CHAPTER 2

Learning how to count and reason

The first chapter of our course starts at the reception of Hilbert's Hotel. Hilbert's Hotel is a special place, it has rooms $0, 1, 2, \ldots, n, \ldots$, one for each $n \in \mathbb{N}$. The sign outside the hotel clearly reads NO VACANCY, but you're out of gas and desperate for a room. You arrive there around 11pm and beg the receptionist. He thinks for a minute and realises that he can solve your problem. He instructs all residents to exit their rooms and enter the rooms immediately to the right. So the resident of room 0 is now in room 1, the resident of room 1 is now in room 2, etc. At the end, everybody has a new room and the receptionist proudly announces to you that room 0 is empty.

Around midnight, a driver arrives at the reception. He's the driver of von Neumann's Van, and von Neumann's Van is a very long van. It has seats $0, 1, 2 \ldots, n, \ldots$, one for each $n \in \mathbb{N}$. "All the passengers on my van need somewhere to spend the night" he tells the receptionist. The receptionist, tired from his long shift, reluctantly decides to accommodate the bus driver. After thinking about it, he ask everyone to get out of their rooms again, and instructs the resident of room 1 to go to room 2, the resident of room 2 to go to room 4, the resident of room 3 to go to room 6, etc. At the end, you're still in room 0, but rooms $1, 3, 5, \ldots$ are all empty, and each passenger on Cantor's Bus can spend the night in room $2n + 1$, where n was their seat number.

"What now?" the receptionist asks the new arrival. "Sorry sir" he says, "I'm the pilot of Peano's Plane. On my plane there are rows numbered by $0, 1, 2, \ldots, n, \ldots$, one for each $n \in \mathbb{N}$, and on each row there are seas $0, 1, \ldots, n, \ldots$, again one for each $n \in \mathbb{N}$. All my passengers need somewhere to spend the night!" The receptionist who is by now fed up with all these people ignoring the NO VACANCY sign outside the hotel, comes to your room:

Exercise. Find a way to fit every passenger on Peano's Plane in the hotel.

It's now past 3am, and Cantor's Cruise Ship (aptly named *"The Paradise"*) has reached the nearest port. It has rooms numbered by $x \in \mathbb{R}$. The captain approaches the reception. Before he even gets the chance to say anything, the receptionist looks at him, points at the NO VACANCY sign, through the window, and goes to sleep.

1. Pretty naive set theory

1.1. The basics. If you were excited to learn the formal definition of a set, you should probably curb your enthusiasm: We will not *formally* define what a set is (at least not in this class). It's more important, for now, to learn how to use sets. If that makes you feel a bit uneasy, keep in mind that people more-or-less understood what sets were for quite some time before Zermelo, Fraenkel, and the Axiom of Choice gave a formal definition.

For now, a **set** will just mean some "collection" of (mathematical) objects. Even though it is in quotes, the word collection is important here. **Sets are not lists**.

There are two usual ways that we can write a set down:

$$\Big\{ \text{ A (MORALLY) COMPLETE LIST OF ITS OBJECTS } \Big\},$$

or:

$$\Big\{ \text{ OBJECTS WITH SOME PROPERTY } \Big\}.$$

In the latter case, if P is a property we may write $\{x : P(x)\}$ to mean the set of all objects x satisfying property P. If X is a set and x some (mathematical) object, we write:

- $x \in X$ to mean that x is an *element of* X.
- $x \notin X$ to mean that x is *not an element of* X.

Example 1.1.1. Here's a couple of sets, illustrating the two ways we can write sets down, from above:

- $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ – a *morally* complete list of the elements of the set of **natural numbers**. The natural numbers are the starting point of most things,[1] so throughout this chapter we'll discuss them in much more detail.
- A less important example is the following $E = \{n : n \in \mathbb{N} \text{ and } P(n)\}$, where $P(n)$ holds of a natural number if and only if it is even – objects with some property.

Here is the first crucial property of sets. It is important enough to have a pretentious name:

> **Extensionality**: Two sets are the same if and only if they have the same elements.

[1] "God created the ~~integers~~ natural numbers, the rest is the work of man" and so on and so on.

A bit more mathematically put:

$$X = Y \text{ if and only if: } \begin{aligned}&(1)\ \textit{for all } x \in X, \textit{ we have } x \in Y\\ &\qquad\qquad \text{and}\\ &(2)\ \textit{for all } y \in Y, \textit{ we have } y \in X.\end{aligned}$$

That's kinda abstract, but it really is not saying all that much – well, actually, it's saying quite a fair bit. Here's an example:

Example 1.1.2. Consider the sets:

$$A = \{1, 2, 3, 4\} \text{ and } B = \{1, 3, 1, 4, 2, 2\}$$

Observe that for all $x \in A$ we have that $x \in B$ and for all $x \in B$ we have that $x \in A$. So, by extensionality, $A = B$. So, as I said before: *Sets are not lists!* The moral of the story is that order and repetition do not matter.

Extensionality has one funny consequence. It implies that if there is a set with no elements, then it is unique. As silly as it may sound, we will need to convince ourselves of the following:

Empty set: There is a set with no elements.

We will denote the unique set from above by $\emptyset$.

Definition 1.1.3. Let X and Y be sets. We say that X is a **subset** of Y if for all $x \in X$ we have that $x \in Y$. We denote this by $X \subseteq Y$.

Example 1.1.4. For all sets X, $X \subseteq X$ and $\emptyset \subseteq X$. The former one is clear, but maybe we should think a bit about the latter: Since the empty set has no elements, every element of the empty set is an element of any set. Now say that phrase three more times and see if the words "element" and "set" have lost all meaning.

With our new and updated terminology, extensionality can be written more succintly as follows:

$$X = Y \text{ if and only if: } (1)\ \ X \subseteq Y \text{ and } (2)\ \ Y \subseteq X$$

Comprehension. When we wrote that sets are collections of the form

$$\Big\{ \text{ OBJECTS WITH SOME PROPERTY } \Big\},$$

we made a big concession. We assumed the "axiom schema" of **full comprehension**, which can lead us into trouble if we are not careful. Consider for example the set R named after English philosopher[2] Bertrand Russel:

$$R = \{x : x \text{ is a set and } x \notin x\},$$

that is, the set of all elements that do not contain themselves.

Question: Does R contain itself?

Hmm... Okay let's not worry too much about all that now.

– End of digression –

Let's distract ourselves with some sets:

(1) $\mathbb{N} := \{0, 1, 2, \dots\}$, yeah that one showed up before...
(2) $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$, the set of **integers**.
(3) $\mathbb{Q}$, the set of **rational numbers** (fractions of integers).
(4) $\mathbb{R}$, the set of **real numbers** (integers followed by infinite decimal expansions).

1.2. Operations on sets. Okay great. So we have a few sets. How can we build more sets, from the sets we already have?

Definition 1.2.1. Let A and B be sets.

