CHAPTER 1

Week 1
Sets, Classes, and The Cumulative Hierarchy

1. Sets

Sets are collections of objects. The collection of all students registered in the Advanced Logic module at King's College London in 2020 is a set. The collection of all prime numbers less than 1,000,000 is a set. This definition of set relies on a clear notion of 'object'. This, in turn, requires an uncontroversial metaphysics of objects, which we don't have. Luckily, for the purpose of developing mathematics – and therefore scientifically applicable mathematics – and portions of formal philosophy, we don't need that much.

There are two main ways to refer to sets. The first is by listing their elements – by enclosing them in curly brackets $\{, \}$ –, as in

$$\{\text{Caravaggio}\}, \qquad \{\text{Michelangelo Merisi}\},$$
$$\{1, 2, 3\}, \qquad \{\sqrt{2}, \pi\}.$$

The second is by description, by means of the so called *abstraction terms*:

$$\{x \mid x \text{ is the painter of } \textit{The Calling of Saint Matthew}\}$$
$$\{x \mid x \text{ is a positive integer smaller than } 4\}$$

Notice, however, that *not all descriptions give rise to sets!* We will treat this point more carefully in the next section.

The identity conditions for sets – cf. Quine's 'there's no entity without identity' [**vOQ69**, p. 23] – are encapsulated in the well-known principle of extensionality:

PRINCIPLE OF EXTENSIONALITY: Two sets are identical if and only if they have the same elements. In symbols:

$$x = y \leftrightarrow \forall u(u \in x \leftrightarrow u \in y)$$

Sets are then completely characterized by their elements, and not by their *mode of presentation* – i.e. the way we refer to them. For instance:

$$\{\text{Caravaggio}\} = \{x \mid x \text{ is the painter of } \textit{The Calling of Saint Matthew}\},$$
$$\{x \mid x \text{ is a positive integer smaller than } 4\} = \{1, 2, 3\}.$$

Similarly, one has that (as far as we know):

$$\{x \mid x \text{ is an animal with a heart}\} = \{x \mid x \text{ is an animal with a kidney}\}.$$

Therefore, even though arguably the property of having a heart is not the same as the property of having a kidney, the *extension* of these two properties, i.e. the set of objects that have such properties, is the same. The philosophical question whether a scientific (broadly construed) worldview needs properties or concepts together with sets is still substantially open.[1]

As anticipated above, however, the entire universe of mathematical objects can be constructed without resorting to objects located in space and time such as Caravaggio. One of the fundamental building blocks of such construction is the following *set-existence principle*:

EMPTY SET PRINCIPLE: there is a set containing no elements:

$$\exists y \forall u (u \notin y)$$

EXERCISE 1.1. Show, using the EXTENSIONALITY PRINCIPLE, that there is a unique empty set.

Since there is only one empty set, we can safely denote it with the 'proper name' $\emptyset$. We collect a few simple facts concerning the empty set. As usual, we let

$$x \subseteq y :\leftrightarrow \forall u (u \in x \longrightarrow u \in y)$$
$$x \subset y :\leftrightarrow x \subseteq y \wedge x \neq y$$

FACT 1.

(i) *For any collection* A, $\emptyset \subseteq A$

(ii) $\emptyset \subseteq \emptyset$, $\emptyset \notin \emptyset$, $\{\emptyset\} \in \{\{\emptyset\}\}$, $\{\emptyset\} \not\subseteq \{\{\emptyset\}\}$.

PROOF. Exercise. *qed*

DEFINITION 1 (POWER SET). *The power set* $\mathcal{P}(x)$ *of a set* x *is the set* $\{y \mid y \subseteq x\}$.

[1] For an overview, see [**BM03**].

EXERCISE 1.2. Show that $\mathcal{P}(x)$, if it exists, is unique.

From the empty set, we can basically construct the entire universe of sets (and thus of mathematics). We only need to close the empty set under two basic operations. One is implicit in the definition of power set, and tells us that one *can always* collect subsets of a given set.

POWER SET PRINCIPLE: For any set x, $\mathcal{P}(x)$ exists and is a set.

The second operation is is a generalization of the well-known notion of *union* of sets. We know that

$$x \cup y := \{u \mid u \in x \vee u \in y\}$$

This operation (iterated finitely many times) enables us to define *finite unions*:

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

But what about infinite ones? Suppose $x := \{x_i \mid i \in \mathbb{N}\}$. How do we collect all x_i together? We need a further operation

$$\bigcup x = \{u \mid (\exists y \in x)(u \in y)\}$$

Again implicit in this definition is the possibility of forming unions of all elements of any given set.[2]

PRINCIPLE OF UNION: For any set x, $\bigcup x$ exists and is a set.

Using the PRINCIPLE OF UNION, we can readily obtain the infinite union of all elements of x_i.

EXERCISE.

(i) Which collection is $\bigcup\{\{1\}, \{1, 2\}, \{1\}\}$?

(ii) Which collection is $\bigcup\{\{0\}, \{1\}, \{2\}, \{3\}, \ldots\}$?

The empty set, power sets, and unions give us basically all we need to outline the so called *cumulative hierarchy of sets*, the universe of all sets. I said 'outline', because a precise definition will only be available once the full axiomatic development of set theory will be given. However, at this

[2] A quick remark on notation. We write

$$(\exists u \in y)\, \varphi \qquad \text{for} \qquad \exists u(u \in y \wedge \varphi),$$

and

$$(\forall u \in y)\, \varphi \qquad \text{for} \qquad \forall u(u \in y \to \varphi).$$

stage we can start by giving the first steps of this process, and leave the details for later.

We want to define the universe V of all sets in stages. So we put

$$\mathrm{V}_0 := \emptyset$$
$$\mathrm{V}_{n+1} := \mathcal{P}(\mathrm{V}_n)$$

Once we have completed the finite stages, we need to find a way to index later stages. This is the job of *Cantor's theory of ordinal numbers*, that we can only briefly introduce now. Roughly speaking, ordinals extend natural numbers $\mathbb{N}$ by introducing *infinite* numbers. The first infinite numbers is denoted with ω (read 'omega'). So we put:

$$\mathrm{V}_\omega := \bigcup\{\mathrm{V}_i \mid i \in \mathbb{N}\} = \{u \mid (\exists i \in \mathbb{N})(u \in \mathrm{V}_i)\}$$

The definition can then be iterated. So $\mathrm{V}_\omega + 1 := \mathcal{P}(\mathrm{V}_\omega)\ldots$. The resulting picture of V is displayed in Figure 1.

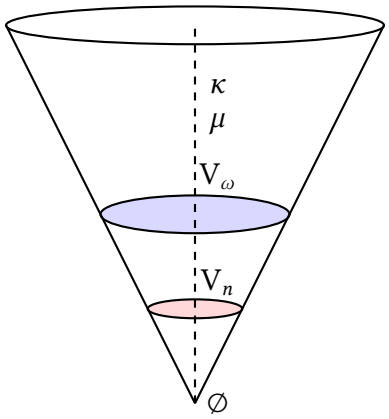

FIGURE 1. Cumulative hierarchy

EXERCISE. Show that $\mathrm{V}_3 = \mathrm{V}_2 \cup \mathcal{P}(\mathrm{V}_2)$. (Notice that this is true for every stage).

EXERCISE. Define the *rank* of a set x, written $\rho(x)$, to be the least α such that $x \subseteq \mathrm{V}_\alpha$. What is $\rho(\{\emptyset\})$? What is $\rho(\{\{\emptyset\}, \{\{\{\emptyset\}\}\}\})$?

2. Classes

In introducing expressions of the form $\{x \mid \Phi(x)\}$, we have put no limits to the possibility of forming sets given a suitable specifying condition Φ. Such limits, however, exist.

FACT 2 (RUSSELL'S PARADOX). *The collection* $\mathrm{R} = \{x \mid x \notin x\}$ *is not a set.*

PROOF. Suppose R is a set. Since we assume classical logic, either $\mathrm{R} \in \mathrm{R}$ or $\mathrm{R} \notin \mathrm{R}$. If the former, then R is one of the sets that satisfies the condition $x \notin x$; therefore $\mathrm{R} \notin \mathrm{R}$, and so both $\mathrm{R} \in \mathrm{R}$ and $\mathrm{R} \notin \mathrm{R}$. If the latter, again R is one of the sets that satisfies the condition $x \notin x$. Therefore, $\mathrm{R} \in \mathrm{R}$ and $\mathrm{R} \notin \mathrm{R}$. In either case we obtain a contradiction. So R cannot be a set. *qed*

The totality of the collections we refer to are called *classes*. So all sets are classes, but not viceversa: classes that are not sets are called *proper classes*. It then follows from FACT 2 that R is a proper class.

