Math 434: Set Theory

Matthew Harrison-Trainor

May 10, 2019

Part I
Foundations

1 Axioms of ZFC

1.1 History

Set theory is the study of *sets*, or collections of objects. The modern study of set theory began in the 1870's with Cantor and Dedekind. They studied set theory in a naive way, without formally setting out the rules. One principal they used was that if P was a property, then there is a set $\{x \mid P(x)\}$ of all sets satisfying P. In 1901, Russell discovered the following paradox: Let R be the set $R = \{x \mid x \notin x\}$; then $R \in R$ if and only if $R \notin R$. This led to somewhat of a crisis in mathematics, as mathematicians wanted to be sure that what they were doing was consistent.

Mathematics had to be formalized; the idea was to propose a simple base theory in which all of mathematics could be constructed. Moreover, this base theory should be consistent, i.e., it should have no contradictions. The solution that became most popular was axiomatic set theory. The idea is to represent all mathematical objects—functions, groups, fields, the natural numbers, the real numbers, etc. as sets, and then to use certain *axioms*—assumptions—to prove theorems about these. In 1908 Zermelo set out various axioms, but Fraenkel pointed out in 1921 that they were not sufficient to carry out all mathematics. In 1922 Fraenkel and Skolem independently added the last few axioms. The resulting axiomatization is known as Zermelo-Fraenkel set theory, or ZF, and was published by Zermelo in 1930. Adding the Axiom of Choice leads to ZFC.

Around the same time, David Hilbert suggested the following program: Find a set of axioms in which to set mathematics, and then prove that these axioms are consistent. Unfortunately, Gödel proved (with his incompleteness theorems) that this is impossible; no sufficiently strong system can prove its own consistency. Thus, while most mathematicians believe that ZFC is consistent, we can never prove that it is.

What we can prove is relative consistency results; for example, we can prove that if ZF is consistent, then so is ZFC. So the Axiom of Choice cannot add any inconsistencies that were not already present.

1.2 The Axioms

The language of ZFC is the first-order language with a single binary relation symbol $\in$ (though of course we always have $=$ as well). We can also define more symbols such as $\cup, \cap, \varnothing$, etc. in the obvious way. Because we are in the setting of first-order logic, we have access to the completeness theorem, compactness, etc. The axioms of ZFC are:

Axiom of Extensionality: Two sets are equal if and only if they have the same elements.

$$(\forall x)(\forall y)[(\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x = y].$$

Axiom of Pairing:

$$(\forall x)(\forall y)(\exists z)[x \in z \wedge y \in z \wedge (\forall w \in z)(w = x \vee w = y)].$$

We denote the set z by $\{x, y\}$.

Axiom of Infinity: There is an inductive set.

$$(\exists x)[\varnothing \in x \wedge (\forall y \in x)(\{y\} \cup y \in x)].$$

This is the only axiom which supposes that any sets exist (without relying on some other set existing), though usually this is taken for granted in first-order logic. Note that by the Axiom of Extensionality, there is a unique empty set which we denote by $\varnothing$.

Axiom of Union: Given a set x, there is a set y that is the union of all elements of x.

$$(\forall x)(\exists y)(\forall z)[z \in y \leftrightarrow (\exists w \in x)(z \in w)].$$

We denote the set z by $\bigcup x$.

We can define further symbols like "$\subseteq$": $x \subseteq y$ just means $\forall z(z \in x \rightarrow z \in y)$.

Axiom of Powerset: Given any set x, there is a set containing all subsets of x.

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

So far, the axioms have all been a single expression. The next two will be schema: they are actually a list of axioms.

Axiom Schema of Separation: For any first-order formula $\varphi(x, z, w_1, \ldots, x_n)$,

$$\forall z \forall w_1 \cdots \forall w_n \exists y \forall x[x \in y \longleftrightarrow (x \in z \ \wedge \ \varphi)].$$

This implies that given z and a formula φ, we can find a set

$$y = \{x \in z : \varphi(x)\}.$$

Given Russell's paradox, this implies that there is no set of all sets. Note that the relation defined by φ is not necessarily a set.