(1) The *union* of A and B denoted $A \cup B$, is defined by:
$$x \in A \cup B \text{ if and only if } x \in A \ \underline{\text{or}}\ \ x \in B.$$
(2) The *intersection* of A and B denoted $A \cap B$, is defined by:
$$x \in A \cap B \text{ if and only if } x \in A \ \underline{\text{and}}\ \ x \in B.$$
(3) The *relative complement* of A and B, denoted $A \setminus B$, is defined by:
$$x \in A \setminus B \text{ if and only if } x \in A \ \underline{\text{but}}\ \ x \notin B.$$

Example 1.2.2. Let $P \subseteq \mathbb{N}$ be the set of all prime numbers. Then $P \cap E = \{2\}$.

Example 1.2.3. This is all frivolous vocabulary to describe Venn diagrams. We all remember our Venn diagrams, don't we? At some point in high-school I'm sure we all had to draw Venn diagrams proving DeMorgan's Laws:

[2]who was joined by Bob Dylan in the club of people that have won a Noble prize in literature.

(1) $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

(2) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$.

More generally:

Exercise 1.2.4. Let I be a set and for each $i \in I$, suppose we are given a set A_i. Define:

$$\bigcup_{i \in I} A_i \text{ and } \bigcap_{i \in I} A_i,$$

in the obvious way, that is:

$$x \in \bigcup_{i \in I} A_i \text{ if and only if } x \in A_i \text{ for some } i \in I,$$

and

$$x \in \bigcap_{i \in I} A_i \text{ if and only if } x \in A_i \text{ for all } i \in I.$$

Let X, I and A_i (for $i \in I$) be sets. Prove DeMorgan's laws, i.e:

(1) $X \setminus \left(\bigcap_{i \in I} A_i\right) = \bigcup_{i \in I}(X \setminus A_i)$.

(2) $X \setminus \left(\bigcup_{i \in I} A_i\right) = \bigcap_{i \in I}(X \setminus A_i)$

Before we go on with more sets, let's go back to square zero. And then move to square one. And then square two. And then...

1.3. What's in a number. You may have heard something along the lines of "set theory acts as a foundation of mathematics" or some other abstract nonsense to that extent. Let's adopt this view for a bit. What this means is that in a sense **everything is a set**(?!). We will consider very briefly how every natural number is a set.

If everything is a set, the smallest natural number "should" be the smallest set. So, let's define:

$$0 := \emptyset.$$

Okay. But then 1 should be the next smallest set, right? Okay great, let's define:

$$1 := \{\emptyset\}$$

But at this point, we have done something rather interesting. We have identified 1 with the set consisting of all numbers smaller than it (which is just the set consisting of 0, which by the previous definition is just the set consisting of $\emptyset$). But this is a great idea, let's do it for 2 as well:

$$2 := \{0, 1\} = \{\emptyset, \{\emptyset\}\}.$$

If you stop to thing about what we just did for a minute you'll notice that we have done a second rather interesting thing! We have identified 2 not just with the set of numbers smaller than it, but also with the set $\{\emptyset\} \cup \{\{\emptyset\}\}$, that is, with the set that represents 1 together with the set containing the set that represents 1. So:

$$2 = 1 \cup \{1\}.$$

Let's turn this process into a definition:

Definition 1.3.1. Let X be a set. We define the *successor* of X, denoted X^+, to be the set:

$$X^+ := X \cup \{X\}.$$

Explicitly, $x \in X^+$ if and only if $x \in X$ or $x = X$.

So now, suppose we have defined the number n. To define the natural number $n+1$ we will simply set $n + 1 := n^+$, i.e. $n^+ = n \cup \{n\}$.

Exercise 1.3.2. Write down in full what the natural numbers $3, 4$ and 5 are.

THEOREM 1.3.3. *If* $m, n \in \mathbb{N}$ *then* $m \cup n = \max\{m, n\}$.

PROOF. We show, by induction on n that if $m \leq n$ then $m \subseteq n$. The base case $n = 0$ is trivial. Suppose that this is true for n and that $m \leq n + 1$. Then, either $m \leq n$ or $m = n + 1$. If $m = n + 1$, then we are done. If $m \leq n$, then, by induction $m \subseteq n$, so $m \subseteq n \cup \{n\} = n + 1$. □

– End of digression –

1.4. Relations and functions. We will now build up a little bit of vocabulary which you may have seen before, but it's important to all be on the same page:

Definition 1.4.1. Let A and B be sets.

(1) The *Cartesian product* of A and B, denoted by $A \times B$ is the set:

$$(x, y) \in A \times B \text{ if and only if } x \in A,\ y \in B.$$ [3]

[3] We should probably have some kind of "set theoretic" notion of ordered pair. One that always does the trick is the following $(x, y) := \{x, \{x, y\}\}$. What does that even mean?

(2) A *relation* R from A to B is a subset of $A \times B$.

(3) If R is a relation from A to B we define the *domain* of R, denoted $\mathsf{dom}(R)$, to be

$$\mathsf{dom}(R) := \{a \in A : \text{there is some } b \in B \text{ s.t. } (a, b) \in R\},$$

and the *range* of R, denoted $\mathsf{range}(R)$, to be

$$\mathsf{range}(R) := \{b \in B : \text{there is some } a \in A \text{ s.t. } (a, b) \in R\}.$$

(4) If $R \subseteq A \times B$ is a relation and $a \in \mathsf{dom}(R)$, then the *fibre* of a, denoted R_a, is the set:

$$R_a := \{b \in B : (a, b) \in R\}.$$

(5) A relation R from A to B is a *function*, if for all $x \in A$ there is a unique $y \in B$ such that $(x, y) \in R$. If f is a function from A to B, we may refer to the range of f as the *image* of f, and denote it $\mathsf{im}(f)$.

So a relation $R \subseteq A \times B$ is a function if and only if $A = \mathsf{dom}(R)$ and, for all $a \in A$, R_a is a singleton. If $R \subseteq A \times B$ is a relation such that for all $a \in \mathsf{dom}(R)$ we have that R_a is a singleton, but $A \neq \mathsf{dom}(R)$, then we call R a **partial function** (it is a function on the set $A' \subseteq A$, where $A' = \mathsf{dom}(R)$).

Remark 1.4.2. If $R \subseteq A \times B$ is a relation, we may write $R(a, b)$ to mean that $(a, b) \in R$. If $f \subseteq A \times B$ is a function, we denote this by writing $f : A \to B$, and we write $f(a) = b$ to mean that $(a, b) \in f$.

Let's look at an easy way of obtaining new relations from old:

Example 1.4.3. Given a relation $R \subseteq A \times B$, and subsets $A' \subseteq A, B' \subseteq B$ we define the **restriction** of R to $A' \times B'$, denoted $R \restriction_{A' \times B'}$ to be:

$$R \restriction_{A' \times B'} := R \cap (A' \times B').$$

The next example discusses one of the most common kinds of relations in mathematics. These will occupy our minds a bunch later on, so better introduce them as fast as possible.

Example 1.4.4. We say that a relation $R \subseteq A \times A$ is an **equivalence relation** if it satisfies the following three conditions:

(1) *Reflexivity*: For all $a \in A$ we have $R(a, a)$.