But if not all conditions define a set, which ones do? The following principle gives us an answer.

> SUBSET PRINCIPLE: let $\Phi(\cdot)$ be a determinate property and x a set. Then $\{u \in x \mid \Phi(u)\}$ is a set.

But for the answer to be satisfactory, one also has to know what a determinate property is. To keep things simple,[3] the standard reply that set theory offers is that a property $\Phi(x)$ is determinate when it can be expressed in a specific formal language, the language $\mathcal{L}_\in$ of set theory. The SUBSET PRINCIPLE is then based on a *syntactic* restriction: the only properties that can be employed in defining sets are the ones that belong to a specific class of syntactic objects.

QUESTION 1. Is this notion of determinacy satisfactory?

COROLLARY 1. V *is not a set.*

PROOF. Suppose it is. Then by the SUBSET PRINCIPLE, the class $\{u \in \mathrm{V} \mid u \notin u\}$ is a set. But this contradicts FACT 2. *qed*

EXERCISE. Fill in the details of the proof of Corollary 1.

EXERCISE. Consider a set x and the set $y := \{u \in x \mid u \notin u\}$. Show that $y \notin x$.

[3]Notice that, depending on the answer, one might end up with a formulation of the subset principle that is difficult to even write down. For instance,

CHAPTER 2

Week 2
The Axioms of Set theory

The set theory that we study in the first part of the module – as we have seen, inspired by the cumulative hierarchy –, is called ZFC, standing for 'Zermelo-Fraenkel set theory with the axiom of Choice'. ZFC is a first-order theory. That is, it is obtained by extending pure predicate (aka first-order) logic with identity with a collection of first-order sentences, its *axioms*.

The language $\mathcal{L}_\in$ of ZFC is a language containing the usual *logical symbols*[1]

$$\neg, \vee, \wedge, \longrightarrow, \leftrightarrow, \forall, \exists, =$$

and *only one* nonlogical symbols, the membership (binary) relation $\in$. Therefore, formulas of $\mathcal{L}_\in$ are built from atomic formulas $x \in y$ and $x = y$ by closing them under the logical operations. When writing $\varphi(x_1, \dots, x_n)$, we will assume that all free variables in φ are among $x_1, \dots, x_n$.

Although we only have one type of objects in ZFC, we will avail ourselves with a way of referring to *classes*, especially *proper classes* – in particular, we will use capital letters for classes. In fact, this is what we have already done in the previous section, when for instance referring to

$$\mathrm{V} = \{x \mid x = x\},$$

or

$$\mathrm{R} = \{x \mid x \notin x\}.$$

Recall that the main difference between classes and sets is the way they are formed. The former satisfy the so-called *comprehension principle*:

$$\exists X \forall u (u \in X \leftrightarrow \varphi(u, v_1, \dots, v_n)),$$

where $v_1, \dots, v_n$ are called *parameters* (essentially, class talk is shorthand for formulas that *define them*, so that R is shorthand for $x \notin x$, and V is shorthand for $x = x$). The latter (sets) satisfy restricted principles such as

[1] Notice that identity is treated as a logical symbol.

the SUBSET PRINCIPLE that we have considered before, and that will correspond to a specific axiom schema that we will introduce in a moment.

1. Extensionality

The first axiom is a familiar one. It is often thought to be *constitutive* of the *concept* of set. Sets are fully characterized by their elements:

(EXTENSIONALITY) $\forall x \forall y(x = y \leftrightarrow \forall u(u \in x \leftrightarrow u \in y))$

'sets are identical iff they have the same elements'.

EXERCISE 2.1. In fact, only the direction

$$\forall x \forall y(\forall u(u \in x \leftrightarrow u \in y) \rightarrow x = y)$$

of extensionality would suffice for our development of set theory. Why?

2. Pairing

The first set-existence axiom that we consider enables us to form sets containing two given elements.

(PAIRING) $\forall x \forall y \exists z \forall u(u \in z \leftrightarrow (u = x \vee u = y))$

'Given x and y, there is the set $\{x, y\}$.'

LEMMA 1.

(i) *Given sets x, y, the set $\{x, y\}$ obtained by pairing is unique.*

(ii) *Given a set x, the* singleton $\{x\}$ *exists and is unique.*

PROOF. Exercise. For (ii), you need to use both PAIRING and EXTENSIONALITY. *qed*

It is an immediate consequence of EXTENSIONALITY that, for any x, y, $\{x, y\} = \{y, x\}$. In fact, the PAIRING AXIOM is often referred to as the axiom of *unordered* pairing. *Ordered pairs* – written (x, y) – are characterized by the principle

(1) $$(x, y) = (u, v) \leftrightarrow (x = u \wedge y = v)$$

Ordered pairs (x, y) can be defined in ZFC as the set $\{\{x\}, \{x, y\}\}$ – this is the so-called Kuratowski definition of ordered pair. Of course, given any sets x, y, PAIRING enable us to construct $\{\{x\}, \{x, y\}\}$, and EXTENSIONALITY ensures its uniqueness.

EXERCISE 2.2. Show that Kuratowski's definition of ordered pair satisfies (1).

What we have said generalizes to triples, quadruples, quintuples, ...It suffices to put

$$\begin{aligned}(x, y, z) &:= ((x, y), z) := \{\{(x, y)\}, \{(x, y), z\}\}\\ (x, y, z, u) &:= ((x, y, z), u) := \{\{(x, y, z)\}, \{(x, y, z), u\}\}\\ \vdots &\qquad\quad \vdots\end{aligned}$$

Exercise 2.3. Show that

$$(x_1, \dots, x_n) = (y_1, \dots, y_n) \leftrightarrow x_1 = y_1 \wedge \dots \wedge x_n = y_n.$$

3. Union

The principle of union considered in the previous chapter is also an axiom of ZFC:

(union)

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

'Give a set x, the set $\bigcup x = \{u \mid (\exists z \in x)\, u \in z\} = \bigcup \{z \mid z \in x\}$ exists'.

In the previous chapter we have introduced the union of x, $\bigcup x$, as a generalization of the union of two sets. In ZFC, the order is reversed. We can define:

$$\begin{aligned}x \cup y &:= \bigcup \{x, y\}\\ x \cup y \cup z &:= (x \cup y) \cup z\\ &\vdots\end{aligned}$$

Exercise 2.4. Show that: $\bigcup \{x\} = x$; $\bigcup (x \cup y) = \bigcup x \cup \bigcup y$.

Exercise 2.5. Show that $\mathrm{P} := \{(x, y) \mid x, y \in \mathrm{V}\}$, the class of all ordered pairs, is a proper class. (Hint: we need to use the axiom of union twice).

4. Separation (Subset)

The next set-existence axiom is the precise counterpart of the subset principle that we have considered in the previous section: for any formula $\varphi(u, v)$ of $\mathcal{L}_\in$:

(separation) $\quad \forall x \exists y \forall u (u \in y \leftrightarrow u \in x \wedge \varphi(u, v))$

'given a set x, the subset of x satisfying φ is a set'.

The first thing to notice about separation is that, unlike pairing, it is not a single sentence, but is is an *axiom schema*, that is a recipe to generate

infinitely many sentences. In fact, the expression $\varphi(u)$ in SEPARATION is a placeholder for an arbitrary formula of $\mathcal{L}_\in$.[2]

An immediate consequence of SEPARATION is the existence of the *intersection* and of the set-theoretic difference of two sets. Given sets x and y, the former can be defined as

$$x \cap y := \{u \in x \mid u \in y\},$$

whereas the latter as

$$x \setminus y := \{u \in x \mid u \notin y\}.$$

EXERCISE 2.6. Show that, given sets x, y, $x \cap y$ and $x \setminus y$ are unique.

EXERCISE 2.7. The *symmetric difference* of x and y, written $x \triangle y$, is defined as

$$(x \setminus y) \cup (y \setminus x).$$

Is the equation $\{\emptyset, \{\emptyset\}\} \triangle \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{\emptyset\}$ true or false?

The operation of intersection gives us a way to understand why the axiom schema of SEPARATION is often called 'Subset Axiom'. Consider in fact a class:

$$X = \{u \mid \varphi(u, v_1, \dots, v_n)\}.$$

Then SEPARATION tells us that

$$\exists y \forall u (u \in y \leftrightarrow u \in x \wedge \varphi(u, v_1, \dots, v_n)),$$

that is that the set

$$y = \{u \mid u \in x \wedge u \in X\} = X \cap x$$

exists.