Axiom Schema of Replacement: Given a first-order formula $\varphi(x, y, w_1, \ldots, w_n)$ that defines a function from x to y with parameters $w_1, \ldots, w_n$ (i.e., for all $x, w_1, \ldots, w_n$, there is a unique y with $\varphi(x, y, w_1, \ldots, w_n)$), the range of any set under this function is again a set:

$$\forall a \forall w_1 \forall w_2 \ldots \forall w_n \big[\forall x (x \in a \to \exists! y\, \varphi) \longrightarrow \exists b\ \forall x \big(x \in a \Rightarrow \exists y (y \in b \wedge \varphi)\big)\big].$$

Note that the function defined by φ is not necessarily a set.

Axiom of Foundation: Any non-empty set x has an element y with $x \cap y = \varnothing$:

$$\forall x [\exists a (a \in x) \longrightarrow \exists y (y \in x \wedge \neg \exists z (z \in y \wedge z \in x))].$$

This implies that there is no infinite sequence $x_0 \ni x_1 \ni x_2 \ni \cdots$. This axiom can be useful for doing metamathematics, but otherwise has no effect on most of mathematics.

Axiom of Choice: Given a collection x of non-empty sets, there is a function f which picks out a single element from set in x:

$$\forall x [\varnothing \notin x \quad \longrightarrow \quad \exists f \colon x \to \bigcup x \quad \forall a \in x (f(a) \in a)].$$

Sometimes we will want to talk about all of the sets satisfying a particular property, even though there might not be a single set containing all of these. We call such a collection a *class*.

Definition 1.1. Given a formula $\varphi(x)$, we call $\{x : \varphi(x)\}$ a class.

Lemma 1.2. *There are classes that are not sets, for example, the universe* $V := \{x : x = x\}$ *and* $R := \{x : x \notin x\}$

Proof. If R was a set, then $R \in R \longleftrightarrow R \notin R$. So R cannot be a set. If V was a set, then we could define R by the Axiom of Comprehension. □

1.3 Basic Constructions with Sets

From the axioms, we need to show that we can do various small things that we usually take for granted. For example, given sets X and Y, we want to show that $X \cup Y = \{z : z \in X \vee z \in Y\}$ exists. We will be pedantic as this is our first such argument. First, define

$$\text{union}(x, y, z) \quad \text{ if and only if } \quad \forall u [(u \in x \vee u \in y) \longleftrightarrow u \in z].$$

Then we need to show:

Lemma 1.3. $\forall x, y \exists! z\, \text{union}(x, y, z)$.

Proof. To prove that some such z exists, by the Axiom of Pairing there is a set $\{x, y\}$ and then $z = \bigcup\{x, y\}$ exists by the Axiom of Union. Then z is unique by the Axiom of Extensionality. □

Definition 1.4. We write $x \cup y$ for the unique set z satisfying union(x, y, z).

Similarly, we can define $x \cap y := \{u \in x : u \in y\}$ and $\bigcap x := \{u \in \bigcup u : \forall y \in x(u \in y)\}$.

Given two sets a and b, we can use the Axiom of Pairing to define the set $\{a, b\}$. This is the unordered pair containing a and b. We also want to define the ordered pair $(a, b) = \{\{a\}, \{a, b\}\}$. We will show that (a, b) exists, and prove some of its properties, on the assignment.

We also want to form the product $X \times Y = \{(x, y) : x \in X, y \in Y\}$. To do this, we will have to use the Powerset Axiom. This says that given a set x, there is a set containing all of the subsets of x; we call this the powerset of x, and write $\mathcal{P}(x)$.

Lemma 1.5. *For all sets X and Y, $X \times Y$ is a set.*

Proof. We know that $X \cup Y$ is a set. Applying the powerset axiom twice, we get a set $Z = \mathcal{P}(\mathcal{P}(X \cup Y))$. Given $x \in X$ and $y \in Y$, $\{x\}, \{x, y\} \in \mathcal{P}(a \cup b)$ and so $(x, y) \in Z$. Then by the Axiom of Separation,

$$X \times Y = \{z \in Z : (\exists x \in X)(\exists y \in Y)\ z = (x, y)\}$$

is a set. □

1.4 The Natural Numbers

We want to represent the natural numbers by sets. We start with $0 = \varnothing$ (the set $\varnothing$ exists by the Axiom of Infinity). By pairing, there is a set $\{\varnothing\}$; we call this 1. Note that $1 = \{0\}$. By pairing again, $\{1, \{1\}\}$; by union, $\{0, 1\}$ is a set. We call this set 2. In general, we define the *ordinal successor function* $S(x) = x \cup \{x\}$. Then, in general we want to say that

$$n + 1 = S(n) = \{0, 1, \ldots, n\}.$$

One feature of this definition is that $n < m$ if and only if $n \in m$. (The statement $n + 1$, while helpful, is in some sense meaningless; we are currently defining the natural numbers, whereas the left-hand-side assumes that they have already been defined.)