(2) *Symmetry*: For all $a, a' \in A$, if we have that $R(a, a')$ then we have that $R(a', a)$.

(3) *Transitivity*: For all $a, a', a'' \in A$ if we have that $R(a, a')$ and $R(a', a'')$ then we have that $R(a, a'')$.

For $a \in A$, we write $[a]_R$ for the set $\{b \in A : R(a, b)\}$, which we call the **equivalence class** of a.

Equivalence relations are useful, because they allow us to **quotient** things. Indeed, let A be a set and R an equivalence relation on A. Then, there is a set A/R, defined as follows:

$$A/R := \{[a]_R : a \in A\}.$$

We call a a **representative** of the equivalence class $[a]_R$. Equivalence classes can (and usually do) have many representatives.

Exercise 1.4.5. Let A be a set and R an equivalence relation on A. Prove that $R(a, b)$ if and only if $[a]_R = [b]_R$.

If A and B are sets, then A^B is the set of all functions from B to A, i.e.

$$A^B := \{f \subseteq A \times B : f \text{ is a function}\}.$$

Some people denote this by ${}^B A$, but I think this notation is rather obnoxious.[4]

Definition 1.4.6. A function $f : A \to B$ is called:

(1) *injective* if for all $b \in B$ there is at most one $a \in A$ such that $f(a) = b$.

(2) *surjective* if for all $b \in B$ there at least one $a \in A$ such that $f(a) = b$.

(3) *bijective* if it is both injective and surjective.

Given any relation $R \subseteq A \times B$, we can define a relation $R^{-1} \subseteq B \times A$ by:

$$R^{-1}(b, a) \text{ if and only if } f(a) = b.$$

Remark 1.4.7. If $f : A \to B$ is a bijective function, then for all $a \in A$ there is a unique $b \in B$ such that $f(a) = b$. In this case, f^{-1}, from above, is a function. We call f^{-1} the **inverse function** of f.

Exercise 1.4.8. Suppose that $f : A \to B$ is a bijective function. Check that:

(1) For all $a \in A$ we have that $f^{-1}(f(a)) = a$.

[4] I have strong opinions and I'm not ashamed of that.

(2) For all $b \in B$ we have that $f(f^{-1}(b)) = b$.

Definition 1.4.9. The *powerset* of a set A, denoted $\mathcal{P}(A)$, is defined by:

$$x \in \mathcal{P}(A) \text{ if and only if } x \subseteq A.$$

Remark 1.4.10. For all sets A, $\emptyset \in \mathcal{P}(A)$.

Exercise 1.4.11. Let A be a set. Show that there is a bijection between $\mathcal{P}(A)$ and 2^A.[5] (Recall that $2 = \{0, 1\}$ so $2^A = \{0, 1\}^A$ is the set of all functions from A to $\{0, 1\}$.)

1.5. Infinities come in different sizes. Before we define what the size of a set is, we should figure out when two sets have the same size. For finite sets we definitely understand this:

Two finite sets have the same size if they have the same number of elements.

Yeah, I also didn't think this course would be so deep! But wait... there is really a point here. A set X has n elements, if and only if there is a bijection between X and n (remember, or if you skipped Section 1.3, quickly learn that from now on $n = \{0, \ldots, n-1\}$). We don't know what the infinite analogues of the natural numbers are (yet) , but, here is the point. If two finite sets A and B have the same number of elements, then there is some $n \in \mathbb{N}$ and bijections:

$$f : A \to \{0, \ldots, n-1\} \text{ and } g : B \to \{0, \ldots, n-1\}$$

Thus, there is a bijection from $g^{-1} \circ f : A \to B$. This is good enough to be our starting definition:

Definition 1.5.1. We say that two sets A and B are *equinumerous*[6] if there is a bijection $f : A \to B$. We denote this by $A \sim Y$. We say that A is *subnumerous*[7] to B if there is some injective function $f : A \to B$, and denote this by $A \preceq B$.

For finite sets, this just says that A and B are equinumerous if and only if they have the same number of elements,[8] which is great, as this is what we wanted to generalise. What does it say for arbitrary sets? Well...

[5] The name *Potenzmenge* ("powerset") appears to come form *Untersuchungen über die Grundlagen der Mengenlehre* (1908) by Ernst Zermelo. For example, in 1906, Gerhard Hessenberg used the name *Menge der Teilmengen*, i.e. set of subsets. I hope this exercise justifies the naming convention.

[6] "Equi-" as in "equal" and "-numerous" as in "number".

[7] I wonder what the etymology of "subnumerous" is...

[8] The notion of subnumerosity is also pretty obviously an extension of what it means for a finite set to be of smaller size than another finite set.

Example 1.5.2. $\mathbb{N}$ and $\mathbb{Z}$ are equinumerous. For example, the function $f : \mathbb{N} \to \mathbb{Z}$ given by:

$$x \mapsto \begin{cases} \frac{x}{2} & \text{if } x \text{ is even;} \\ \frac{-(x+1)}{2} & \text{if } x \text{ is odd.} \end{cases}$$

is a bijection.

So it really doesn't make all that much sense to ask that two infinite sets have "the same number of elements", but it's close enough:

Exercise 1.5.3. Prove that $\sim$ is an equivalence relation.[9]

The main idea behind all of this is that we would somehow like to think of the "size" of a set A as the equivalence class of A modulo the equivalence relation $\sim$. This is not so simple (first of all, because there is no set of all sets, and secondly because there is no guarantee that given two sets one will be subnumerous to the other).

Exercise 1.5.4. Show that $\mathbb{N} \sim \mathbb{N} \times \mathbb{N}$. Deduce that $\mathbb{N} \sim \mathbb{Q}$.

[*Hint.* The following zig-zag idea may be helpful:

$$\begin{tikzcd}
(0,0) \arrow[r] & (0,1) \arrow[dl] & (0,2) \arrow[r] & (0,3) \arrow[dl] & (0,4) \arrow[r] & (0,5) \arrow[dl] \\
(1,0) \arrow[d] & (1,1) \arrow[ur] & (1,2) \arrow[dl] & (1,3) \arrow[ur] & (1,4) \arrow[dl] & \ddots \\
(2,0) \arrow[ur] & (2,1) \arrow[dl] & (2,2) \arrow[ur] & (2,3) \arrow[dl] & \ddots & \\
(3,0) \arrow[d] & (3,1) \arrow[ur] & (3,2) \arrow[dl] & \ddots & & \\
(4,0) \arrow[ur] & (4,1) \arrow[dl] & \ddots & & & \\
(5,0) & \ddots & & & &
\end{tikzcd}$$

Remember, the goal is to define a bijection from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$.]

All of this begs the obvious question:

Question. What about $\mathbb{N}$ and $\mathbb{R}$?

THEOREM 1.5.5 (Cantor). *Let A be a set. There is no surjective function from A to $\mathcal{P}(A)$. In particular $A \not\sim \mathcal{P}(A)$.*

---

[9] If you study more set theory you may learn that this exercise is "formally" wrong, at least according to some people. I don't really care.