Two sets x and y are called *disjoint* if $x \cap y = \emptyset$. We define, for x a nonempty set:

$$\bigcap x := \{u \mid (\forall y \in x)\, u \in y\}.$$

FACT 3.

(i) *For any x, y: $x \cap y = \bigcap\{x, y\}$.*

(ii) *For any class A of sets, $\bigcap A$ is a set.*

PROOF. Exercise. (Hint for (ii): it's important that A is a class of *sets*, and that there is an $a \in A$). *qed*

[2]This situation should not be unfamiliar: the rules of natural deduction are in fact rule schemata.

The notation just introduced enables us to extend our conceptual inventory for pairs and sequences. Give a pair $x = (u, v)$, we let

$$(x)_0 = u, \qquad (x)_1 = v$$

These operations – often called *projection functions* – can be properly defined in our setting by letting

$$(2) \qquad (x)_0 = \bigcup\bigcap x, \qquad (x)_1 = \begin{cases} \bigcup(\bigcup x \setminus \bigcap x), & \text{if } \bigcup x \neq \bigcap x; \\ \bigcup\bigcup x, & \text{otherwise.} \end{cases}$$

Exercise 2.8. Show that, for $x = (u, v)$, our definition of the projection functions (2) does indeed yield the correct results.

A final important aspect of our formulation of SEPARATION is that it only contains one parameter – the variable v in our formulation. However, we can show that, given PAIRING and projection, a more general formulation of SEPARATION is available, featuring arbitrarily many (finite) parameters:

(SEPARATION*) $\qquad \forall x \exists y \forall u(u \in y \leftrightarrow u \in x \wedge \chi(u, v_1, \dots, v_n))$

for any formula $\chi(u, v_1, \dots, v_n)$ of $\mathcal{L}_\in$.

In fact, we can let

$$\psi(u, v) \ := \exists v_1 \dots \exists v_n(v = (v_1, \dots, v_n) \wedge \chi(u, v_1, \dots, v_n)).$$

SEPARATION then entails that

$$\exists y \forall u(u \in y \leftrightarrow u \in x \wedge \psi(u, v_1, \dots, v_n)).$$

Exercise 2.9. Fill in the details in this argument showing that SEPARATION entails SEPARATION*.

5. Emptyset

In the previous chapter we have introduced the EMPTY SET PRINCIPLE. As an axiom, it is redundant, as it is a theorem of ZFC as presented here:

(EMPTYSET) $\qquad \exists y \forall u(u \notin y).$

The axiom of INFINITY below in fact will entail the existence of at least one set, that is it will entail that $\exists x\, x = x$.

Lemma 2. *If* $\exists x\, x = x$, *then* $\exists y \forall u(u \notin y)$.

PROOF. We know by assumption that there is at least one set. Call one such set a. Define $\emptyset = \{u \in a \mid u \neq u\}$. I leave it as an exercise to complete the details (i.e.: which principles are employed in my reasoning?). *qed*

6. Power Set

Also the power set principle has a counterpart in the axioms of ZFC:

(POWER SET) $\forall x \exists y \forall u (u \in y \leftrightarrow u \subseteq x)$

'for any x, its power set $\mathcal{P}(x)$ exists and is a set.'

The *product* of two sets is defined as:

$$x \times y := \{(u, v) \mid u \in x \wedge v \in y\},$$

that is, $x \times y$ is the set of pairs whose first component belongs to x, and whose second component belongs to y. Notice that

$$x \times y \subseteq \mathcal{P}\mathcal{P}(x \cup y),$$

so that we can use SEPARATION to show that $x \times y$ is always a set:

$$x \times y = \{u \in \mathcal{P}\mathcal{P}(x \cup y) \mid \exists v \exists w (u = (v, w) \wedge v \in x \wedge w \in y\}. \tag{3}$$

EXERCISE 2.10. Explain why one needs to employ $\subseteq \mathcal{P}\mathcal{P}(x \cup y)$ in (3) to apply SEPARATION.

The definition of product can be generalized in a straighforward way:

$$x^n := x_1 \times \ldots \times x_n := (x_1 \times \ldots \times x_{n-1}) \times x_n.$$

DEFINITION 2. *An n-*ary relation *is a set of tuples $(u_1, \ldots, u_n)$, and as such is a subset of some $x_1 \times \ldots \times x_n$. An n-*ary relation *R^n on a set x is such that $R^n \subseteq x^n$.*

The *domain* of a binary relation R is the set

$$\mathrm{dom}(R) = \{u \mid (u, v) \in R\},$$

and its *range* the set

$$\mathrm{ran}(R) = \{v \mid (u, v) \in R\}.$$

EXERCISE 2.11. Show that $\mathrm{dom}(R) \subseteq \bigcup\bigcup R$ and that $\mathrm{ran}(R) \subseteq \bigcup\bigcup R$.

As a consequence, $\mathrm{dom}(R)$ and $\mathrm{ran}(R)$ are sets by SEPARATION and POWER SET (why?).

A binary relation $R \subseteq x^2$ is:

reflexive, if $(\forall u \in x)\, Ruu$,

irreflexive,	if $(\forall u \in x)\, \neg Ruu$,
symmetric,	if $(\forall u, v \in x)(Ruv \longrightarrow Rvu)$,
antisymmetric,	if $(\forall u, v \in x)((Ruv \wedge Rvu) \longrightarrow u = v)$,
connected,	if $(\forall u, v \in x)(u = v \vee Ruv \vee Rvu)$,
transitive,	if $(\forall u, v, w \in x)(Ruv \wedge Rvw \longrightarrow Ruw)$.

A binary relation is an *equivalence relation* if it is reflexive, symmetric, and transitive. For $\equiv$ an equivalence relation on x, and for any $u \in x$,

$$[u] = \{v \in x \mid u \equiv v\}$$

is called the *equivalence class* of u.

DEFINITION 3. *A binary relation f is* a function *if*

$$(x, y) \in f \wedge (x, z) \in f \longrightarrow y = z.$$

We introduce some notation for functions:

- f is a function *on* x if $x = \mathrm{dom}(f)$.
- f is an n-ary function on x if $\mathrm{dom}(f) = x^n$.
- f is a function *from* x to y (written: $f : x \longrightarrow y$) if $\mathrm{dom}(f) = x$ and $\mathrm{ran}(f) \subseteq y$.
- the set of *all functions* from x to y is written y^x.

  EXERCISE 2.12. Considering that $y^x \subseteq \mathcal{P}(x \times y)$, show that y^x is a set.

- f is a function *onto* y, if $\mathrm{ran}(f) = y$.
- f is *one-to-one* if $f(x) = f(y) \longrightarrow x = y$.
- the *restriction* of f to x, written $f \restriction x$, is defined as:

$$f \restriction x = \{(u, v) \in f \mid u \in x\}.$$

- for f and g functions and $\mathrm{ran}(g) \subseteq \mathrm{dom}(f)$, the *composition* of f and g, written $f \circ g$, is the function with $\mathrm{dom}(f \circ g) = \mathrm{dom}(g)$ such that

$$\text{for all } x \in \mathrm{dom}(g),\ f \circ g(x) = f(g(x)).$$

- the *image* of a set x under f is

$$f''(x) = \{v \mid (\exists u \in x)\, v = f(u)\},$$

  and the *inverse image* of x is

$$f_{-1}(x) = \{u : f(u) \in x\}$$

- for f a one-to-one function, the the *inverse* of f, written f^{-1}, is such that

$$f^{-1}(x) = y \text{ if and only if } f(y) = x.$$

A *family* F of sets – i.e. a collection of sets – is *disjoint* iff, for all $x, y \in F$, $x \cap y = \emptyset$. A *partition* of a set x is a disjoint family P of nonempty sets such that

$$x = \bigcup \{u \mid u \in P\}.$$

Exercise 2.13. For $\equiv$ an equivalence relation on x, define the *quotient of x by $\equiv$* as

$$x/\equiv := \{[y] \mid y \in x\}.$$

Show that $x/\equiv$ is a partition of x.

Finally, consider a partition P of x. It can be shown that it defines an equivalence relation, by letting:

$$u \equiv v \text{ if and only if } (\exists y \in P)(u \in y \wedge v \in y)$$

(exercise: verify this last claim, that is verify that it indeed satisfies the properties of an equivalence relation).

7. Other axioms

There are four additional axioms of ZFC that we will consider in detail in the following weeks, but that we will only mention for completeness.

The first is the main set existence axiom of ZFC – and the one that enables us to dispense with the empty set axiom: it states that there is an infinite set. Call a set x *inductive* if

$$\emptyset \in x \wedge (\forall u \in x)(u \cup \{u\} \in x).$$

Then:

(INFINITY) There is an inductive set.

