

Definition 8.11. An **ordinal** is a transitive set which is well-ordered by $\in$.

That is, an ordinal is a transitive set α such that $(\alpha, <)$ is a well-ordered[8] set, where $< := \{\langle \beta, \gamma \rangle \in \alpha \times \alpha : \beta \in \gamma\}$.

We use $<$ and $\in$ interchangeably to denote the order on an ordinal.

We denote the class of ordinals by **ON**.

By Theorem 5.11 and the Remark following it, $\mathbb{N}$ is an ordinal. We use ω to denote $\mathbb{N}$ when we consider it as an ordinal.

Lemma 8.12. *Any element of an ordinal is an ordinal.*

Proof. Let $\beta \in \alpha \in \mathbf{ON}$. Then $\beta \subseteq \alpha$ by transitivity of α, so the restriction $(\beta, \in)$ is a well-order. But β is transitive, since if $x \in y \in \beta$ then $y \in \alpha$ and $x \in \alpha$ by transitivity of α, so then $x \in \beta$ by the transitivity property of the order $\in$ on α. So β is an ordinal. □

Lemma 8.13. *Let $\beta \in \mathbf{ON}$.*

(i) If $\alpha \in \beta$ then $\alpha = \beta_{<\alpha}$. In particular, the elements of β are precisely the proper initial segments of β.

(ii) If α is transitive (in particular, if $\alpha \in \mathbf{ON}$), then $\alpha \subsetneq \beta$ iff $\alpha \in \beta$.

Proof. (i) $\beta_{<\alpha} = \{\gamma \in \beta : \gamma \in \alpha\} = \alpha$, since $\gamma \in \alpha \Rightarrow \gamma \in \beta$ by transitivity of β.

(ii) If $\alpha \subseteq \beta$ then α is an initial segment of β by transitivity of α.

So $\alpha \subsetneq \beta$ iff α is a proper initial segment of β, so we conclude by (i).

□

Theorem 8.14. *The class **ON** is well-ordered by $\in$, i.e. for all $\alpha, \beta, \gamma \in \mathbf{ON}$:*

(i) $\alpha \notin \alpha$ (Irreflexivity)

(ii) $\alpha \in \beta \in \gamma \Rightarrow \alpha \in \gamma$ (Transitivity)

(iii) $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$ (Totality)

(iv) Any non-empty class of ordinals has an $\in$-least element. (Well-foundedness)

Proof. (i) By Lemma 8.13(ii).

(ii) By transitivity of γ.

(iii) By Lemma 8.13(ii), it suffices to show that $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$. Suppose not. Then $\gamma := \alpha \cap \beta$ is a proper subset of α and of β. But γ is transitive since α and β are, so $\gamma \in \alpha \cap \beta$ by Lemma 8.13(ii), so $\gamma \in \gamma$, which contradicts (i) since γ is an ordinal by Lemma 8.12.

[8] Using Foundation, we could equivalently say “totally ordered”

(iv) This is immediate from Foundation, but we can also argue directly as follows. Given a non-empty class $\mathbf{\Gamma}$ of ordinals and $\gamma \in \mathbf{\Gamma}$, if $\mathbf{\Gamma} \cap \gamma = \emptyset$ then $\min \mathbf{\Gamma} = \gamma$, and otherwise $\min \mathbf{\Gamma} = \min(\mathbf{\Gamma} \cap \gamma)$, which exists since γ is an ordinal.

□

Corollary 8.15. *Any transitive set of ordinals is an ordinal.*

Proof. Theorem 8.14 shows that $\in$ defines a well-order on any set of ordinals. □

Theorem 8.16. **ON** *is a proper class.*[9]

Proof. Suppose **ON** is a set. Then **ON** is transitive by Lemma 8.12, so **ON** is an ordinal by Corollary 8.15. But then $\mathbf{ON} \in \mathbf{ON}$, contradicting Theorem 8.14(i).

□

Lemma 8.17. *Isomorphic ordinals are equal.*

Proof. By Theorem 8.14(iii) and Lemma 8.13(i), if $\alpha, \beta \in \mathbf{ON}$ are not equal then one is a proper initial segment of the other, so by Lemma 8.7 they are not isomorphic. □

Theorem 8.18 (Hartogs' Theorem)**.** *If X is a set, then there exists $\alpha \in \mathbf{ON}$ with $|\alpha| \not\leq |X|$.*

Proof. Suppose for a contradiction that $|\alpha| \leq |X|$ for all $\alpha \in \mathbf{ON}$. Then for every $\alpha \in \mathbf{ON}$, there is an injection $f : \alpha \to X$, and then $f(x) < f(y) \Leftrightarrow x \in y$ defines a well-order on $f[X] \subseteq X$ which is isomorphic to α.

Considering a well-order $(Y, <)$ as an ordered pair $\langle Y, < \rangle$, and let $W \subseteq \mathcal{P}(X) \times \mathcal{P}(X \times X)$ be the set of all well-orders $(Y, <)$ with $Y \subseteq X$ such that $(Y, <)$ is isomorphic to some ordinal, and let $\mathbf{F} : W \to \mathbf{ON}$ be the class function such that $\mathbf{F}((Y, <))$ is the ordinal isomorphic to $(Y, <)$, which is unique by Lemma 8.17. Then $\mathbf{F}[W] = \mathbf{ON}$ by the previous paragraph, so **ON** is a set by Replacement, contradicting Theorem 8.16. □

Theorem 8.19. *Every well-ordered set $(X, <)$ is isomorphic to a unique ordinal by a unique isomorphism.*

Proof. Uniqueness of the ordinal is by Lemma 8.17, and uniqueness of the isomorphism is by Lemma 8.8.

For existence, by Theorem 8.18 say $\alpha \in \mathbf{ON}$ with $|\alpha| \not\leq |X|$. Then α is not isomorphic to an initial segment of X, so by Theorem 8.9, X is isomorphic to an initial segment of α, which is an ordinal by Lemma 8.13. □

Lemma 8.20.

(a) *(i)* $0 = \emptyset$ *is an ordinal.*

(ii) If α is an ordinal, then so is its ***successor*** $\alpha^+ = \alpha \cup \{\alpha\}$.

(iii) If Γ is a set of ordinals, then $\bigcup \Gamma$ is an ordinal.