The proof is often referred to as **Cantor's diagonal argument** – we've actually kind of encountered this kind of logic already in Russel's paradox. Two of the most important proofs in this course will involve funky variations on diagonal arguments, so it's perhaps rather important to understand the grandfather of them all.

PROOF OF CANTOR'S THEOREM, TAKE 1. Let $B := \{x \in A : x \notin f(x)\}$.[10] Suppose towards a contradiction that $f(x) = B$ for some $x \in X$. Then we have that $x \in B$ if and only if $x \notin B$, a contradiction. $\square$

This proof was like very easy, but I claim that we actually showed a bunch of cool stuff:

Corollaries of Cantor's theorem

(1) SOME INFINITIES ARE BIGGER THAN OTHERS: Well, we did show that there are infinite sets which are not equinumerous. For any set A, we have that A is subnumerous to $\mathcal{P}(A)$ [Why?] but by Cantor's theorem they are not equinumerous. Let's start with a "small-ish" infinite set, say $\mathbb{N}$:

$$\mathbb{N} \prec \mathcal{P}(\mathbb{N}) \prec \mathcal{P}(\mathcal{P}(\mathbb{N})) \prec \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N}))) \prec \cdots ,$$

where $A \prec B$ means $A \preceq B$ and $A \not\sim B$.

(2) THERE IS NO SET OF ALL SETS: Suppose that V is "the set of all sets". Since any set of subsets of V is a set, $\mathcal{P}(V) \subseteq V$, so $\mathcal{P}(V) \preceq V$, contradicting Cantor's theorem.

(3) THERE ARE MORE REAL NUMBERS THAN NATURALS: Remember that $\mathbb{R}$ is the set of all integers followed by infinite binary expansions:

$$n.a_0a_1a_2a_3\ldots, a_i \in \{0,1\}.$$

The proof of Cantor's theorem may have gone a bit fast. Can we try it again?

PROOF OF CANTOR'S THEOREM, TAKE 2. Let's just prove that the set $2^{\mathbb{N}}$ of infinite 0-1 sequences is not in bijection with $\mathbb{N}$. Let $f : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ be *any* map.

[10] This is the diagonalisation part of the proof (see TAKE 2 and the discussion after it, in the next page).

We can just write down the elements in the image of f, as follows:

$$\begin{aligned}
f(1) &= a_1^1 a_2^1 a_3^1 a_4^1 \dots a_n^1 \dots \\
f(2) &= a_1^2 a_2^2 a_3^2 a_4^2 \dots a_n^2 \dots \\
f(3) &= a_1^3 a_2^3 a_3^3 a_4^3 \dots a_n^3 \dots \\
&\vdots \qquad\qquad \ddots \quad \vdots \\
f(n) &= a_1^n a_2^n a_3^n a_4^n \dots a_n^n \dots \\
&\vdots
\end{aligned}$$

where $a_i^j \in \{0,1\}$, for all $i, j \in \mathbb{N}$.

Now consider the following element of $\{0,1\}^{\mathbb{N}}$:

$$\bar{b} := b_1 b_2 b_3 \dots b_n \dots,$$

where:

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

So $\bar{b} \neq f(n)$, for each $n \in \mathbb{N}$. Indeed, $f(n) = a_1^n a_2^n \dots a_n^n \dots$, $\bar{b} = b_1 \dots b_n \dots$ and by definition:

$$b_n \neq a_n^n.$$

So, the function we started with was not a bijection. □

I hope we all now (literally) see why this is called a *diagonal* argument. We always look at the n-th coordinate of $f(n)$. How does this proof compare to the first proof that we gave? Well, we can view $f(n)$ as a subset of $\mathbb{N}$, namely $f(n) = \{i \in \mathbb{N} : f(n)_i = 1\}$, where $f(n)_i = a_i^n$. With this in mind, $\bar{b}$ is the set $\{i \in \mathbb{N} : f(i)_i = 0\} = \{i \in \mathbb{N} : i \notin f(i)\}$ (cf. with Take 1 of the proof). A somewhat subtle point emphasised by the Take 1 argument is that we did not need to assume that the set we started with was **countable** (see the next definition).

Definition 1.5.6. A set A is *countable* if it is subnumerous to $\mathbb{N}$. If $\mathbb{N}$ is not subnumerous to A, we say that A is *finite*. If A is not countable, we say that it is *uncountable*.

By definition, $\mathbb{N}$ is countable and by our discussion above, $\mathbb{R}$ is uncountable. The intuition is that a set is countable if we can write its elements as a "nice" list of the form $0, 1, 2, 3, \dots$.

If you're curious individuals, maybe at this point you're asking yourselves:

Question. How much bigger than $\mathbb{N}$ is $\mathbb{R}$?

Cantor-Bernstein. We will briefly discuss at the end of this section, that this is a fool's errand, but for now, we want to define the infinite analogues of natural numbers (the *ordinals*) and their corresponding sizes (the *cardinals*).[11] First, let's show that $\preceq$ behaves kind of like what we intuitively understand an order to be, and then let's define what an order actually is.

THEOREM 1.5.7 (Cantor-Bernstein). *Let A and B be sets. If $A \preceq B$ and $B \preceq A$ then $A \sim B$.*

PROOF. Recall that $A \preceq B$ means that there is some injective map $f : A \to B$ and that $B \preceq A$ means that there is some injective map $g : B \to A$. We want to show that $A \sim B$, i.e. that there is a bijection from A to B.

It suffices to show that $\mathsf{im}(f) \sim B$ (since by definition, $A \sim \mathsf{im}(f)$ and you have shown that $\sim$ is transitive), so without loss of generality, we may assume that We can replace $A = \mathsf{im}(f) \subseteq B$.

Let $C := \{g^n(x) : x \in B \setminus A\}$, where

$$g^n(x) = \underbrace{g \circ g \circ \cdots \circ g(x)}_{n \text{ times}}.$$

Define:

$$h : B \mapsto A$$

$$b \mapsto \begin{cases} g(b) & \text{if } b \in C \\ b & \text{if } b \in B \setminus C \end{cases}$$

It is easy to see that h is injective (since the composition of injective functions is injective), so we need only show that it is surjective. But indeed, if $x \in A$, then either $x \in A \cap C$ in which case $x = g(y)$ for some $y \in C$, and if $x \notin c$ then $h(x) = x$. $\square$

– End of Digression –

[11] As we will see, this distinction disappears in the finite case.

Homework 1

With that all out of the way, we'll go into another digression, which will be followed by another digression. In fact:

This is formally as much of this part of the course as we will cover in class.

I got rather carried away when I was writing these notes, but whatever... The rest is digression after digression.[12]

1.6. Ordinals. By the end of this little section, we will have uncovered the key properties that make numbers so good at listing things by (what a daft statement) – and through this we will be able to define their infinite counterparts (ah he was making a point). We will have to be a bit abstract at the beginning.