The remaining two axioms are:

(REPLACEMENT)

$$\forall x \forall u \forall v(\varphi(x, u, w) \wedge \varphi(x, v, w) \longrightarrow u = v) \longrightarrow \forall x \exists y \forall u(u \in y \leftrightarrow (\exists v \in x)\varphi(v, u, w))$$

'If the class F defined by φ is a function and $\mathrm{dom}(F)$ is a set, then $\mathrm{ran}(F)$ is a set'.

(CHOICE) For every family of nonempty sets there is a choice function.

(FOUNDATION) Every non-empty set has an $\in$-minimal element.

DEFINITION 4. ZFC *is the first-order theory in* $\mathcal{L}_\in$ *whose axioms are* EXTENSIONALITY, PAIRING, UNION, SEPARATION, POWER-SET, INFINITY, REPLACEMENT, CHOICE, FOUNDATION.

CHAPTER 3

Week 3
Ordinal and Cardinal Numbers

We anticipated in previous chapters that one of the fascinating tools that set theory offered was the possibility of generalizing the natural numbers to *infinite numbers*. In this chapter we explore such generalizations.

1. Ordinals

1.1. Partial and Linear Orderings.

DEFINITION 5. *Let P be a set and $<$ a binary relation on P.*

(i) $<$ *is a* strict partial ordering *of P iff it is:*
- irreflexive*:* $(\forall x \in P)\, x \not< x$*;*
- transitive*:* $(\forall x, y, z \in P)(x < y \wedge y < z \longrightarrow x < z)$.

$<$ *is a* strict linear ordering *of P if is also connected:* $(\forall x, y \in P)(x < y \vee y < x \vee x = y)$.

(ii) *A* partial ordering *of P is a binary relation on P which is transitive, reflexive and antisymmetric. A* linear *ordering, is, in addition, connected.*

EXAMPLE 1.

(i) Given any set a, the relation $\subseteq$ of inclusion on $\mathcal{P}a$ is a partial ordering.

(ii) The usual relations $<$ and $\leq$ on the natural numbers $\mathbb{N}$ are, respectively, a strict linear ordering and a linear ordering of $\mathbb{N}$.

EXERCISE 3.1. Given any set a, what ordering of $\mathcal{P}a$ is given by *proper inclusion*?

Some elements in partial orderings possess a special status. Given a *partially ordered set* $(P, \leq)$, and some $x \subseteq P$,

- a *maximal element* of x is some $a \in x$ such that $(\forall u \in x)\, a \not\leq u$;
- a *minimal element* of x is some $a \in x$ such that $(\forall u \in x)\, u \not\leq a$;

- the *greatest element* of x is some $a \in x$ such that $(\forall u \in x)\ u \leq a$;
- the *least element* of x is some $a \in x$ such that $(\forall u \in x)\ a \leq u$;
- $a \in P$ is an *upper bound* of x if $(\forall u \in x)\ u \leq a$;
- $a \in P$ is a *lower bound* of x if $(\forall u \in x)\ a \leq u$;
- the *supremum* of x, denoted with $\sup(x)$, is the least upper bound of x: that is an upper bound a of x such that, for any other upper bound b of x, $a \leq b$;
- the *infimum* of x, denoted with $\inf(x)$, is the greatest lower bound of x: that is a lower bound a of x such that, for any other lower bound b of x, $b \leq a$.

In what follows, we will be sloppy about the distinction between partial and strict partial orders, as this will not matter much for later developments.

EXERCISE 3.2. Consider the set $A = \{a, b\}$, with $a \neq b$. Find maximal, minimal, greatest, least, supremum, infimum of $(\mathcal{P}(A), \subseteq)$.

EXERCISE 3.3. Let $(P, \leq)$ be a linear order. Then show that the minimal element of P is unique and that coincides with its least element.

We now turn to comparing partially ordered sets.

DEFINITION 6. *Let $(P, <)$ and $(Q, <)$ be partially ordered sets, and $f : P \longrightarrow Q$. Then:*

(i) *f is* order-preserving*, if for all $x, y \in P$, if $x < y$, then $f(x) < f(y)$;*

(ii) *if in addition $(P, <)$ and $(Q, <)$ are linear orders, then f is called* increasing*;*

(iii) *if f is one-to-one, and both f and f^{-1} are order-preserving, f is called an* isomorphism of P and Q *− and therefore, $(P, <)$ and $(Q, <)$ are* isomorphic*. An isomorphism $f : P \longrightarrow P$ is called an* automorphism.

1.2. Well-orderings. One can see the previous definitions as preparatory work to a fundamental concept of set theory, the one of a *well-ordering*. It is important because it is a fundamental tool to compare sets by their 'length'. Such lengths will be called *order-type*. *Ordinal numbers*, in particular, will simply be such order-types. Let's see precisely how this all works.

DEFINITION 7. *Let $(P, <)$ is a linearly ordered set. $<$ is called a* well-ordering *if every non-empty subset of P has an $<$-least element. In symbols:*

$$\forall x(x \subseteq P \wedge x \neq \emptyset \longrightarrow (\exists y \in x)(\forall u \in x)y \leq u).$$

Given a well-ordered set $(W, <)$, we call *the initial segment of W given by $x \in W$* the set $\{u \in W \mid u < x\}$. We turn now to the main result about well-ordered set of this subsection, Theorem 1 below. This is also the first real proof that we consider.

THEOREM 1. *If W_1 and W_2 are well-ordered sets, then exactly one of the following claims hold:*

(i) *W_1 and W_2 are isomorphic;*

(ii) *W_1 is isomorphic to an initial part of W_2;*

(iii) *W_2 is isomorphic to an initial part of W_1.*

PROOF. We break down the proof in several claims. The first is:

(4) No well-ordered set is isomorphic to an initial segment of itself.

To establish (4), consider a well-ordered set $(W, <)$.

We first notice that, if $f : W \longrightarrow W$ is increasing, then $f(x) \geq x$ for any $x \in W$. Suppose in fact there are indeed elements $u \in W$ such that $f(u) < u$. This means that the subset of W, $\{u \in W \mid f(u) < u\}$ is nonempty. By the definition of well-ordering, it has an $<$-least element. Let's call it u_0. Now, since $f(u_0) < u_0$ and f is increasing, also $f(f(u_0)) < f(u_0)$. However, again since $f(u_0) < u_0$ and u_0 is the least member of $\{u \in W \mid f(u) < u\}$, also $f(u_0) \leq f(f(u_0))$. This is a contradiction. So $\{u \in W \mid f(u) < u\} = \emptyset$.

We can then establish (4): let $x \in W$ and consider $Wx := \{u \in W \mid u < x\}$. Suppose there is an isomorphism of W and Wx. Then $f(x) \in Wx$, and $f(x) < x$, which is a contradiction by what we have just shown.

(4) immediately entails that (i), (ii), (iii) are mutually exclusive.

EXERCISE 3.4. Show that (4) entails that (i), (ii), (iii) in Theorem 1 are mutually exclusive.

Now that the preliminary work is over, we can complete the proof of Theorem 1. Let's define a function $f : W \longrightarrow W$ as

$$f := \{(u, v) \in W_1 \times W_2 \mid W_1 u \text{ and } W_2 v \text{ are isomorphic}\}.$$

The first thing to notice about f so-defined is that it is one-to-one: suppose in fact that $f(u_0) = f(u_1)$, for $u_0, u_1 \in W_1$. Suppose further that $u_0 \neq u_1$.

Since $<$ is a well-ordering, either $u_0 < u_1$ or $u_1 < u_0$. We can safely assume that $u_0 < u_1$ (the argument for $u_1 < u_0$ is symmetric). By definition of f, $W_1 u_0$ is isomorphic to $W_2 f(u_0)$, which by assumption is isomorphic to $W_2 f(u_1)$, which again is isomorphic to $W_1 u_1$. Therefore, $W_1 u_0$ is isomorphic to $W_1 u_1$, which contradicts (4).

A further feature of f is that it is order-preserving. Suppose in fact that, for $u_0, u_1 \in W_1$, $u_0 < u_1$, but $f(u_0) > f(u_1)$. By definition of f, $W_1 u_1$ is isomorphic to $W_2 f(u_1)$. Call this isomorphism g: then $W_1 u_0$ and $W_1 g(u_0)$ are isomorphic, and since $g(u_0) < f(u_1)$, also $f(u_0) < f(u_1)$.

The proof can now be completed as follows by distinguishing several, mutually exclusive cases.

If $\mathrm{dom}(f) = W_1$ and $\mathrm{ran}(f) = W_2$, then W_1 and W_2 are isomorphic – because each single initial segment of W_1 is isomorphic to a unique initial segment of W_2 – and the proof is complete.