(b) Every $\beta \in \mathbf{ON}$ is of precisely one of the following three types:

[9] This is known as the *Burali-Forti paradox.*

(i) Zero ordinal: $\beta = 0$.

(ii) Successor ordinal: $\beta = \alpha^+$ *for some* $\alpha \in \mathbf{ON}$.

(iii) Limit ordinal: $\beta = \bigcup \beta$ *and* $\beta \neq 0$.

End of lecture 12

Proof. (a) By Lemma 8.12, 0, α^+, and $\bigcup \Gamma$ are sets of ordinals, so by Corollary 8.15 it remains only to show that they are transitive. We leave this verification as an exercise (Sheet 1).

(b) 0 is not a successor ordinal nor a limit ordinal. A successor ordinal α^+ is not a limit ordinal, since $\alpha \notin \bigcup \alpha^+$.

Suppose $\beta \in \mathbf{ON}$ is not a successor ordinal, and let $\alpha \in \beta$. Then $\beta \neq \alpha^+$, and so $\alpha^+ \in \beta$ by Theorem 8.14(iii), so $\alpha \in \bigcup \beta$. Conversely, $\bigcup \beta \subseteq \beta$ by transitivity of β. So $\beta = \bigcup \beta$ is either 0 or a limit ordinal.

□

Example 8.21.

- $\omega = \mathbb{N}$ is the first limit ordinal, since $\omega = \bigcup \omega$ and every $n \in \omega$ is either zero or a successor.
- We have ordinals $\omega^+, \omega^{++}, \ldots$ and their limit $\bigcup\{\omega, \omega^+, \omega^{++}, \ldots\}$ (defining this set by recursion on ω). We will soon define ordinal addition and write these as $\omega + 1, \omega + 2, \ldots$ and $\omega + \omega$.

8.3 Transfinite recursion

Theorem 8.22 (Transfinite Induction). *Let $\phi(x)$ be a formula with parameters. Suppose that $\phi(\beta)$ holds for every $\beta \in \mathbf{ON}$ for which $\phi(\gamma)$ holds for all $\gamma \in \beta$. Then $\phi(\alpha)$ holds for all $\alpha \in \mathbf{ON}$.*

Proof. Otherwise, let $\beta := \min\{\beta \in \mathbf{ON} : \neg\phi(\beta)\}$ (using Theorem 8.14(iv)). Then $\phi(\gamma)$ holds for all $\gamma \in \beta$ by the minimality of β, so $\phi(\beta)$ holds, contradiction. □

Theorem 8.23 (Transfinite Recursion). *Let $\mathbf{G} : \mathbf{V} \to \mathbf{V}$ be a class function. Then there exists a unique class function $\mathbf{F} : \mathbf{ON} \to \mathbf{V}$ such that for all $\alpha \in \mathbf{ON}$*

$$\mathbf{F}(\alpha) = \mathbf{G}(\mathbf{F}|_\alpha).$$

(This make sense because, by Replacement, $\mathbf{F}|_\alpha : \alpha \to \mathbf{F}[\alpha]$ is a set for any $\alpha \in \mathbf{ON}$.)

Sketch proof (not examinable). Analogous to the proof of Theorem 5.12.

For $\alpha \in \mathbf{ON}$, define an *α-approximation* to be a function $f_\alpha : \alpha^+ \to \mathbf{V}$ such that $f_\alpha(\beta) = \mathbf{G}(f_\alpha|_\beta)$ for all $\beta \in \alpha^+$.

We show by transfinite induction that a unique α-approximation f_α exists for each $\alpha \in \mathbf{ON}$. Indeed, let $\alpha \in \mathbf{ON}$ and suppose f_β is the unique β-approximation for each $\beta \in \alpha$. Note then that $f_\beta|_{\gamma^+} = f_\gamma$ whenever $\gamma \in \beta \in \alpha$, since it is a γ-approximation. So $g_\alpha := \bigcup\{f_\beta : \beta \in \alpha\}$ is a function (using Replacement), with domain $\bigcup\{\beta^+ : \beta \in \alpha\} = \alpha$. If f_α is an α-approximation and $\beta \in \alpha$, then again $f_\alpha|_{\beta^+} = f_\beta$, so $f_\alpha := g_\alpha \cup \{(\alpha, \mathbf{G}(g_\alpha))\}$ is the unique α-approximation.

Now $\mathbf{F}(\alpha) := f_\alpha(\alpha)$ is the unique class function in the statement – unique because again the restriction of $\mathbf{F}$ to any α^+ is an α-approximation. □

We typically apply this in the following form.

Corollary 8.24. *Let x_0 be a set and let $\mathbf{S} : \mathbf{V} \to \mathbf{V}$ be a class function. Then there is a unique class function $\mathbf{F} : \mathbf{ON} \to \mathbf{V}$ such that:*

- $\mathbf{F}(0) = x_0$.
- $\mathbf{F}(\alpha^+) = \mathbf{S}(\mathbf{F}(\alpha))$ *for all* $\alpha \in \mathbf{ON}$.
- *If* $\eta \in \mathbf{ON}$ *is a limit ordinal, then* $\mathbf{F}(\eta) = \bigcup\{\mathbf{F}(\beta) : \beta \in \eta\} = \bigcup \mathbf{F}[\eta]$.

Proof. Define $\mathbf{G}(f)$ as follows. If f is a function with domain an ordinal β: if $\beta = 0$ then set $\mathbf{G}(f) = x_0$, else if $\beta = \alpha^+$ is a successor (i.e. has a largest element α) then set $\mathbf{G}(f) := \mathbf{S}(f(\alpha))$, else set $\mathbf{G}(f) := \bigcup \operatorname{ran}(f)$. Otherwise, set $\mathbf{G}(f) := \emptyset$ (say).

Now apply Theorem 8.23 to obtain $\mathbf{F}$ with $\mathbf{F}(\beta) = \mathbf{G}(\mathbf{F}|_\beta)$, and note that it is as required. □

Remark 8.25. In fact the proof of Theorem 8.23 yields a uniform version (analogous to Corollary 5.13): if $\mathbf{G} = \mathbf{G}_b$ has a parameter b, then $\mathbf{F} = \mathbf{F}_b$ also has this parameter, and $\mathbf{F}_b(\alpha) = \mathbf{G}_b(\mathbf{F}_b|_\alpha)$ holds for all b. Hence also in Corollary 8.24, $\mathbf{S}$ and x_0 may depend on a parameter.