Definition 1.6.1. A *partial order* on a set X is a relation $\prec \subseteq X \times X$ satisfying the following properties:

(1) *Irreflexivity*: For all $x \in X$, $x \not\prec x$.

(2) *Transitivity*: For all $x, y, z \in X$, if $x \prec y$ and $y \prec z$, then $x \prec z$.

We say that $\prec$ is a total order, if, in addition we have the following:

(3) *Trichotomy*: For all $x, y \in X$, if $x \not\prec y$ and $y \not\prec x$, then $x = y$.

Example 1.6.2. Let X be a set. Then, $\subsetneq$ (where $A \subsetneq B$ if and only if $A \subseteq B$ and $A \neq B$, for all $A, B \in \mathcal{P}(X)$) is a partial order on $\mathcal{P}(X)$. If X has at least two elements, then this is not a total order.

Example 1.6.3. Let n be a natural number, viewed as a set. Define a relation $\prec$ on $n = \{0, \dots, n-1\}$ by $x \prec y$ if and only if $x \in y$ (recall that the elements of n are themselves sets of natural numbers). For example, on $2 = \{0, 1\}$ we have that:

$$0 \in 1, 1 \notin 0.$$

More generally:

THEOREM 1.6.4. *Let $n \in \mathbb{N}$ and $m, m' \in n$, then, the following are equivalent:*

(1) $m \in m'$.

(2) $m < m'$ (in the usual order of natural numbers).

[12] Yeah, I suck at keeping stories short. I know.

Proof. We show this by induction. It is obvious for $\emptyset$ (as it has no elements! I will keep stressing this). So suppose that it is true for $n \in \mathbb{N}$. We need to show that it is true for $n+1 = n \cup \{n\}$. Suppose that $m, m' \in n+1$. Then, there are two cases to consider:

- Case 1. Both $m, m' \in n$. In this case, we are done, by induction.
- Case 2. $m \in n$ and $m' = n$. In this case, we are done, by assumption.

$\square$

Exercise 1.6.5. Let $(X, \prec)$ be a **poset**.[13] Define a new relation $\preceq \subseteq X \times X$ as follows:

$$x \preceq y \text{ if and only if } x \prec y \text{ or } x = y.$$

(1) Show that $\preceq$ satisfies the following:

  (a) *Reflexivity*: For all $x \in X$, $x \preceq x$.

  (b) *Antisymmetry*: For all $x, y \in X$, if $x \preceq y$ and $y \preceq x$ then $x = y$.

  (c) *Transitivity*: For all $x, y, z \in X$, if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

(2) Suppose that X is a set and $\trianglelefteq$ is a **binary relation** (i.e. a subset of $X \times X$ satisfying (a)-(c) from above. Show that the relation $\triangleleft$ defined by:

$$x \triangleleft y \text{ if and only if } x \trianglelefteq y \text{ and } x \neq y$$

is a partial order on X.

By the exercise above, I will not pay too much attention on whether my partial orders allow equality or not (it should be clear from the notation though).

Definition 1.6.6. Let $(X, \prec)$ be a poset and $Y \subseteq X$. We say that $y \in Y$ is a *minimal element* of Y if there is no $y' \in Y$ such that $y' \prec y$. We say that $y \in Y$ is a *maximal element* of Y if there is no $y' \in Y$ such that $y \prec y'$. We call $x \in X$ a *lower* (resp. *upper*) *bound* for Y if for all $y \in Y$ we have that $x \prec y$ (resp. $y \prec x$).

Exercise 1.6.7. Let $(X, \prec)$ be a poset and $Y \subseteq X$. Define the following:

(1) $y \in Y$ is a **smallest element** of Y if for all $y' \in Y$ we have that $y \preceq y'$.

(2) $y \in Y$ is a **largest element** of Y if for all $y' \in Y$ we have that $y' \preceq y$.

[13] This is slang for X is a set, $\prec$ is a partial order on X

Prove that smallest and largest elements of a subset of a poset (if they exist) are unique.[14] Give examples of why this need not be the case for minimal and maximal elements.[15]

Definition 1.6.8. We say that a poset $(X, \prec)$ is *well-founded* if every non-empty subset of X has a minimal element. We say that $(X, \prec)$ is a *well-order* if it is a well-founded partial order.

Example 1.6.9. Here are some examples and non-examples:

(1) Every natural number (in the sense of Section 1.3) is well-ordered.

(2) $(\mathbb{N}, \leq)$ is well-ordered.

(3) The total order $(\mathbb{Z}, \leq)$ is not well-founded.

The third example above, essentially generalises to allow us to characterise all the well-founded partial orders. First, a small definition. We say that a subset $C \subseteq X$ of a poset $(X, \prec)$ is a **chain** if $(C, \prec\restriction_{C\times C})$ is a total order. A chain is called **descending** if it has no minimal (i.e. smallest) element.

THEOREM 1.6.10. *Let $(X, \prec)$ be a poset. Then, the following are equivalent:*

(1) $(X, \prec)$ is well-founded.

(2) There are no descending chains in $(X, \prec)$.

PROOF.

(1) $\implies$ (2). Suppose that $C \subseteq X$ is a descending chain in X. Then C has no minimal element.

(2) $\implies$ (1). Suppose that $(X, \prec)$ is not well-founded. Define a map:

$$\begin{aligned} f : X &\to \mathcal{P}(X) \\ x &\mapsto \{y \in X : y \prec x\} \end{aligned}$$

Since X is not well-founded, there is some $Y \subseteq X$ which has no minimal element. Let $y_0 \in Y$. Then, by assumption, y_0 is not minimal in Y, so there is some $y_1 \in Y$ such that $y_1 \prec y_0$, i.e. some $y_1 \in f(y_0) \cap Y$. But then, $y_1 \in Y$ so it is not minimal, so there is some $y_2 \in f(y_1) \cap Y = f(y_0) \cap f(y_1) \cap Y$. Continuing like this, we can construct for each $n \in \mathbb{N}$ some y_n such that $y_n \in Y \cap \bigcap_{i<n} f(y_i)$. Thus $\{y_0, y_1, \ldots, \}$ is a descending chain in X. $\square$

[14]Thus we may refer to them as **the** smallest and largest elements of said subset.

[15]Thus we may *not* refer to them as the minimal and maximal elements.

Exercise 1.6.11. The proof of Theorem 1.6.10 is written rather informally. Write out the inductive proof carefully.

Exercise 1.6.12. Show that the following are equivalent for a set X:

(1) X is equinumerous to some natural number $n \in \mathbb{N}$.

(2) $\mathcal{P}(X)$ is well-founded.

Now that we understand well-orders rather well, let's try to single out the unique property of natural numbers we want to extend to infinite sets:

Definition 1.6.13. A set X is *transitive* if for all $x \in X$ we have that $x \subseteq X$. Explicitly, this says that for all $x \in X$, if $y \in x$ then $y \in X$.