If $\mathrm{ran}(f) \neq W_2$, then $W_2 \setminus \mathrm{ran}(f)$ is nonempty. Let v_0 be the $<$-minimal element of $W_2 \setminus \mathrm{ran}(f)$. This entails that $\mathrm{dom}(f) = W_1$, because if $W_1 \setminus \mathrm{dom}(f) \neq \emptyset$ and u_0 is the $<$-minimal of its members, then one would have that $W_1 u_0$ and $W_1 v_0$ are isomorphic, and therefore $u_0 \in \mathrm{dom}(f)$ and $v_0 \in \mathrm{ran}(f)$ after all. Therefore, W_1 and $W_2 v_0$, an initial segment of W_2, are isomorphic and case (ii) is realized.

If $W_1 \setminus \mathrm{dom}(f) \neq \emptyset$, one can reason as in the previous case to obtain case (iii).

qed

EXERCISE 3.5. Complete the proof of Theorem (1) by showing that claim (iii) is forced by the case in which $W_1 \setminus \mathrm{dom}(f) \neq \emptyset$.

When we compare well-ordered sets by means of isomorphism, we say that we are comparing their *order-type*: two well-ordered sets $(W_1, <)$ and $(W_2, <)$ have the same order-type if they are isomorphic. *Ordinal numbers* are just such order types. We will, however, present a formal definition which is equivalent, although slightly different.

1.3. Ordinal numbers. Ordinal numbers are defined as a special subclass of *transitive sets*:

DEFINITION 8. *A set is* transitive *if every element of it is also a subset of it:*

$$\mathrm{Trans}(x) \ :\leftrightarrow\ (\forall u \in x)\, u \subseteq x$$

EXERCISE 3.6. Show that $\mathrm{Trans}(x) \leftrightarrow \bigcup x \subseteq x$.

DEFINITION 9. *x is an* ordinal *iff it is transitive and well-ordered by $\in$.*

Equivalently (assuming the axiom of foundation/regularity), one can say that x is an ordinal number if it is transitive and all its members are transitive.

The elements of the class Ord of ordinals are usually denoted with Greek letters $\alpha, \beta, \gamma, \ldots$. We let:

$$\alpha < \beta :\leftrightarrow \alpha \in \beta.$$

LEMMA 3. *If $x \neq \emptyset$ is a set of ordinals, then $\bigcup x$ is an ordinal, and $\bigcup x = \sup(x)$.*

PROOF. $\bigcup x = \{u \mid (\exists y \in x)\, u \in y\}$. Therefore, since any element of an ordinal is itself an ordinal (why?), $\bigcup x$ is a set of ordinals. Therefore, for any $\alpha, \beta \in \bigcup x$, $\alpha < \beta$ or $\beta < \alpha$ or $\alpha = \beta$. So $\bigcup x$ is well-ordered by $<$. To show that $\bigcup x$ is itself an ordinal, it suffices to show that $\bigcup x$ is transitive. Now suppose that $a \in \bigcup x$. Then for some $\alpha \in x$, $a \in \alpha$. But α is an ordinal, so $a \subseteq \alpha$. Therefore $a \subseteq \bigcup x$, as required.

It remains to be shown that $\bigcup x = \sup(x)$. That is, we need to show that $(\forall u \in x)\, u \leq \bigcup x$, and that any other upper bound b of x is such that $\bigcup x \leq b$. For the former, if u is some ordinal $\beta \in x$, then $\beta \subseteq \bigcup x$: since $\bigcup x$ is an ordinal, $\beta = \bigcup x$ or $\beta \in \bigcup x$.[1] Finally, if b is an upper bound of x and $b < \bigcup x$, then b is also an ordinal $\in \bigcup x$, and therefore there is an ordinal $u \in x$ such that $b < u$. So b is not an upper bound after all. *qed*

Given an ordinal $\alpha = \{\beta \mid \beta < \alpha\}$, the ordinal $\alpha \cup \{\alpha\}$ (written $\alpha + 1$), is called *the successor of α*. It is an ordinal (proof as an exercise). An ordinal of the form $\beta + 1$, for β an ordinal, is called a *successor ordinal.*

PROPOSITION 1 (BURALI-FORTI). *The class* Ord *is a proper class.*

PROOF. Suppose that Ord is a set. Then sup(Ord) is a set as well by Lemma 3. But then sup(Ord) $\cup$ {sup(Ord)} is an ordinal, therefore it is an element, and therefore a subset, of Ord. So Ord $\in$ Ord, which contradicts the definition of an ordinal, because no ordinal is self-membered.[2] *qed*

The next theorem gives substance to the idea – envisaged earlier in the notes – that ordinals are the 'lengths' of well-ordered sets.

[1] Notice that here we are employing the following fact: if α and β are ordinals and $\alpha \subseteq \beta$, then $\alpha = \beta$ or $\alpha \in \beta$.

[2] More generally, as we shall see, this also contradicts the axiom of Foundation.

THEOREM 2. *Every well-ordered set is isomorphic to a unique ordinal number.*

PROOF. Let A be the set of all $a \in W$ such that Wa is isomorphic to an ordinal. By (4), this ordinal is unique, since if a well-ordered set was isormorphic to two distinct ordinals, one of these would be isomorphic to an initial segment of itself. Call α_a the unique ordinal number associated with $a \in W$. Let $B = \{\alpha_a \mid a \in A\}$. Since the mapping $a \mapsto \alpha_a$ – let's call it F – is a class function, by the axiom of replacement its range, namely B, exists. In addition it is a set of ordinals that is well-ordered by $\in$ – as all ordinals are –, and it is transitive, so there is some $\beta \in \mathrm{Ord}$ such that $\beta = B$.

It remains to show that F is an isomorphism between W and β. The definition of A entails that A and β are isomorphic. In addition, we notice that A is downwards closed under $<$: that is if $a \in A$ and $c < a$, then $c \in A$. Therefore, either $A = W$ or $A = Wd$ for some $d \in W$. If the former, the proof is complete. But the latter is impossible, because if $A = Wd$, then $d \in \{x \in W \mid x < d\}$. This completes the proof. *qed*

An ordinal that is not a successor has the form

$$\lambda = \sup\{\xi \mid \xi < \lambda\} = \bigcup \lambda.$$

Therefore, $\emptyset$ is a limit ordinal.

DEFINITION 10 (NATURAL NUMBERS). *The least limit ordinal different from $\emptyset$ is called ω (a.k.a. $\mathbb{N}$). We let:*

$$0 := \emptyset \qquad 1 := 0 + 1 = \emptyset \cup \{\emptyset\} = \{\emptyset\}$$
$$2 := 1 + 1 = \{\emptyset\} \cup \{\{\emptyset\}\} \qquad 3 := 2 + 1$$
$$\vdots \qquad \vdots$$

DEFINITION 11 (FINITE, INFINITE). *A set x is* finite *if there is a one-to-one function from x onto some $n \in \omega$. A set is* infinite *if it is not finite.*

PROPOSITION 2. *ω exists and it is the least inductive set.*

PROOF. The axiom of infinity would give us the existence claim immediately, if we could show that ω is inductive. But this is immediate: certainly $\emptyset \in \omega$, and if $n \in \omega$ then $n + 1 \in \omega$. So ω exists.

It remains to be shown that if S is inductive, and then $\omega \subseteq S$. Suppose not, then $\omega \setminus S \neq \emptyset$. Pick the least $n \in \omega$ such that $n \notin S$. But then n is either $\emptyset \in S$, or is a successor ordinal, which is also in S.

qed

THEOREM 3 (TRANSFINITE INDUCTION). *Let $\varphi(u)$ specifies a class of ordinals satisfying:*

(i) $\varphi(0)$

(ii) *for all α: if $\varphi(\alpha)$, then $\varphi(\alpha + 1)$;*

(iii) *if $\lambda \neq \emptyset$ is a limit ordinal and $\varphi(\beta)$ for all $\beta < \lambda$, then $\varphi(\lambda)$.*

Then φ holds of all $\alpha \in$ Ord.

PROOF. Suppose $\text{Ord} \setminus \{\alpha \mid \varphi(\alpha)\} \neq \emptyset$, and let α be the least such ordinal. Then, since an ordinal can only be $\emptyset$, a successor, or a limit, the three assumptions all lead to a contradiction. *qed*

COROLLARY 2 (TRANSFINITE INDUCTION (SECOND FORM)). *Let φ define a class. If $(\forall \beta < \alpha)\varphi(\beta)$ entails $\varphi(\alpha)$, then $\varphi(\alpha)$ for all $\alpha \in$ Ord.*

PROOF. Exercise. *qed*

A function whose domain is some ordinal α is called a *transfinite sequence* and denoted with:

$$\langle a_\xi \mid \xi < \alpha \rangle$$

A *sequence* is simply a function whose domain is ω, whereas a *finite sequence* a function whose domain is some $n \in \omega$.