Example 8.26 (Cumulative Hierarchy (not on syllabus)). Apply Corollary 8.24 with $\mathbf{S} = \mathcal{P}$ to obtain a class function $\mathbf{F} : \mathbf{ON} \to \mathbf{V}$ such that, writing V_α for $\mathbf{F}(\alpha)$, we have

- $V_0 = \emptyset$
- $V_{\alpha^+} = \mathcal{P}(V_\alpha)$
- $V_\eta = \bigcup\{V_\beta : \beta \in \eta\}$ if η is a limit ordinal.

This is called the *von Neumann cumulative hierarchy.* One proves by transfinite induction that $V_\alpha \subseteq V_\beta$ for $\alpha \subseteq \beta$. The *rank* of a set x is then defined as the least $\alpha \in \mathbf{ON}$ such that $x \subseteq V_\alpha$ (i.e. $x \in V_{\alpha^+}$), if such exists. The axiom of Foundation is equivalent, modulo the other axioms of ZF, to the statement that every set is an element of some V_α, i.e. that every set has a rank.

8.4 Ordinal arithmetic

We now extend our recursive definitions of the arithmetic operations from ω to $\mathbf{ON}$:

Definition 8.27. Define by Corollary 8.24 (and Remark 8.25) the unique class functions $+, \cdot, ^\frown : \mathbf{ON} \times \mathbf{ON} \to \mathbf{ON}$ such that for all $\alpha, \beta \in \mathbf{ON}$:

-
  - $\alpha + 0 = \alpha$
  - $\alpha + \beta^+ = (\alpha + \beta)^+$

  - $\alpha + \eta = \bigcup\{\alpha + \beta : \beta \in \eta\}$ for η a limit ordinal.

- - $\alpha \cdot 0 = 0$
  - $\alpha \cdot \beta^+ = \alpha \cdot \beta + \alpha$
  - $\alpha \cdot \eta = \bigcup\{\alpha \cdot \beta : \beta \in \eta\}$ for η a limit ordinal.

- - $\alpha^0 = 1$
  - $\alpha^{\beta^+} = (\alpha^\beta) \cdot \alpha$
  - $\alpha^\eta = \bigcup\{\alpha^\beta : \beta \in \eta\}$ for η a limit ordinal.

Example 8.28.

- $1 + \omega = \bigcup_{n\in\omega} 1 + n = \omega \neq \omega^+ = \omega + 1$.
- $\alpha \cdot 1 = \alpha \cdot 0 + \alpha = 0 + \alpha = \alpha$, where the last equality holds by transfinite induction on α.
- $2 \cdot \omega = \bigcup_{n\in\omega} 2 \cdot n = \omega \neq \omega + \omega = \omega \cdot 2$.
- $2^\omega = \bigcup_{n\in\omega} 2^n = \omega \neq \omega \cdot \omega = \omega^2$.
- $2^\omega = \omega$ is countable, so it is not in bijection with the set of functions $\omega \to 2$ – beware this conflict in notation!

Fact 8.29. *The set of countable ordinals is closed under these arithmetic operations. Uncountable ordinals do nonetheless exist, by Hartogs' theorem.*

Definition 8.30. Let $(A, <_A)$ and $(B, <_B)$ be linear orders.

- The **reverse lexicographic product order** (or just **product order**) is the linear order $(A, <_A) \times (B, <_B) := (A \times B, <_\times)$ where
$$(a, b) <_\times (a', b') \Leftrightarrow (b <_B b' \vee (b = b' \wedge a <_A a')).$$
- The **sum order** is the linear order $(A, <_A) + (B, <_B) := ((A \times \{0\}) \cup (B \times \{1\}), <_+)$ where for all $a, a' \in A$ and $b, b' \in B$:
$$\begin{aligned}(a, 0) &<_+ (a', 0) \Leftrightarrow a <_A a' \\ (b, 1) &<_+ (b', 1) \Leftrightarrow b <_B b' \\ (a, 0) &<_+ (b, 1).\end{aligned}$$

Theorem 8.31. *Let $\alpha, \beta \in$* **ON**.

(a) $(\alpha + \beta, \in) \cong (\alpha, \in) + (\beta, \in)$.

(b) $(\alpha \cdot \beta, \in) \cong (\alpha, \in) \times (\beta, \in)$.

End of lecture 13

Proof. (a) By transfinite induction on β for a fixed α:

- $\beta = 0$: Immediate.
- $\beta = \gamma^+$: $\alpha + \beta = (\alpha + \gamma)^+$, which inductively is isomorphic to the extension of $(\alpha, \in)+(\gamma, \in)$ by a new greatest element, which is isomorphic to $(\alpha, \in) + (\gamma^+, \in)$.

- β limit: $\alpha+\beta = \bigcup_{\gamma\in\beta}(\alpha+\gamma)$, and inductively $(\alpha+\gamma,\in) \cong (\alpha,\in)+(\gamma,\in)$ for each $\gamma \in \beta$.

  Let $\sigma_\gamma : (\alpha,\in)+(\gamma,\in) \to (\alpha+\gamma,\in)$ be the unique (by Lemma 8.8) isomorphisms. Then they form a chain: if $\delta \in \gamma$ then σ_γ restricts to an isomorphism of $(\alpha,\in)+(\delta,\in)$ with an initial segment of $\alpha+\gamma$, which is also an ordinal and so must be $\alpha+\delta$ (by the uniqueness in Theorem 8.19); hence σ_γ extends σ_δ.

  So their union $\sigma := \bigcup_{\gamma\in\beta}\sigma_\gamma$ (which is a set by Replacement) is an isomorphism of $\bigcup_{\gamma\in\beta}((\alpha,\in)+(\gamma,\in)) = (\alpha,\in)+(\beta,\in)$ with $\bigcup_{\gamma\in\beta}(\alpha+\gamma) = \alpha+\beta$.

(b) By the same argument, except that for the successor stage we argue as follows:

$$(\alpha\cdot\gamma^+,\in) = (\alpha\cdot\gamma+\alpha,\in) \cong (\alpha,\in)\times(\gamma,\in)+(\alpha,\in) \cong (\alpha,\in)\times(\beta,\in),$$

where the penultimate isomorphism uses the IH and (a), and the final isomorphism is by the definitions of the sum and product orders.