Wait, what? Yes, this is admittedly very confusing. But let's go back to our friends the natural numbers:

$$0 = \emptyset, \quad 1 = \{0\}, \quad 2 = \{0, 1\}, \quad 3 = \{0, 1, 2\}, \quad 4 = \{0, 1, 2, 3\}, \ldots, \\ n = \{0, 1, 2, \ldots, n - 1\}, \ldots$$

Let's look at 2 first. Suppose that $x \in 2$. Then $x = 0$ or $x = 1$. If $x = 0$, then $x = \emptyset$, so $x \subseteq 2$ (the empty set is a subset of every set). On the other hand, if $x = 1$ then $x = \{0\}$ and $\{0\} \subseteq \{0, 1\}$. More generally:

Theorem 1.6.14. *Every natural number is a transitive set.*

Proof. We prove this by induction. Obviously $\emptyset$ is transitive (it has no elements!). So, suppose that n is transitive. We have to show that $n + 1$ is transitive. By definition:

$$n + 1 = n \cup \{n\}.$$

So, if $x \in n + 1$ then either $x \in n$ or $x = n$. On the one hand if $x \in n$ then $x \subseteq n$, by induction. On the other hand, if $x = n$ then, for all $y \in x$ we have that $y \in n$ so $y \in n \cup \{n\} = n + 1$. □

Definition 1.6.15. A set X is an *ordinal* if it is transitive, and the relation $\prec$ given by:

$$x \prec y \text{ if an only if } x \in y,$$

is a total order on X.

So if we put together Theorems 1.6.4 and 1.6.14 we have actually shown the following:

Corollary 1.6.16. *Every natural number is an ordinal.*

Let's start building more ordinals. First, we will show that there is at least one infinite ordinal:

THEOREM 1.6.17. *The set of all natural numbers*[16] *is an ordinal.*

PROOF. We need to show that $\mathbb{N}$ is transitive and that $\in$ is a total order on $\mathbb{N}$.

- TRANSITIVE: First, we will show that if $n \in \mathbb{N}$ then $n \subseteq \mathbb{N}$, by induction. This is obvious when $n = \emptyset$ (for the n-th time, $\emptyset$ is a subset of every set). So we need to show that for $n+1 \in \mathbb{N}$, $n+1 \subseteq \mathbb{N}$. But $n+1 = n \cup \{n\}$. By induction, we know that $n \subseteq \mathbb{N}$ and by assumption, we know that $n \in \mathbb{N}$, so $n \cup \{n\} \subseteq \mathbb{N}$.
- TOTALLY ORDERED BY $\in$: We have to show that for all $m, n \in \mathbb{N}$ if $n \notin m$ and $m \notin n$ then $m = n$. Let $m' = m \cup n = \max\{m, n\}$ (where the maximum is taken in the usual order of natural numbers, by Theorem 1.3.3). Then, by definition, $m, n \in m+1$ and, by Theorem 1.6.4, $\in$ is a total order on $m+1$.

□

Okay great we have at least one infinite ordinal, and you should be asking yourselves now. Can we build more?

Proposition 1.6.18. *Let α be an ordinal. Then α^+ is an ordinal.*

Exercise 1.6.19. Prove Proposition 1.6.18

So $\mathbb{N}$ is an ordinal, and so is $\mathbb{N}^+ = \mathbb{N} \cup \{\mathbb{N}\}$ and so is $(\mathbb{N}^+)^+ = \mathbb{N} \cup \{\mathbb{N}\} \cup \{\mathbb{N} \cup \{\mathbb{N}\}\}$, etc.

Exercise 1.6.20. Prove that $\mathbb{N}^+$ is a countable set. More generally, prove that if α is an infinite ordinal, then α and α^+ are equinumerous.

[16] Okay, we haven't formally explained why the set of natural numbers is a set, but trust me, for now, it is one.

Okay that's all good and well, but can we build uncountable ordinals?

Well... yes, I know we're all excited, but let's try to be a bit more patient. We'll get there, but it may take a minute.

Lemma 1.6.21. *Let α, β be ordinals. Then:*

(1) $\alpha \notin \alpha$.

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

(3) $\beta \subseteq \alpha$ *if and only if* $\beta \in \alpha$ *or* $\beta = \alpha$.

Proof.

(1) By definition, $(\alpha, \in)$ is a partial order. Thus, for all $\beta \in \alpha$ we have that $\beta \notin \beta$, by irreflexivity. Suppose that $\alpha \in \alpha$. Then $\alpha \notin \alpha$, a contradiction.

(2) Suppose that $x \in \alpha$. Then $x \subseteq \alpha$, so $\in\restriction_{x \times x}$ is a well-order. To see that x is transitive, suppose that $y \in x$. We have to show that $y \subseteq x$, so suppose that $z \in y$, then $z, y \in \alpha$ and by transitivity of $\in$ we have that $z \in x$.

(3) On the one hand, suppose that $\beta \subsetneq \alpha$. Let $x \in \alpha \setminus \beta$ be minimal. By minimality assumption, if $y \in \alpha$ is such that $y \in x$, then $y \in \beta$. On the other hand, if $y \in \beta$ then $y \in x$ (otherwise $x \in y$, since $\in$ is a total order on α), and thus $x \in \beta$, contradicting the assumption that $x \in \alpha \setminus \beta$. Thus $\beta = x \in \alpha$. Conversely, suppose that $\beta \in \alpha$ then, since α is transitive, $\beta \subseteq \alpha$, and we are done.

□

Exercise 1.6.22. Let X be a non-empty set of ordinals. Prove that $\bigcap X := \bigcap_{\alpha \in X} \alpha$ is the smallest element of X.

[*Hint.* First, prove that $\bigcap X$ is an ordinal. Then, show that $\bigcap X \in X$, using Lemma 1.6.21(3).]

The intuition behind the next theorem is that $\in$ behaves like a total order around ordinals:

Theorem 1.6.23. *Let α and β be ordinals. Then, exactly one of the following holds:*

$$(1)\ \alpha \in \beta; \quad (2)\ \alpha = \beta; \quad (2)\ \beta \in \alpha$$

PROOF. Let $X = \{\alpha, \beta\}$. By Exercise 1.6.22, $\cap X = \alpha \cap \beta$ is an ordinal, and it is the least of α and β. If $\alpha \cap \beta = \alpha$, then $\alpha \subseteq \beta$ so by Lemma 1.6.21(3), $\alpha \in \beta$ or $\alpha = \beta$. Similarly, if $\alpha \cap \beta = \beta$, then $\alpha = \beta$ or $\beta \in \alpha$. The fact that exactly one of the three cases holds is immediate because $\in$ is a partial order. $\square$

As is usual in mathematics, once we have defined some class of objects and have some examples of objects in the class, we want general ways of constructing new objects in the class. For ordinals, we have shown that successors of ordinals , elements of ordinals, and intersections of ordinals are all ordinals. The next proposition allows us to build even bigger ordinals:

Proposition 1.6.24. *Let X be a set of ordinals. Then $\bigcup X := \bigcup_{\alpha \in X} \alpha$ is an ordinal.*

PROOF. Observe that the union of transitive sets is transitive (if $y \in \bigcup X$ then $y \in x$ for some $x \in X$, but x is transitive, so $y \subseteq x$, so $y \subseteq \bigcup X$). Since $\bigcup X$ consists of ordinals, by Theorem 1.6.23, $\in$ is a total order on $\bigcup X$. Let $Y \subseteq X$ be non-empty. Then $\bigcap Y$ is the smallest element of Y by Exercise 1.6.22. $\square$

Exercise 1.6.25. Prove that if X is a set of ordinals, then, for all ordinals γ, if $\gamma \in \bigcup X$ then there is some $\alpha \in X$ such that $\gamma \in \alpha$.