A fundamental tool to construct functions in ZFC is the following.

THEOREM 4 (TRANSFINITE RECURSION). *Let $G : \mathrm{V} \longrightarrow \mathrm{V}$ be a function. Then, there is a unique function F with $\mathrm{dom}(F) = \text{Ord}$ such that:*

$$F(\alpha) = G(F \restriction \alpha).$$

PROOF. We want to define F as a the result of collecting together a series of approximating functions, that will take the form of specific sequences. In particular, for an ordinal α, we want

$$(\star) \qquad F(\alpha) = x \;:\leftrightarrow\; \text{There is a sequence } \langle a_\xi \mid \xi < \alpha \rangle \text{ such that}$$
$$a_\xi = G(\langle a_\eta \mid \eta < \xi \rangle) \text{ for all } \xi < \alpha, \text{ and}$$
$$x = G(\langle a_\xi \mid \xi < \alpha \rangle).$$

Now we verify that, for each α, $F(\alpha)$ *exists and is unique*. Both claims make use of the transfinite induction principle. For the induction hypothesis, let's assume that $F(\beta)$ exists and is unique for any $\beta < \alpha$.

Existence:

Consider two sequences

$$\langle a_\xi \mid \xi < \alpha \rangle \quad \text{and} \quad \langle b_\xi \mid \xi < \alpha \rangle$$

satisfying ($\star$). Now, since $G(\emptyset)$ is uniformly determined, and so are successor and limit stages in terms of G, by transfinite induction on ξ they must be the same. So, if $F(\alpha)$ exists, it is a function.

Again by transfinite induction on α, one shows that: for every α there is an α-sequence satisfying ($\star$). This is trivial for $\emptyset$, since $F(\emptyset)$ is simply $G(\emptyset)$, which is defined by assumption. The successor stage is readily obtained. For the limit stages $\alpha := \lambda$, we are not guaranteed that the sequence

$$\bigcup_{\beta<\lambda} \langle a_\xi \mid \xi < \beta \rangle$$

exists, so one needs to employ the axiom schema of replacement to obtain the existence of such limits. But this is no problem. Therefore,

$$F(\alpha) = G(\langle a_\xi \mid \xi < \alpha \rangle) = G(F \restriction \alpha).$$

Uniqueness: Suppose that there is an $H : \mathrm{Ord} \longrightarrow V$ such that $H(\alpha) = G(H \restriction \alpha)$. Again, since $F(\emptyset) = H(\emptyset)$, and if successor steps and limit stages are uniquely determined in terms of previous steps and G, one concludes by transfinite induction that $F = G$. *qed*

Transfinite recursion enables us to define the generalizations of arithmetical operations to ordinals.

DEFINITION 12 (ORDINAL ARITHMETIC). *Let λ be a nonzero limit:*

(i) *Addition:*

$$\begin{aligned} \alpha + 0 &= \alpha \\ \alpha + (\beta + 1) &= (\alpha + \beta) + 1 \\ \alpha + \lambda &= \sup\{\alpha + \beta \mid \beta < \lambda\} \end{aligned}$$

(ii) *Multiplication:*

$$\begin{aligned} \alpha \cdot 0 &= 0 \\ \alpha \cdot (\beta + 1) &= (\alpha \cdot \beta) + \alpha \\ \alpha \cdot \lambda &= \sup\{\alpha \cdot \beta \mid \beta < \lambda\} \end{aligned}$$

(iii) *Exponentiation:*

$$\begin{aligned} \alpha^0 &= 1 \\ \alpha^{\beta+1} &= (\alpha^\beta) \cdot \alpha \\ \alpha^\lambda &= \sup\{\alpha^\beta \mid \beta < \lambda\} \end{aligned}$$

The most basic and salient properties of ordinals are described by the following lemma.

LEMMA 4.

(i) $\cdot$ *and* $+$ *are associative;*

(ii) $\cdot$ *and* $+$ *are* not *commutative.*

2. Cardinals

Like ordinals, cardinal numbers can also be seen as 'measures' of certain sets. However, we are now interested in the pure *size* of sets, regardless of their structure given by some ordering relations.

DEFINITION 13 (CARDINALITY). *Two sets x and y have the same* cardinality *– in symbols, $|x| = |y|$ –, if there is a one-to-one mapping of x onto y.*

A set x is *finite* if, for some $n \in \omega$, $|x| = |n|$. In this case one says that x has n elements.

EXERCISE 3.7. For $n, m \in \omega$, show that $|n| = |m|$ iff $n = m$.

Therefore, finite cardinals are just the natural numbers as defined in ZFC. More generally:

DEFINITION 14 (CARDINAL NUMBER). *An ordinal α is called a cardinal number if for all $\beta < \alpha$, $|\alpha| \neq |\beta|$.*

EXAMPLE 2. For instance, $\omega < \omega + \omega$, but $|\omega| = |\omega + \omega|$, because there is a one-to-one mapping between ω and $\omega + \omega$ (Think of the one-to-one mapping that there is between $\mathbb{N}$ and the result of listing first the even numbers, and then the odd numbers).

Cardinals are denoted with the first letter of the Hebrew alphabet $\aleph$. The ordinal ω is the least infinite cardinal, and is denoted with $\aleph_0$. Sets with cardinality $\aleph_0$ are called *countable*. Infinite sets that are not countable are called *uncountable*.

Sets can be ordered by cardinality:

$$|x| \leq |y| :\leftrightarrow \text{There is a one-to-one mapping of } x \text{ into } y.$$

As usual, $|x| < |y|$ means that $|x| \leq |y|$ and $|x| \neq |y|$. This suffices to introduce one of the most well-known theorems of set theory.

THEOREM 5 (CANTOR). *For every set x, $|x| < |\mathcal{P}x|$.*

Proof. One first shows that $|x| \leq |\mathcal{P}x|$. For this, it suffices to find a one-to-one function of x onto $\mathcal{P}x$. But the function $x \mapsto \{x\}$ is precisely this function.

It remains to be shown that $|x| \neq |\mathcal{P}x|$. Suppose that $|x| = |\mathcal{P}x|$, and consider a one-to-one function f from x *onto* $\mathcal{P}x$. Recall that this means that $\mathrm{ran}(f) = \mathcal{P}(X)$. However, consider the set $y \subseteq x$ defined as:

$$y = \{u \in x \mid u \notin f(u)\}.$$

If $y \in \mathrm{ran}(f)$, then there is some $u \in x$ such that $y = f(u)$. Now, if $u \in y$, then $u \notin f(u)$, so $u \notin y$. Therefore, $u \notin y$. But then $u \in x \wedge u \notin f(u)$, so $u \in y$. Contradiction. *qed*

Cardinal arithmetic has slight different properties than ordinal arithmetic, and can be directly defined:

$$\begin{aligned} |x| + |y| &:= |x \cup y| && \text{for disjoint } x, y, \\ |x| \cdot |y| &:= |x \times y| \\ |x|^{|y|} &:= |x^y|. \end{aligned}$$

Exercise 3.8. Show that + and × are associative and commutative.

The definitions of the basic arithmetical operations on cardinals enable us to formulate Cantor's theorem in a more familiar form.

Corollary 3. *For any set x, $|\mathcal{P}x| = 2^{|x|}$.*

Proof. The *characteristic function* of a set $u \subseteq x$, called χ_u, is such that

$$\chi_u(v) = \begin{cases} 1, & \text{if } v \in u \\ 0, & \text{if } v \in x \setminus u. \end{cases}$$

The mapping $u \mapsto \chi_u$ is a one-to-one mapping of $\mathcal{P}x$ onto $\{0,1\}^x$, the set of functions from x to $\{0,1\}$. *qed*

For yet another formulation:

Corollary 4. *For every cardinal κ, $\kappa < 2^\kappa$.*

CHAPTER 4

Week 4
Real Numbers and The Axiom of Choice

1. Reals

In this section we outline the set theoretic construction of some of the most well-known mathematical objects. We have already constructed in the previous chapters the set $\mathbb{N}$ of natural numbers. We now see how to construct in ZFC the integers $\mathbb{Z}$, the rationals $\mathbb{Q}$, and the reals $\mathbb{R}$.