□

Lemma 8.32. *If $B \subseteq \alpha \in \mathbf{ON}$ is a subset of an ordinal α, then the induced order $(B,\in)$ is isomorphic to some $\beta \leq \alpha$.*

Proof. Let β be the ordinal isomorphic to $(B,\in)$. If $\beta \not\leq \alpha$, then $\alpha < \beta$, so α is isomorphic to a proper initial segment of B, say $B_{<}b$. But then α embeds into the proper initial segment $\alpha_{<}b$ of α, contradicting Lemma 8.6. □

Theorem 8.33. *For all $\alpha,\beta,\gamma \in \mathbf{ON}$:*

(a) *(i)* $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$.

(ii) $\beta < \gamma \Rightarrow \alpha+\beta < \alpha+\gamma$.

(iii) $\alpha \leq \gamma \Rightarrow \alpha+\beta \leq \gamma+\beta$.

(iv) $\alpha+\beta = \alpha+\gamma \Rightarrow \beta = \gamma$.

(v) $\alpha \leq \beta \Rightarrow \exists\delta \leq \beta\ \alpha+\delta = \beta$.

(b) *(i)* $(\alpha\cdot\beta)\cdot\gamma = \alpha\cdot(\beta\cdot\gamma)$.

(ii) $\alpha\cdot(\beta+\gamma) = \alpha\cdot\beta+\alpha\cdot\gamma$.

(iii) *For* $\alpha \neq 0$, $\beta < \gamma \Rightarrow \alpha\cdot\beta < \alpha\cdot\gamma$.

(iv) $\alpha \leq \gamma \Rightarrow \alpha\cdot\beta \leq \gamma\cdot\beta$.

Proof. (a) (i) By the corresponding associativity of the sum of orders.

(ii) If $\beta < \gamma$ then $(\beta,\in)$ is a proper initial segment of $(\gamma,\in)$, so $(\alpha,\in)+(\beta,\in)$ is a proper initial segment of $(\alpha,\in)+(\gamma,\in)$, so $\alpha+\beta < \alpha+\gamma$.

(iii) $(\alpha+\beta,\in)$ is isomorphic to a suborder of $(\gamma+\beta,\in)$ by considering ordered sums, so $\alpha+\beta \leq \gamma+\beta$ by Lemma 8.32.

(iv) By (ii) and totality.

(v) By Lemma 8.32, $(\beta\setminus\alpha, \in)$ is isomorphic to $(\delta, \in)$ for some $\delta \leq \beta$. Then

$$(\beta, \in) \cong (\alpha, \in) + (\beta \setminus \alpha, \in) \cong (\alpha + \delta, \in),$$

so $\beta = \alpha + \delta$.

(b) Exercise. Consider product orders, and apply Lemma 8.32 for (iv). (iii) can also be proven by induction.

□

9 The Axiom of Choice

Definition 9.1. The **Axiom of Choice**, AC, is the following statement:

If X is a set of disjoint non-empty sets, then there exists a set C such that $|C \cap a| = 1$ for all $a \in X$.

So Y "chooses" an element of each $a \in X$.

We first give some immediate reformulations of AC.

Lemma 9.2. *The following are equivalent:*

(a) AC

(b) Every set X has a ***choice function****, a function $h : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ such that $h(A) \in A$ for all $\emptyset \neq A \subseteq X$.*

——————————— *End of lecture 14*

Proof.

- (a) ⇒ (b): The set

$$Y := \{\{A\} \times A : \emptyset \neq A \subseteq X\}$$

is a set of disjoint non-empty sets, so say C is such that $|C \cap (\{A\} \times A)| = 1$ for all $\emptyset \neq A \subseteq X$. So for each $\emptyset \neq A \subseteq X$ there is precisely one a such that $a \in A$ and $\langle A, a\rangle \in C$, so setting $h(A) := a$ defines a choice function. Explicitly,

$$h = \{\langle A, a\rangle \in C : a \in A \wedge A \in \mathcal{P}(X) \setminus \{\emptyset\}\}.$$

- (b) ⇒ (a): If X is a set of disjoint non-empty sets and h is a choice function for $\bigcup X$, then $C := h[X]$ is as required.

□

9.1 The well-ordering principle

Definition 9.3. The **well-ordering principle**, WO, is the statement that every set can be well-ordered, i.e. that for every X there exists $<$ such that $(X, <)$ is a well-order.

Lemma 9.4. *WO holds if and only if every set is equinumerous with an ordinal.*

Proof. Let X be a set. If X can be well-ordered, then it is in bijection with an ordinal by Theorem 8.19. Conversely, if $f : X \to \alpha$ is a bijection with an ordinal α, then $x < y \Leftrightarrow f(x) \in f(y)$ defines a well-order on X. □

Theorem 9.5. *$AC \Leftrightarrow$ WO.*

Proof.$\Leftarrow$: Let X be a set. By WO, X can be well-ordered, and then

$$\min : \mathcal{P}(X) \setminus \{\emptyset\} \to X$$

is a choice function.

$\Rightarrow$ (*Zermelo's theorem*): Let $h : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ be a choice function. Define by Corollary 8.24 a chain of injections $(f_\alpha)_{\alpha \in \mathbf{ON}}$ from ordinals to X such that

- $f_0 = \emptyset$
- $f_{\alpha^+} = \begin{cases} f_\alpha \cup \{\langle \alpha, h(X \setminus \operatorname{ran}(f_\alpha))\rangle\} & \text{if } X \setminus \operatorname{ran}(f_\alpha) \neq \emptyset \\ f_\alpha & \text{else} \end{cases}$
- $f_\eta = \bigcup_{\beta \in \eta} f_\beta$ for η a limit ordinal.

Then by transfinite induction, for all $\alpha \in \mathbf{ON}$:

- Either $\operatorname{dom}(f_\alpha) = \alpha$, or $\operatorname{ran}(f_\beta) = X$ for some $\beta < \alpha$.

By Hartogs' theorem, the second case must occur for some $\alpha \in \mathbf{ON}$, so let $\beta \in \mathbf{ON}$ be least such that $\operatorname{ran}(f_\beta) = X$ (which exists by Theorem 8.14(iv)). Then $\operatorname{dom}(f_\beta) = \beta$, so $f_\beta : \beta \to X$ is a bijection, and we conclude by Lemma 9.4.