Definition 1.6.26. Let $(X, <_X)$ and $(Y, <_Y)$ be two partial orders. We say that they are *order-isomorphic* if there is some bijection $f : X \to Y$ such that:

$$x <_X x' \text{ if and only if } f(x) < Y f(x').$$

The next theorem is crucial. Indeed, part of the point of why we've been doing all of this abstract work is that putting Theorem 1.6.27 together with the Theorem 1.6.23 and the *well-ordering principle* (see next section) will allow us to define an honest to god way of sizing up infinite sets.

THEOREM 1.6.27. *Every well-ordered set is order-isomorphic to a unique ordinal.*

PROOF. Optional HW. $\square$

And now for the main course (somewhat due to Hartogs):

THEOREM 1.6.28 (Baby Hartogs). *Uncountable ordinals exist.*

Proof. Let X be the set of all well-orders on $\mathbb{N}$ (i.e. X consists of all subsets of $\mathbb{N} \times \mathbb{N}$ which are well-founded total orders). Define an equivalence relation $\sim$ on X by

$$\prec_1 \sim \prec_2 \text{ if and only if there is an order-isomorphism from } (\mathbb{N}, \sim_1) \text{ to } (\mathbb{N}, \sim_2)$$

Exercise 1.6.29. Prove that $\sim$ is an equivalence relation.

Let $Y = X/\sim$. For all $[\prec]_\sim \in Y$ there is, by Theorem 1.6.27, a unique countable ordinal $\alpha_\prec$ which is order-isomorphic with $(\mathbb{N}, \prec)$. Let Z be the set of all such ordinals (i.e. Z is a set consisting of countable ordinals, each of which is a representative of an equivalence class of $\sim$).

By Proposition 1.6.24, $\bigcup Z$ is an ordinal. By Exercise 1.6.25, every ordinal $\alpha < \zeta$ is order-isomorphic to some well-ordering of $\mathbb{N}$.

We claim that $\zeta = \bigcup Z$ is uncountable. Suppose not. Then ζ^+ would also be countable (by Exercise 1.6.20), but if ζ^+ were countable, then by definition, we would have an injection $f : \zeta^+ \to \mathbb{N}$. Clearly, ζ^+ is infinite, so there is an injection $g : \mathbb{N} \to \zeta^+$, and thus by Cantor-Bernstein there is a bijection $h : \zeta^+ \to \mathbb{N}$. But then, define an order $\prec$ on $\mathbb{N}$ by

$$n \prec m \text{ if and only if } g^{-1}(n) \in g^{-1}(m)$$

This is a well-order, so ζ^+ is order-isomorphic to a well-order on $\mathbb{N}$, and hence $\zeta^+ \in \zeta$. This is a contradiction.[17] □

Let's list some of the ordinals we have seen so far:

- Natural numbers are ordinals. For $n \in \mathbb{N} \setminus \{0\}$, the ordinal n is of the form $m + 1$, for some $m \in \mathbb{N}$.
- The set of all natural numbers is an ordinal.
- $\mathbb{N}^+$, $(\mathbb{N}^+)^+$), etc. are all ordinals. Let's write $\omega + n$ for the set obtained by iterating $(\cdot)^+$ on $\mathbb{N}$, n times.
- $\bigcup_{n \in \mathbb{N}} (\omega + n)$ is an ordinal. Let's call this $\omega + \omega$.
- $\bigcup_{n \in \mathbb{N}} (\omega + \omega + n)$ is an ordinal. Let's call this set $\omega + \omega + \omega$. Similarly, we can build $\underbrace{\omega + \cdots + \omega)}_{n \text{ times}}$, for $n \in \mathbb{N}$

[17] This is the easiest proof of Hartogs we can give with the assumptions we have so far. Once we learn about the *well-ordering principle*, we will be able to give a one-line proof.

- $\bigcup_{n\in\mathbb{N}}(\underbrace{\omega+\cdots+\omega}_{n \text{ times}})$. Let's call this set $\omega\times\omega$.
- We can keep going ad infinitum (quite literally).

The keen eyed amongst you may have seen already that ordinals feel like they come in two different flavours. They are either of the form α^+ for some ordinal α or they are a(n ordinal-indexed) union of smaller ordinals:

Definition 1.6.30. An ordinal of the form α^+ is called a *successor* ordinal. If λ is an ordinal that is not a successor ordinal, then we say that λ is a limit ordinal.

Example 1.6.31. $\mathbb{N}$ is an ordinal. If it were a successor ordinal, then it would be of the form α^+, for some $\alpha\in\mathbb{N}$. But, for every $\alpha\in\mathbb{N}$ we know that α^+ is finite, which is a contradiction. Thus, $\mathbb{N}$ is a limit ordinal.

The next exercise justifies the discussion before Definition 1.6.30:

Exercise 1.6.32. Let λ be a non-empty ordinal. Prove that the following are equivalent:

(1) λ is a limit ordinal.

(2) $\lambda=\bigcup_{\mu<\lambda}\mu$.

Exercise 1.6.33. Prove that $\mathbb{N}$ is the smallest non-empty limit ordinal.

We kinda sorta saw how to multiply ordinals (ish). The last exercise in this section introduces the concept of ordinal exponentiation:

Exercise 1.6.34 (Baby Ordinal Exponentiation)**.** Let $(X,<_X)$ and $(Y,<_Y)$ be totally ordered sets, and assume that X admits a smallest element $0\in X$. Let $X^{(Y)}$ be the set of all functions from Y to X of **finite support**, that is, for all $f\in X^{(Y)}$ we have

$$\mathsf{supp}(f):=\{y\in Y: f(y)\neq 0\}$$

is finite. Define a relation $\prec$ on $X^{(Y)}$ by

$f\prec g$ if and only if

$$\text{there is some } y\in Y \text{ s.t. } f(y)<_X g(y) \text{ and } f(y')=g(y') \text{ if } y'<_Y y.$$

Prove that:

(1) $\prec$ is a total order on $X^{(Y)}$.

(2) If $(X,<_X)$ and $(Y,<_Y)$ are both well-founded then so is $(X^{(Y)},\prec)$.