If one pictures $\mathbb{N}$ as a list, it's clear that the integers can be pictured as extending the natural numbers with another copy of $\{1, 2, 3, \ldots\}$ in the 'reversed direction':

$$\ldots -3, -2, -1, 0, 1, 2, 3, \ldots$$

This is nothing more than a picture, but it gives rise to the idea of treating the integers as pairs of natural numbers that result in a unique location in this list. Intuitively, we would like to define integers as pairs of natural numbers (n, m) such that, for all $(n, m) = (k, l)$ iff $n - m = k - l$. So $(6, 9)$ and $(12, 15)$ are the same integer because $6 - 9 = 12 - 15$. Of course, since subtraction is not defined (yet) for negative integers – because they are yet to be defined themselves –, we need to find a shortcut, which is based on the fact that $6 - 9 = 12 - 15$ iff $6 + 15 = 9 + 12$.

DEFINITION 15 (THE SET $\mathbb{Z}$ OF INTEGERS). *Consider the set $\mathbb{N}^2$, and define the binary relation on it*

$$(k, l) \equiv (m, n) \text{ if and only if } k + n = l + m.$$

We let $\mathbb{Z} := \mathbb{N}^2/\equiv$.

Addition and multiplication for integers can be defined inductively as:

$$[(a, b) + (c, d)] = [(a_c, b_d)],$$
$$[(a, b)] \cdot [(c, d)] = [((a \cdot c) + (b \cdot d)), ((a \cdot d) + (b \cdot c))].$$

An integer a is divisible by an integer b if there is a unique $x \in \mathbb{Z}$ such that $a = x \cdot b$. Divisibility is key to define rational numbers, intuitively, as *fractions* of integers where we avoid division by 0.

DEFINITION 16 (THE SET $\mathbb{Q}$ OF RATIONAL NUMBERS). *Let* $Q := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) = \{(p, q) \in \mathbb{Z}^2 \mid q \neq 0\}$. *We define the binary relation on* Q*:*

$$\frac{a}{b} \sim \frac{c}{d} \text{ if and only if } a \cdot d = b \cdot c.$$

Finally we let $\mathbb{Q} := Q/\sim$.

We are not focusing so much on the operations (the so-called field properties) of $\mathbb{Q}$, but only on the order properties of it. The usual ordering of $\mathbb{Q}$ is defined, for $b, d > 0$, as:

$$[a/b] < [c/d] \text{ iff } a \cdot d < b \cdot c.$$

We say that a linear ordering $(P, <)$ is *bounded from above* if it has an upper bound, and that it is *bounded from below* if it has a lower bound. $(P, <)$ is *bounded* if it has both upper and lower bounds.

DEFINITION 17 (DENSE, COMPLETE LINEAR ORDERINGS).

(i) *A linear ordering* $(P, <)$ *is* dense *if for all* $p, q \in P$*, if* $p < q$ *then there is a* r *such that* $p < r < q$.

(ii) *A linear ordering* $(P, <)$ *is* complete *if every nonempty* $x \subseteq P$ *bounded from above has a supremum (i.e. if* x *has an upper bound, it has a least upper bound).*

EXAMPLE 3. $\mathbb{Q}$ is dense, and does not have least or greatest elements, but not complete. Consider $\{x \in \mathbb{Q} \mid x < \sqrt{2}\}$. It has an upper bound in $\mathbb{Q}$ (e.g. $3/4$), but no least upper bound in $\mathbb{Q}$, because we can always choose a smaller upper bound. So $(\mathbb{Q}, <)$ is said to have 'gaps'.

DEFINITION 18 (CUT, DEDEKIND CUT). *Let* $(P, <)$ *be a dense, linearly ordered set without least or greatest elements:*

(i) *A* cut *in* P *is a pair* (A, B) *of sets such that:*
   (a) *A and B are nonempty and disjoint subsets of P, and $A \cup B = P$.*
   (b) *if $u \in A$ and $v \in B$, then $u < v$.*

(ii) *A cut* (A, B) *is, in addition, a Dedekind cut if* A *does not have a greatest element.*

The following important theorem is essential to construct real numbers from rationals. It basically constructs complete, dense linear orderings from non-complete ones by 'filling the gaps'.

THEOREM 6. *Let $(P, <)$ be a dense, linearly ordered set without least or greatest elements. Then there is a complete linearly ordered set $(C, \prec)$ such that:*

(1) $P \subseteq C$;
(2) $<$ and $\prec$ coincide on P: for all $p, q \in P$, $p < q$ iff $p \prec q$;
(3) P is dense in C: for all $c, d \in C$, there is a $p \in P$ such that $c \prec p \prec d$;
(4) C does not have greatest and least elements.

In addition, $(C, \prec)$ is unique, modulo isomorphism, over P: for any $(C^*, \prec^*)$ satisfying the conditions (1)-(4) above there is an isomorphism h between $(C, \prec)$ and $(C^*, \prec^*)$ such that, for all $c \in P$, $h(c) = c$. In virtue of this uniqueness property $(C, \prec)$ is called the completion of $(P, <)$.

PROOF. Given its statement, it's clear that the proof of the theorem splits between an existence and a uniqueness condition.

Uniqueness Condition. Suppose that there are complete linear orderings $(C, \preceq)$ and $(C^*, \preceq^*)$ satisfying 1-4. One has to show that there is an isomorphism h between them such that $h(x) = x$ for all $x \in P$.

For $c \in C$, let $S_c := \{p \in P \mid p \preceq c\}$, and for $c^* \in C^*$, $S_{c^*} := \{p \in P \mid p \preceq^* c^*\}$. For nonempty $S \subseteq P$, $\sup S$ is the least upper bound of S in $(C, \preceq)$, and $\sup^* S$ is the least upper bound of S in $(C^*, \preceq^*)$. Therefore, we have:

$$\sup S_c = \sup\{p \in P \mid p \preceq c\} = c,$$

and

$$\sup^* S_c = \sup\{p \in P \mid p \preceq^* c\} = c^*.$$

Define $h : C \longrightarrow C^*$ as $h(c) = \sup^* S_c = \sup^*\{p \in P \mid p \preceq^* c\}$. Since C is a linear order and the upper bound is uniquely determined, h is a function from C *into* C^*. To show also that it is *onto*, one needs to show that $\mathrm{ran}(h) = C^*$. Pick an arbitrary $c^* \in C^*$, so $c^* = \sup^* S_{c^*}$. Then, since $\preceq^*$ and $\preceq$ both agree with $\leq$ on P, we can let

$$c = \sup S_{c^*} = \sup\{p \in P \mid p \preceq c^*\}.$$

Therefore,

$$\begin{aligned} S_c = \{p \in P \mid p \preceq c\} &= \{p \in P \mid p \preceq \sup S_{c^*}\} \\ &= \{p \in P \mid p \preceq^* \sup S_{c^*}\} = S_{c^*} \end{aligned}$$

The last equality holds because of (2). Therefore $c^* = \sup^* S_{c^*} = \sup^* S_c$.

Next, we show that h is order preserving.[1] Suppose that, for $c, d \in C$, $c \prec d$. By assumption (3), there is a $p \in P$ such that $c \prec p \prec d$, and by assumption (2), $\sup^* S_c \prec p \prec \sup^* S_d$, which entails that $h(c) \preceq^* h(d)$.

Finally, for $x \in P$, we have:

$$x = \sup\{u \in P \mid u < x\} = \sup\{u \in P \mid u \prec^* x\} = h(x).$$

Existence Condition. We let $(C, \preceq)$ be the set of all Dedekind cuts from P and the order $\preceq$ to be defined as

$$(A_0, B_0) \preceq (A_1, B_1) \text{ iff } A_0 \subseteq A_1 \text{ (and therefore } B_1 \subseteq B_0).$$

Exercise 4.1. $(C, \preceq)$ is a linearly ordered set.

The strategy now proceeds as follows: by letting

$$A_p = \{x \in P \mid x < p\} \qquad B_p = \{x \in P \mid x \geq p\}$$

it is clear that $(P' := \{(A_p, B_p) \mid p \in P\}, \preceq)$ is isomorphic to $(P, \leq)$. It then suffices to show that $(C, \preceq)$ is a completion of $(P', \preceq)$. That is, that it satisfies the clauses (1)-(4) above.

Obviously, $P' \subseteq C$, and also (2) is trivially satisfied. It remains to show (3), (4), and the completeness of C. To show that P' is dense in C, we assume that $c, d \in C$ – with $c = (A_0, B_0)$, $d = (A_1, B_1)$ – such that $c \prec d$. By the definition of $\prec$, this means that $A_0 \subset A_1$. By selecting a $p \in P$ such that $p \in A_1 \setminus A_0$, we have that

$$(A_0, B_0) \prec (A_p, B_p) \prec (A_1, B_1)$$

which shows that P' is dense in C.