□

9.2 Cardinal comparability

Definition 9.6. Cardinal comparability, CC, is the statement that the ordering $<$ on cardinalities is total, i.e. for any two sets X and Y, either $|X| \leq |Y|$ or $|Y| \leq |X|$.

Theorem 9.7. *WO $\Leftrightarrow$ CC.*

Proof. $\Rightarrow$: By comparability of well-orders (Theorem 8.9), if sets X and Y can be well-ordered then one admits an injection to the other.

$\Leftarrow$: Let X be a set. By Hartogs' Theorem, say $|\alpha| \not\leq |X|$. By CC, $|X| \leq |\alpha|$, so there exists an injection $f : X \to \alpha$, and then $x < y \Leftrightarrow f(x) \in f(y)$ defines a well-order on X.

□

9.3 Zorn's lemma

Definition 9.8. A **chain** in a partially ordered set $(X, <)$ is a subset $C \subseteq X$ which is totally ordered by $<$. An **upper bound** for a subset $A \subseteq X$ is an element $u \in X$ such that $u \geq a$ for all $a \in A$. An element $m \in X$ is *maximal* if $m \not< x$ for all $x \in X$.

Zorn's Lemma, ZL, is the statement:

- If $(X, <)$ is a partially ordered set in which every chain has an upper bound, then $(X, <)$ has a maximal element.

Theorem 9.9. *AC $\Leftrightarrow$ ZL.*

Proof. $\Rightarrow$: Let $(X, <)$ be a partial order in which every chain has an upper bound.

By WO (and Lemma 9.4), there exists a bijection $\theta : \alpha \to X$ for some ordinal α.

Define an increasing sequence of chains by transfinite recursion (Corollary 8.24):

- $C_0 := \emptyset$;
- $C_{\beta+} := \begin{cases} C_\beta \cup \{\theta(\beta)\} & \text{if } \beta \in \alpha \text{ and } \theta(\beta) > x \text{ for all } x \in C_\beta \\ C_\beta & \text{else;} \end{cases}$
- $C_\eta := \bigcup_{\beta\in\eta} C_\beta$ if η is a limit ordinal.

Then, by transfinite induction, $C_\beta \subseteq C_\gamma$ if $\beta \leq \gamma$, and each C_β is a chain.

In particular, C_α is a chain, so say $u \in X$ is an upper bound for C_α. Suppose u is not maximal, say $x > u$. Let $\beta = \theta^{-1}(x) \in \alpha$. Then $x = \theta(\beta) \in C_{\beta+} \subseteq C_\alpha$ by definition of $C_{\beta+}$, contradicting u being an upper bound for C_α. So u is a maximal element.

$\Leftarrow$: Let X be a set. Let $P' := \mathcal{P}(X) \setminus \{\emptyset\}$. Say $h \subseteq P' \times X$ is a *partial choice function* if it is a function with $\mathrm{dom}(h) \subseteq P'$ and such that $h(A) \in A$ for all $A \in \mathrm{dom}(h)$. Then the partial choice functions form a partial order with respect to inclusion, and any chain has an upper bound, namely the union of the chain. So by Zorn's Lemma, a maximal partial choice function h exists.

We conclude by showing that h is a choice function, i.e. that $\mathrm{dom}(h) = P'$. Suppose not, say $A \in P' \setminus \mathrm{dom}(h)$. Then $A \neq \emptyset$ by definition of P', so say $a \in A$. But then $h \cup \{\langle A, a\rangle\}$ is a partial choice function properly extending h, contradicting maximality of h.

□

Remark 9.10. Zorn's Lemma is often applied in the following special form (which the above proof shows is actually equivalent to our statement): if a is a set and $X \subseteq \mathcal{P}(a)$ is a non-empty set of subsets of a which is closed under unions of non-empty chains, i.e. if $\emptyset \neq C \subseteq X$ is totally ordered by inclusion then $\bigcup C \in X$, then X has a maximal element with respect to inclusion. This follows from our statement of Zorn's Lemma by considering the partial order $(X, \subseteq)$; indeed, the empty chain has an upper bound since X is non-empty, and any non-empty chain is upper-bounded by its union.

9.4 ZFC

From now on, we assume AC. We could take any of the above equivalent forms as the axiom; we use our first formulation.

AC (Choice): If X is a set of disjoint non-empty sets, then there exists a set C such that $|C \cap a| = 1$ for all $a \in X$:

$$\forall x\ (\forall y \in x\ (y \neq \emptyset \wedge \forall y' \in x\ y \cap y' = \emptyset) \rightarrow \exists z\ \forall y \in x\ (\exists u \in z \cap y\ \forall v \in z \cap y\ u = v)$$

This completes our axiom system ZFC = ZF + AC.

Fact 9.11. *Assume ZF is consistent. Gödel proved (using the* constructible universe*) that ZFC is then also consistent, i.e. that ZF does not prove ¬AC; this is covered in the part C course Axiomatic Set Theory. Paul Cohen later proved (using* forcing*) that ZF doesn't prove AC either.*

Even the weak form of Choice in which every element of X is of cardinality 2 is not a consequence of ZF (if ZF is consistent). As Russell put it: "To choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, but for shoes the Axiom is not needed" (the idea being that we can consider the set of left shoes, but the elements of a pair of socks are indistinguishable).

End of lecture 15

10 Cardinal numbers

By WO, every set is equinumerous with an ordinal (this was Lemma 9.4). Using this, we now redefine our notation $|X|$:

Definition 10.1. The **cardinality** $|X|$ of a set X is the smallest ordinal equinumerous with X:

$$|X| := \min\{\alpha \in \mathbf{ON} : \alpha \sim X\}.$$

This accords with our previous notation:

Lemma 10.2. *Let X and Y be sets.*

(i) $|X| = |Y| \Leftrightarrow X \sim Y$.

(ii) $|X| \leq |Y|$ if and only if and only if an injection $X \rightarrow Y$ exists.

Proof. (i) Immediate.

(ii) By Lemma 8.32, an injection $|X| \rightarrow |Y|$ exists iff $|X| \leq |Y|$, and the result follows.

□

Lemma 10.3. *For an ordinal α, the following are equivalent.*

(i) $\alpha = |X|$ for some set X.

(ii) $\alpha = |\alpha|$.

(iii) For all $\beta \in \alpha$, $\beta \not\sim \alpha$.