1.7. The Axiom of Choice. We previously defined the Cartesian product of two sets. This, of course, allows us to define the Cartesian product of finitely many sets, by induction:

$$A_0 \times A_1 \times A_2 = (A_0 \times A_1) \times A_2,$$

and we already know what $A_0 \times A_1$ is. More generally:

$$A_0 \times A_1 \times \cdots \times A_n \times A_{n+1} := (A_0 \times \cdots \times A_n) \times A_{n+1}.$$

But what about Cartesian products of infinitely many sets? The induction thing we pulled off above won't work anymore, so we have to be more clever. Let's see what's so special about the Cartesian product of two sets:

$$A_0 \times A_1 = \{(a_0, a_1) : a_0 \in A_0, a_1 \in A_1\},$$

so in a stupid sense, every element of $A_0 \times A_1$ is just a function from $\{0,1\}$ to $A_0 \cup A_1$ such that $f(0) \in A_0$ and $f(1) \in A_1$, and these are all the elements of $A_0 \times A_1$. Right? Functions from $\{0,1\}$ to a set X are just sets $\{(0,x),(1,x')\}$, for $x, x \in X$.

This is kind of silly, but really powerful. Indeed, if I is any set and $(X_i)_{i \in I}$ some family of sets, we define

$$\prod_{i \in I} X_i := \left\{ f : I \to \bigcup_{i \in I} : f(i) \in X_i \text{ for all } i \in I \right\}.$$

By our discussion above, when I is finite, this is nothing new, but now we can talk about arbitrary products of sets.

Axiom of Choice: The product of a family of non-empty sets is non-empty.

This innocuous statement has many many consequences. The next theorem (which we won't prove here) is really fundamental:

THEOREM 1.7.1. *The following are equivalent:*[18]

1. *The Axiom of Choice.*
2. *The* ***Well-ordering Principle****: Every set can be well-ordered.*[19]
3. ***Zorn's Lemma****: Let $(X, <)$ be a poset. If every chain in X has an upper bound in X then X has a maximal element.*

PROOF (SKETCH).

[18]Assuming the axioms of Zermelo-Fraenkel set theory (which you can google if you want)

[19]This means that for every set X there is a binary relation $\preceq$ on X which is a well-order.

"The Axiom of Choice is obviously true, the Well-ordering principle obviously false, and who can tell about Zorn's lemma?"

– J. Bona

□

1.8. Cardinals. Now that we have orders, we'd like to discuss their "sizes". As we saw previously, there are many infinite countable ordinals, but $\mathbb{N}$ is the smallest one. Somehow we want to make this canonical:

Definition 1.8.1. An ordinal is a *cardinal* if it's not equinumerous to any smaller ordinal.

Example 1.8.2. Any finite ordinal is a cardinal. $\mathbb{N}$ is also a cardinal. When considered as a cardinal, I will sometimes denote it by $\aleph_0$. And since we're here, when considered as an ordinal I will sometimes denote it by ω. I already pulled this trick in a previous paragraph.

Exercise 1.8.3. If α is an infinite ordinal, then α^+ is not a cardinal. [*Hint.* Prove that α and α^+ are equinumerous.]

Proposition 1.8.4. *Any set X is equinumerous to a unique cardinal.*

Proof. By the well-ordering principle, every set is equinumerous to an ordinal α. Let $\beta \le \alpha$ be minimal such that β is equinumerous to α. Then β is a cardinal and is in bijection with X. Uniqueness follows by minimality. □

We will write $|X|$ for the unique cardinal that X is equinumerous to. We call $|X|$ the **cardinality** of X.

Remark 1.8.5. If we return to our previous definition of countability, we see that: X is countable if and only if $|X| \le \aleph_0$, and X is finite if and only if $|X| < \aleph_0$.

Lemma 1.8.6. *Let X and Y be non-empty sets. Then, the following are equivalent:*

(1) $|X| \le |Y|$.

(2) There is an injective function $f : X \to Y$.

(3) There is a surjective function $g : Y \to X$.

Proof.

(1) $\implies$ (2): Suppose that $|X| = \kappa \leq \lambda = |Y|$. Then, there are bijections $f : X \to \kappa$, $g : Y \to \lambda$ and an injection $h : \kappa \to \lambda$. Thus we have an injection $g^{-1} \circ h \circ f : X \to Y$.

(2) $\implies$ (3): Let $f : X \to Y$ be an injective map. Fix some $x_0 \in X$. Define a map:

$$g : Y \to X$$
$$y \mapsto \begin{cases} x_0 & \text{if } y \notin \mathsf{im}(f) \\ f^{-1}(y) & \text{otherwise.} \end{cases}$$

This is surjective.

(3) $\implies$ (1): Let $f : Y \to X$ be a surjective map. Then, there is a surjective map $g : \lambda \to \kappa$, where $\lambda = |Y|$ and $\kappa = |X|$. Given $\alpha \in \kappa$, there is a minimal $\alpha_m \in \lambda$ such that $g(\alpha_m) = \alpha$ (since λ is an ordinal), so set $f(\alpha) = \alpha_m$. Then $f : \kappa \to \lambda$ is an injection, showing that $\kappa \leq \lambda$. □

Going back to Cantor's theorem, we can show that there is **no largest cardinal**. Indeed, for any cardinal λ, $|\mathcal{P}(\lambda)| \geq \lambda$.

Definition 1.8.7. Let κ be a cardinal. Then, we define the *cardinal successor* of κ, denoted $\kappa + 1$ to be the least element of the set

$$\{\lambda : \lambda > \kappa \text{ is a cardinal, and } \lambda \leq |\mathcal{P}(\kappa)|\}.$$

We can now define, for any ordinal α a cardinal $\aleph_\alpha$, as follows:

(1) $\aleph_0 := \omega$

(2) $\aleph_{\alpha^+} := \aleph_\alpha + 1$

(3) $\aleph_\lambda := \bigcup_{\mu<\lambda} \aleph_\mu$, when λ is a limit ordinal.

To close this little discussion, recall that in Exercise 1.6.34 we kind of defined ordinal exponentiation. For cardinals the situation is easier:

$$\kappa^\lambda = |\kappa^\lambda|.$$

Theorem 1.8.8. *Every infinite cardinal is of the form $\aleph_\alpha$ for some α.*

Proof. Optional HW. □

Remark 1.8.9. By Cantor's theorem $|\mathbb{R}| = 2^{\aleph_0}$, and we know that $\aleph_1 \leq 2^{\aleph_0}$, since $\aleph_1$ is, by definition, the smallest uncountable cardinal. It follows from the previous theorem that $2^{\aleph_0} = \aleph_\alpha$, for some α. But *which* α?

There's not too much we can prove about $2^{\aleph_0}$ without having to strengthen (whatever that means) our set theory. For instance:

FACT. *We* can *prove that* $2^{\aleph_0} \neq \aleph_\omega$.[20]

The **Continuum Hypothesis** is the statement:

$$2^{\aleph_0} = \aleph_1.$$

It turns out, that this statement is *independent* of the axioms of set theory. This will make more sense later, but let's record it here for now:

> There are models of set theory in which the continuum hypothesis holds and models of set theory in which the continuum hypothesis does not hold.

[20]But this will be done in another course.

Optional Homework