Next, we show that C does not have least and greatest elements. Let $(A, B) \in C$: since (A, B) is a Dedekind cut, B may or may not have a least element. In each case – since P is unbounded – we can choose an element $p \in B$ that is not the least element of B. Therefore $A \subset A_p$, and therefore C does not have a greatest element. Similarly, since A does not have a greatest element, we can find A_p (with $p \in P$) such that $A_p \subset A$, and so C does not have a least element.

It remains to be shown that C is complete, that is that every nonempty $S \subseteq C$ that has an upper bound has a supremum. Pick an arbitrary such S: by assumption, we have that there is some $(A_0, B_0) \in C$ such that $A \subseteq A_0$

[1] This will also entail that h^{-1} is order preserving, since both sets are linearly ordered.

for each $(A, B) \in S$. Let

$$A_S = \bigcup\{A \mid (A, B) \in S\} \qquad B_S = \bigcap\{B \mid (A, B) \in S\}$$

Exercise 4.2. (A_S, B_S) is a Dedekind cut.

Since $A_S \supseteq A$ for all $(A, B) \in S$, (A_S, B_S) is an upper bound of S. Let $(A', B') \in C$ be any upper bound of S, then $A' \supseteq \bigcup\{A \mid (A, B) \in S\}$, so $A_S \subseteq A'$. *qed*

We can now readily define the real numbers:

definition 19 (Reals). *The unique (up to isomorphism) completion of* $(\mathbb{Q}, <)$ *is denoted with* $(\mathbb{R}, <)$. *The elements of* $\mathbb{R}$ *are the* real numbers.

The cardinality of the real numbers is denoted with $\mathfrak{c} := \mathbb{R}$.

Proposition 3. $\mathfrak{c} = 2^{\aleph_0}$.

Proof. One first shows that $\mathfrak{c} > \aleph_0$. That is, that $\mathbb{R}$ is uncountable. Here's Cantor's original proof. Suppose that the real numbers are countable, and can be therefore enumerated by a sequence $\langle r_n \mid n \in_\omega \rangle$. Now, any r_n has a decimal expansion of the form $a_0^n, \ldots, a_n^n, \ldots$.[2] Now define a new real number whose decimal expansion $b_0, b_1, b_2, \ldots$ is:

$$b_n = \begin{cases} 1, & \text{if } a_n^n = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Let $r^* := 0.b_0 b_1 b_2, \ldots$. As a consequence, $b_n \neq a_n^n$ for each n, and therefore $r^* \neq r_n$ for each n. So r^* is a real number not in the enumeration.

We have already seen in the previous chapter that $|\mathcal{P}(\mathbb{N})| = 2^{\aleph_0}$. Notice, also, that real numbers r, being cuts of rationals, are of the form (A, B), with $A \subset \mathbb{Q}$. This mapping $r \mapsto A$ is a one-to-one function between $\mathbb{R}$ and $\mathcal{P}(\mathbb{Q})$. But since rationals are pairs of elements of $\mathbb{N}$, $|\mathbb{N}| = |\mathbb{Q}|$. So, $\mathfrak{c} \leq 2^{\aleph_0}$. Finally, each $S \subseteq \mathbb{N}$ can be represented as a real $0.a_0 a_1 a_2 \ldots$ where each a_n is the value of $\chi_S(x)$. This shows that $|\mathcal{P}(\mathbb{N})| \leq \mathfrak{c}$. This entails that $\mathfrak{c} \leq \mathfrak{c}$ and $\mathfrak{c} \geq \mathfrak{c}$, therefore $\mathfrak{c} = \mathfrak{c}$.[3]

qed

We conclude with a claim that we certainly won't be able to prove. We have seen that in ZFC infinite cardinals are *alephs*. What has just been

[2] Assuming that every real has a unique decimal expansion.

[3] Notice that this last step looks trivial but it isn't. We are employing the Cantor-Bernstein theorem.

said (together with the axiom of choice!) therefore entails that $\mathfrak{c} \geq \aleph_1$. Cantor's *Continuum Hypothesis*, one of the most well-known and longstanding conjectures in mathematics, is that

$$\mathfrak{c} = \aleph_1.$$

Kurt Gödel and Paul Cohen showed that the problem is not easy [**Goe40, Coh63**]. Both the continuum hypothesis and its negation are consistent with ZFC, and therefore unprovable from its axioms.

2. Axiom of Choice

One of the most contentious axioms of ZFC is the Axiom of Choice. Unlike other axioms, which state either give basic meaning postulates for the concept of set, or give specific set existence conditions that can then be proved to be unique by extensionality, the Axiom of Choice enables one to construct in a sense more 'arbitrary' set theoretic objects, functions in particular.

A *family* of sets S is simply a set of sets. A *choice* function for a family X is a function f on X such that $f(x) \in x$ for any $x \in X$. Intuitively, a choice function takes one element $x \in X$, and *selects* one element of x as 'representative' of x.

> AXIOM OF CHOICE: Every family of nonempty sets has a choice function.

EXAMPLE 4. Some easy examples of choice functions can be proved in ZF.

(i) Given a family of singletons, the function $f(\{x\}) = x$ is a choice function.

(ii) Given any finite family S of sets, one can construct a choice function by induction on the size of S.

Perhaps the most important consequence of the axiom of choice is the possibility of well-ordering any set. This gives us a useful tool to compare sets via their size, and fits well with our conceptual understanding of the cumulative hierarchy.

THEOREM 7 (ZERMELO'S WELL-ORDERING THEOREM). *Every set can be well-ordered.*

PROOF. Let A be a set. By the power set axiom, $\mathcal{P}(A)$ is a set. By the axiom of choice, there is a choice function h for the nonempty subsets of

A. We would like to define a function that 'lists' all elements of A once at a time. One way to do this would be to have a function:

$$\begin{aligned} f(0) =& g(A) \\ f(1) =& g(A \setminus \operatorname{ran}(f \restriction 1)) \\ =& g(A \setminus \{g(A)\}) \\ & \vdots \end{aligned}$$

But we can use the TRANSFINITE RECURSION THEOREM to define such function:

$$f(\alpha) = \begin{cases} g(A \setminus \{a_\xi \mid \xi < \alpha\}), & \text{if } A \setminus \{a_\xi \mid \xi < \alpha\} \neq \emptyset, \\ \text{some } b \notin A, & \text{otherwise.} \end{cases}$$

Let θ be the least ordinal α such that $f(\alpha) = b$. It remains to be shown that f is one-to-one, and that $A = \operatorname{ran}(f \restriction \theta)$. For the first, suppose that $\alpha \neq \beta$: we need to show that $f(\alpha) \neq f(\beta)$. If $\alpha \neq \beta$, without loss of generality we assume that $\alpha < \beta$ (because the argument for $\beta < \alpha$ is symmetric). Then, $f(\alpha) \in \operatorname{ran}(f \restriction \alpha) \subseteq \operatorname{ran}(f \restriction \beta)$. But $f(\beta) \in A \setminus \operatorname{ran}(f \restriction \beta)$. Therefore, $f(\alpha) \neq f(\beta)$, as required. Finally, by definition of f, $\operatorname{ran}(f \restriction \theta) \subseteq A$. If $\operatorname{ran}(f \restriction \beta) \subset A$, then $A \setminus \operatorname{ran}(f \restriction \beta) \neq \emptyset$, and so $f(\theta) \neq b$. A contradiction. So $A = \operatorname{ran}(f \restriction \theta)$. *qed*

Given our definition of cardinals for well-ordered set, ZERMELO'S WELL-ORDERING THEOREM entails that any two sets can be compared by their cardinality, given that every well-ordered sets can be compared via isomorphism. In particular, the set of real numbers has a cardinality. And this gives full meaning to the later remarks from the previous section, when we claimed that:

COROLLARY 5. $\mathfrak{c} \leq \aleph_1$.

Bibliography

[BM03] George Bealer and Uwe Monnich, *Property theories*, Handbook of Philosophical Logic, Volume 10 (Dov Gabbay and Frans Guenthner, eds.), Kluwer Academic Publishers, 2003, pp. 143–248.

[Coh63] Paul J. Cohen, *The independence of the continuum hypothesis*, Proceedings of the National Academy of Sciences of the United States of America **50** (1963), no. 6, 1143–8.

[Goe40] Kurt Goedel, *The consistency of the continuum hypothesis*, Princeton University Press, 1940.

[vOQ69] Willard van Orman Quine, *Ontological relativity and other essays*, Columbia University Press, 1969.