Each natural number is a set, but is there a set of all natural numbers? We use the Axiom of Infinity: there is a set I such containing $\varnothing$, and such that for all $x \in I$, $S(x) \in I$. We call any such set *inductive*. This set I contains all of the natural numbers we have just defined, but there is no reason that it might not contain more elements. We define $\mathbb{N}$ to be the smallest inductive set:

Definition 1.6. We define the natural numbers

$$\mathbb{N} = \{x : \text{for all inductive sets } J,\ x \in J]\}.$$

To show that this exists, we note that:

$$\mathbb{N} = \{x \in I : \text{for all inductive sets } J,\ x \in J]\}$$

and use the Axiom of Separation.

It is not hard to show that $\mathbb{N}$ is the smallest inductive set. We call the elements of $\mathbb{N}$ natural numbers. One of the most important properties of the natural numbers is induction.

Theorem 1.7 (Induction)**.** *For any non-empty inductive set $X \subseteq \mathbb{N}$, $X = \mathbb{N}$.*

Proof. If X is inductive then by definition of $\mathbb{N}$, $\mathbb{N} \subseteq X$, and so $X = \mathbb{N}$. □

Using induction, to prove $\forall x \varphi(x)$, it suffices to prove $\varphi(\varnothing)$ and $\forall x \varphi(x) \to \varphi(S(x))$.

1.5 Functions as Sets

Like all mathematical objects, we want to represent functions as sets.

Definition 1.8. A set R is a (binary) relation if R is a set of ordered pairs. We write xRy for $(x, y) \in R$.

Definition 1.9. A set f is a *(set) function* if it is a binary relation, and for every x, there is most one y such that $(x, y) \in f$. If $\exists y[(x, y) \in f]$, then we denote by $f(x)$ the unique such y.

If f is a function, we define:

- the domain of f is $\mathrm{dom}(f) = \{x : \exists y (x, y) \in f\}$.
- the range of f is $\mathrm{ran}(f) = \{y : \exists x (x, y) \in f\}$.

Of course, we must prove that these are sets.

The Axiom Schema of Replacement also involves functions, but those functions are defined by formulas and may not be sets. We call such functions class functions to distinguish them from (set) functions.

We return to the natural numbers to define addition and multiplication as functions. First, we show that the successor operator is a function.

Lemma 1.10. *The successor function $s: \mathbb{N} \to \mathbb{N}$ defined by $s(n) = \{n\} \cup n$ is a set.*

Proof. The function s consists of the set of all pairs $(n, m) \in \mathbb{N} \times \mathbb{N}$ such that m has two elements, one of which is n, and the other of which is a set containing a single element n. This can be expressed in first-order logic, and so s is a set by the Axiom of Separation. □

Definition 1.11. Define a function $f(n, m) = n + m$ from $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ as follow. Let A be the set of all partial functions $g: \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that (1) for all n, $g(n, 0) = n$ and (2) whenever $g(n, s(m))$ is defined, $g(n, m)$ is defined and $g(n, s(m)) = s(g(n, m))$. Now define $f = \bigcup A$.

A partial function $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is just a function whose domain is a subset of $\mathbb{N} \times \mathbb{N}$. We must verify that this definition works.

Lemma 1.12. *As defined above, $+$ is a function defined on all of $\mathbb{N} \times \mathbb{N}$.*

Proof. First, we show that + is a (partial) function: given $g, h \in A$, and $n, m \in \mathbb{N}$, if $g(n,m)$ and $h(n,m)$ are defined, then $g(n,m) = h(n,m)$. Fix n. Let $E \subseteq \mathbb{N}$ be the set of integers m such that if $g(n,m)$ and $h(n,m)$ are defined, then $g(n,m) = h(n,m)$. We argue by induction that $E = \mathbb{N}$. First, $0 \in E$ as $g(n,0) = h(n,0) = n$. Second, suppose that $m \in E$; if $g(n,s(m))$ and $h(n,s(m))$ are defined, then $g(n,s(m)) = s(g(n,m)) = s(h(n,m)) = h(n,s(m))$. So $s(m) \in E$.