Definition 10.4. An ordinal satisfying these properties is called a **cardinal** or a **cardinal number**. The infinite cardinals are sometimes also known as *initial ordinals*.

The class of cardinals is denoted **CN**.

Proof.
- (i) $\Rightarrow$ (ii): If $\alpha = |X|$ then $\alpha \sim X$ so $|\alpha| = |X| = \alpha$.
- (ii) $\Rightarrow$ (i): Immediate.
- (ii) $\Leftrightarrow$ (iii): Immediate from the definition, since $\alpha \sim \alpha$.

□

Lemma 10.5. *(i) If κ is a cardinal, then there exists a cardinal greater than κ.*

(ii) If K is a set of cardinals, then $\bigcup K$ is a cardinal.

Proof. (i) By Corollary 6.25, $|\mathcal{P}(\kappa)| > |\kappa| = \kappa$. Alternatively: By Hartogs' Theorem and cardinal comparability, there is an ordinal α such that $|\alpha| > |\kappa| = \kappa$.

(ii) $\bigcup K$ is an ordinal by Lemma 8.20(a)(iii). Suppose $|\bigcup K| \in \bigcup K$. Then $|\bigcup K| \in \kappa$ for some $\kappa \in K$, so $|\bigcup K| < |\kappa|$ since κ is a cardinal, contradicting $\kappa \subseteq \bigcup K$. So $|\bigcup K| = \bigcup K$.

□

This lemma justifies the following definition:

Definition 10.6. Define by transfinite recursion (Corollary 8.24) a class function $\mathbf{ON} \to \mathbf{CN}$; $\alpha \mapsto \aleph_\alpha$ such that:

- $\aleph_0 = \omega$;
- $\aleph_{\alpha^+}$ is the smallest cardinal greater than $\aleph_\alpha$;
- $\aleph_\eta = \bigcup_{\beta \in \eta} \aleph_\beta$ if η is a limit ordinal.

In particular, we redefine $\aleph_0 := \omega = |\mathbb{N}|$.

We also write $\aleph_\alpha$ as ω_α when we think of it as an ordinal rather than a cardinal (see below).

Theorem 10.7.

(i) $\aleph_\alpha \geq \alpha$ for all $\alpha \in \mathbf{ON}$.

(ii) If $\alpha < \beta \in \mathbf{ON}$ then $\aleph_\alpha < \aleph_\beta$.

(iii) Every infinite cardinal is of the form $\aleph_\alpha$ for some $\alpha \in \mathbf{ON}$.

*(iv) **CN** is a proper class*[10].

Proof. (i) By transfinite induction on α.

(ii) By transfinite induction on β.

[10] This is known as *Cantor's paradox.*

(iii) Let κ be an infinite cardinal. Consider the set

$$\alpha := \{\beta : \aleph_\beta < \kappa\} = \{\beta \in \kappa : \aleph_\beta < \kappa\},$$

where the equality is by (i). Then α is an initial segment of κ by (ii), so α is an ordinal. So $\alpha \notin \alpha$, hence $\aleph_\alpha \geq \kappa$. We conclude by showing $\aleph_\alpha \leq \kappa$.

If $\alpha = 0$, this follows from κ being infinite.

If α is a limit ordinal, then $\aleph_\alpha = \bigcup_{\beta\in\alpha} \aleph_\beta \leq \kappa$ since each $\aleph_\beta < \kappa$.

If α is a successor ordinal, say $\alpha = \gamma^+$, then $\aleph_\gamma < \kappa$, so $\aleph_\alpha = \aleph_{\gamma^+} \leq \kappa$ by definition of $\aleph_{\gamma^+}$.

(iv) By (ii) and (iii), $\aleph_\alpha \mapsto \alpha$ is a well-defined surjective class function $\mathbf{CN} \setminus \omega \to \mathbf{ON}$. So if **CN** were a set, then by Replacement so would be **ON**, contradicting Theorem 8.16.

□

10.1 Cardinal arithmetic with Choice

Definition 10.8. We now consider the cardinal arithmetic operations (addition, multiplication, and exponentiation) defined in Definition 6.20 as operations on cardinals:

$$\kappa + \lambda := |(\kappa \times \{0\}) \cup (\lambda \times \{1\})| \quad \kappa \cdot \lambda := |\kappa \times \lambda| \quad \kappa^\lambda := |\{f : \lambda \to \kappa\}|.$$

Warning: This leads to an unfortunate ambiguity, since these cardinal arithmetic operations rarely agree with the ordinal arithmetic operations. In practice, we get around this by notational conventions: we reserve $\kappa, \lambda, \mu, \nu$ and $\aleph_\alpha$ for cardinals, and arithmetic expressions involving these and $|X|$ refer to cardinal arithmetic, while expressions involving $\alpha, \beta, \gamma, \delta$ and ω_α refer to ordinal arithmetic.[11]

In ZFC, cardinal addition and multiplication are very simple:

Theorem 10.9. *Let κ be an infinite cardinal.*

(i) $\kappa \cdot \kappa = \kappa$

(ii) If λ is a cardinal $1 \leq \lambda \leq \kappa$, then $\kappa + \lambda = \kappa = \kappa \cdot \lambda$.

(iii) If λ is an infinite cardinal, then $\kappa + \lambda = \max(\kappa, \lambda) = \kappa \cdot \lambda$.

Proof. (i) By transfinite induction. Assume $|\alpha| \cdot |\alpha| = |\alpha|$ for all infinite ordinals $\alpha < \kappa$. Then

$$|\alpha| \cdot |\alpha| < \kappa \text{ for all } \alpha < \kappa; \tag{*}$$

for finite α, this is because $|\alpha| \cdot |\alpha|$ is finite by Proposition 6.22(b).

Define an ordering $\lhd$ on $\kappa \times \kappa$ by

$$(\alpha, \beta) \lhd (\alpha', \beta') \Leftrightarrow (\alpha, \beta, \max(\alpha, \beta)) <_r (\alpha', \beta', \max(\alpha', \beta'))$$

[11] To add to the confusion, κ^+ is usually defined as the smallest cardinal greater than κ (so $\aleph_{\alpha^+} = (\aleph_\alpha)^+$), which is not the ordinal successor unless κ is finite.

where $<_r$ is reverse lexicographic order.

This is a well-order, so $(\kappa \times \kappa, \lhd)$ is isomorphic to an ordinal γ. Let S be a proper initial segment, say $S = (\kappa \times \kappa)_{\lhd(\alpha,\beta)}$. Set $\delta := \max(\alpha, \beta)$. Then $S \subseteq \delta^+ \times \delta^+$, and $\delta^+ \in \kappa$ (indeed, this holds if δ is finite since κ is infinite, and if δ is infinite then $|\delta^+| = |\delta| < \kappa$), so by (*), $|S| \leq |\delta^+| \cdot |\delta^+| < \kappa$. Hence $\gamma \leq \kappa$, since otherwise γ would have a proper initial segment $\gamma_{<\kappa}$ of cardinality κ.

So $\kappa \cdot \kappa \leq \kappa$. Conversely, $\kappa = |\kappa \cdot \{0\}| \leq \kappa \cdot \kappa$. So $\kappa \cdot \kappa = \kappa$.

(ii) By the monotonicity properties of Proposition 6.22(c) and (i),

$$\kappa \leq \kappa + \lambda \leq \kappa + \kappa = \kappa \cdot 2 \leq \kappa \cdot \kappa = \kappa$$
$$\kappa \leq \kappa \cdot \lambda \leq \kappa \cdot \kappa = \kappa$$

(iii) By (ii) and commutativity of cardinal addition and multiplication (Proposition 6.22(a)).

□

Lemma 10.10. *If $f : X \to Y$ is a surjection, then $|X| \geq |Y|$.*

Proof. Let h be a choice function for X. Then $g(y) := h(\{x \in X : f(x) = y\})$ defines an injection $Y \to X$. So $|Y| \leq |X|$. □

Theorem 10.11. *A countable union of countable sets is countable.*

More generally, if κ is an infinite cardinal, and X is a set such that $|X| \leq \kappa$ and $|a| \leq \kappa$ for all $a \in X$, then $|\bigcup X| \leq \kappa$.

Proof. For every $a \in X$, there exists an injection $a \to \kappa$. By Choice, we can uniformly choose such injections: let I_a be the set of injections $f : a \to \kappa$, let h be a choice function on $\bigcup\{I_a : a \in X\}$, and let $f_a := h(I_a)$.

Let $g : X \to \kappa$ be an injection.

Then $Z := \{\langle g(a), f_a(x)\rangle : x \in a \in X\}$ is a subset of $\kappa \times \kappa$, so $|Z| \leq \kappa \cdot \kappa = \kappa$.

Finally, $\langle g(a), f_a(x)\rangle \mapsto x$ is a surjection $Z \to \bigcup X$, so by Lemma 10.10, $|\bigcup X| \leq |Z| \leq \kappa$. □

Remark. Choice was essential here: ZF does not prove that a countable union of countable sets is countable.

10.2 Cardinal exponentation and CH (off-syllabus)

In contrast, very little is determined by ZFC about cardinal exponentiation.

Definition 10.12.

- The **Continuum Hypothesis** (**CH**) is the assertion: $2^{\aleph_0} = \aleph_1$. In other words: every uncountable subset of $\mathbb{R}$ is in bijection with $\mathbb{R}$.
- The **Generalised Continuum Hypothesis** (**GCH**) is the assertion: $2^{\aleph_\alpha} = \aleph_{\alpha^+}$ for all ordinals α.

Fact 10.13.

- *CH is independent of ZFC. That is, assuming ZFC is consistent, it proves neither CH nor $\neg$CH, so both ZFC+CH and ZFC+$\neg$CH are consistent. The same goes for GCH. As with AC (Fact 9.11), consistency of ZFC+GCH is due to Kurt Gödel and is covered in Part C, and that of ZFC+$\neg$CH (hence also ZFC+$\neg$GCH) is due to Paul Cohen using forcing.*
- *Any counterexample $X \subseteq \mathbb{R}$ to CH has to be "complicated": it can not be Borel, nor the projection of a Borel set (an* analytic *set).*
- $2^{\aleph_0} \neq \aleph_\omega$*. More generally, $2^{\aleph_0}$ is not of the form $\bigcup_{i\in\omega} \alpha_i$ for any ordinals $\alpha_i < 2^{\aleph_0}$ (i.e. $2^{\aleph_0}$ does not have* countable cofinality*). This is all ZFC tells us about $2^{\aleph_0}$, in the sense that for any $\aleph_\alpha$ which is not of this form, it is consistent with ZFC that $2^{\aleph_0} = \aleph_\alpha$.*

End of lecture 16

11 Example: Infinite dimensional vector spaces

In this section, we illustrate the use of set theory in mathematics by developing some of the basic theory of linear algebra without assuming finite dimensionality. (This material is not on the syllabus, but the set theory techniques we use are.)

We use without proof the finite dimensional results (covered in Prelims).

Let V be a vector space over a field K. This implies that V and K are sets, and the associated algebraic operations ($+, \cdot : K \times K \to K$, and $+ : V \times V \to V$, and scalar multiplication $\cdot : K \times V \to V$) are functions.

Definition 11.1. A subset $B \subseteq V$ is

- **linearly independent** if no non-trivial finite linear combination of elements of B is 0, i.e. if for any $n \in \mathbb{N}$, $a_1, \ldots, a_n \in K$, and $b_1, \ldots, b_n \in V$,

$$a_1 \cdot b_1 + \ldots + a_n \cdot b_n = 0 \;\Rightarrow\; a_1 = \ldots = a_n = 0;$$

- **spanning** if $V = \langle B \rangle$ where

$$\langle B \rangle = \{a_1 \cdot b_1 + \ldots + a_n \cdot b_n : n \in \mathbb{N},\ a_1, \ldots, a_n \in K,\ b_1, \ldots, b_n \in V\}.$$

- a **basis** if B is both linearly independent and spanning.

Theorem 11.2. *(i) A basis exists.*

*(ii) Any two bases have the same cardinality. This cardinality is called the **dimension** of V.*

Proof. (i) We apply Zorn's Lemma. Consider the set $\mathcal{I}$ of linearly independent subsets of V as a partial order, ordered by inclusion, $\subseteq$. If $C \subseteq \mathcal{I}$ is a chain, then its union is also linearly independent, since any finitely many elements of $\bigcup C$ are already elements of some $I \in C$. So by Zorn's Lemma, there exists a maximal element $B \in \mathcal{I}$. We conclude by showing that B is spanning. Suppose not, say $v \in V \setminus \langle B \rangle$. Then one verifies directly that $B \cup \{v\}$ is linearly independent, and $v \notin B$, contradicting maximality of B.

(ii) Let B and B' be bases. The case where B or B' is finite was done in Prelims. So suppose B and B' are infinite.

Let $\mathcal{P}^{<\omega}(B) := \{B_0 \subseteq B : |B_0| < \aleph_0\}$ be the set of finite subsets of B. Then $|\mathcal{P}^{<\omega}(B)| = |B|$, since $\mathcal{P}^{<\omega}(B) = \bigcup_{n\in\mathbb{N}} B^{(n)}$ where $B^{(n)} := \{B_0 \subseteq B : |B_0| = n\}$, and $B^n \to B^{(n)}; (b_1, \ldots, b_n) \mapsto \{b_1, \ldots, b_n\}$ is a surjection, so $|B^{(n)}| \leq |B^n| = |B|$ so $|\mathcal{P}^{<\omega}(B)| \leq |B|$ by Theorem 10.11.

If $B_0 \in \mathcal{P}^{<\omega}(B)$, then $\langle B_0\rangle \cap B'$ is finite by the finite-dimensional case. But $V = \langle B\rangle = \bigcup\{\langle B_0\rangle : B_0 \in \mathcal{P}^{<\omega}(B)\}$, so $B' = \bigcup\{\langle B_0\rangle \cap B' : B_0 \in \mathcal{P}^{<\omega}(B)\}$ is a union of $|\mathcal{P}^{<\omega}(B)| = |B|$ finite sets, so by Theorem 10.11 again, $|B'| \leq |B|$.

By symmetry, $|B'| = |B|$.

□

A References

[Copied directly from Jonathan Pila's notes, with a few amendments]

Text to expand and supplement these notes:

[1] D. Goldrei, *Classic set theory,* Chapman and Hall/CRC, Boca Raton, 1998.

Highly recommended general audience sources:

[2] S. Aaronson, The Busy Beaver Frontier, https://scottaaronson.blog/?p=4916

[3] S. Lavine, *Understanding the infinite,* Harvard, 1994.

[4] R. Rucker, *Infinity and the mind,* Princeton University Press, 1995.

[5] J. Stillwell, *Roads to Infinity*, CRC Press, AK Peters, 2010.

Other textbooks (more advanced) and notes:

[6] P. Aczel, *Non-well founded sets,* CSLI Lecture Notes, 14, Stanford, CA. On web.

[7] M. Hils and F. Loeser, *A First Journey through Logic*, AMS Student Mathematical Library volume 89, 2019.

[8] T. Jech, *Set theory,* Academic Press, 1978.

[9] R. Knight, b1 Set Theory Lecture Notes.

[10] K. Kunen, *Set theory,* North-Holland, Amsterdam, 1980.

[11] A. Levy, *Basic Set Theory,* Springer, Berlin, 1979; reprinted by Dover.

Original texts and interpretations:

[12] G. Cantor, *Contributions to the founding of the theory of transfinite numbers* (translated by P. Jourdain), Dover, 1955. p85.

[13] P. Cohen, *Set theory and the continuum hypothesis,* W. A. Benjamin, 1966.

[14] P. Cohen and R. Hersch, *Non-Cantorian set theory, Scientific American* **217** (1967), 104–116.

[15] R. Dedekind, *Essays on the theory of numbers,* Dover, 1963.

[16] K. Gödel, *The consistency of the axiom of choice and the generalized continuum hypothesis with the axioms of set theory,* Annals of Mathematics Studies 3, Princeton University Press, 1940.

[17] K. Gödel, What is Cantor's continuum problem? *American Mathematical Monthly,* **54** (1964), 515–525. Revised version in Benacerraf and Putnam, *Philosophy of mathematics: selected readings,* Prentice-Hall, 1964.

Modern perspectives, histories, and commentaries:

[18] S. Feferman, *In the Light of Logic,* OUP, Oxford, 1998.

[19] A. Fraenkel, Y. Bar Hillel, A. Levy, *Foundations of set theory,* North-Holland, Amsterdam, 1973.

[20] J. Ferreiros, On the relations between Georg Cantor and Richard Dedekind, *Historia Math.* **20** (1993), 343–363.

[21] M. Hallett, Cantorian set theory and limitation of size, Oxford Logic Guides **10**, OUP, 1984.

[22] J. Hamkins, Is the dream solution of the continuum hypothesis attainable? *Notre Dame J. Symbolic Logic* **56** (2015), 135–145.

[23] P. Maddy, *Defending the Axioms,* OUP, 2011.

[24] Y. Manin, George Cantor and his heritage, http://arxiv.org/abs/math/0209244 (just ignore the stuff about $P \neq NP$)

[25] C. McLarty, What does it take to prove Fermat's Last Theorem? Grothendieck and the logic of number theory, *Bull. Symbolic Logic* **16** (2010), 359–377.

[26] G. H. Moore, *Zermelo's axiom of choice,* Springer, 1982.

[27] U. Rehmann (Editor-in-Chief), *Encylopedia of Mathematics*, entry on "Antinomy", https://www.encyclopediaofmath.org/index.php/Antinomy

[28] S. Shelah, Logical dreams, *Bull. Amer. Math. Soc.* **40** (2003), 203–228.

[29] J. Stillwell, The continuum problem, *Amer. Math. Monthly* **109** (2002), 286–297.

[30] W. H. Woodin, Strong axioms of infinity and the search for V, *Proc. ICM Hyderabad* (2010). Available online at http://www.mathunion.org/ICM/ICM2010.1/ Also the corresponding lecture can be viewed online.

A.1 Exam errata

Past exams obtained from some sources are “as given” while those obtained from the Mathematical Institute sometimes have errors (or obscurities) corrected. The latter are therefore recommended. The solutions which are available from the MI archive often contain mistakes or omissions, but are usually useful as at least a rough guide to a solution.

Here are a few errata from recentish exams:

2017.2.b.ii. The set X should be assumed non-empty.

2015.1.b.ii. Remove ‘strictly’, or its wrong!

2015.2.b.ii. Beware that ‘contained’ here means as a subset (not as an element).

2015.3.c.ii. Has a too easy (but correct) solution.

2014.2.c.iii. This is too hard and should not be attempted.