Second, we show that the domain of f is $\mathbb{N} \times \mathbb{N}$. Fix n. Given $m \in \mathbb{N}$, we must show that there is $g \in A$ such that $g(n,m)$ is defined. Let $F \subseteq \mathbb{N}$ be the set of $m \in \mathbb{N}$ such that $g(n,m)$ is defined. We argue by induction that $F = \mathbb{N}$. We first show that $0 \in F$, i.e. that there is a function defined on 0, for example $\{((0,0),0)\}$ (one must, of course, argue that this is a set). Second, suppose that $m \in F$, so that there is $g \in A$ with $g(n,m)$ defined. If $g(n,s(m))$ is already defined, then $s(m) \in F$. Otherwise,

$$g \cup \{((n,s(m)), s(g(n,m)))\}$$

is a partial function in A defined on $(n,s(m))$. So if $s(m) \in F$, and $F = \mathbb{N}$. □

Lemma 1.13. *Addition + is the unique function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ satisfying $f(n,0) = n$ and $f(n,s(m)) = s(f(n,m))$.*

Proof. Exercise. □

2 Well-orders and Ordinals

2.1 Well-orders

Natural numbers also have an ordering. You may have seen before that induction on $\mathbb{N}$ is essentially the same as the fact that any subset of $\mathbb{N}$ has a least element. This property of an ordering is known as being *well-ordered*. In this section, we will extend the natural numbers to *ordinals*, which essentially keep counting "past infinity".

Definition 2.1. Let A be a set and $\leq$ a binary relation on A. Then $\leq$ is a partial order if it is:

(Reflexive) For all $x \in A$, $x \leq x$.

(Transitive) For all $x, y, z \in A$, if $x \leq y$ and $y \leq z$ then $x \leq z$.

(Anti-symmetric) For all $x, y \in A$, if $x \leq y$ and $y \leq x$ then $x = y$.

The partial order $\leq$ is a total order (or linear order) if it is:

(Total) For all $x, y \in A$, $x \leq y$ or $y \leq x$.

To a partial order $\leq$ there is an associated *strict* linear order $<$ defined by $x < y$ if and only if $x \leq y$ and $x \neq y$.

Definition 2.2. A partial order $\leq$ on A is well-founded if every non-empty set has a least element: for every set $E \subseteq A$, there is $x \in E$ such that for all $y \in E$, $x \leq y$. We usually say that a total order which is well-founded is a well-order.

The prototypical examples of well-orders are the finite linear orders and $\mathbb{N}$ (though we have not actually put a linear order on this yet). But there are other examples, for example $\mathbb{N} \cup \{\infty\}$ where ∞ is larger than each $n \in \mathbb{N}$. $\mathbb{Z}$ and $\mathbb{R}$ are not well-orders; e.g. $\mathbb{Z}$ itself has no least element. The interval $[0,1] \subseteq \mathbb{R}$ has a least element, but it is still not a well-order; $(0,1]$ has no least element.

Lemma 2.3. *If $(A, \leq)$ is a well-order, and $B \subseteq A$, then $(B, \leq)$ is also a well-order.*

Proof. Exercise. □

Another useful characterization of well-orders is as follows:

Lemma 2.4. *A partial order $(A, \leq)$ is well-founded if and only if it has no infinite strictly descending sequence $a_0 > a_1 > a_2 > \cdots$.*

Formally, a sequence is a function with domain $\mathbb{N}$.

Proof. Suppose that there is a such a descending sequence $a_0 > a_1 > a_2 > \cdots$. Then the set $\{a_i : i \in \mathbb{N}\}$ has no least element, so $(A, \leq)$ is not well-ordered.

On the other hand, suppose that there is a non-empty set $E \subseteq A$ that has no least element. Pick $a_0 \in E$. Since a_0 is not the least element in E, there is $a_1 < a_0$ in E. Then a_1 is not the least element of E so there is $a_2 < a_1$ in E, and so on. (Formally, we should be defining this sequence the same way we defined addition. Later we will talk more about inductive definitions.) □

To compare linear orders, we have a notion of isomorphism.

Definition 2.5. Let $(A, \leq_A)$ and $(B, \leq_B)$ be two linear orders. An order-isomorphism between A and B is a bijection $f : A \to B$ such that

$$a \leq_A b \iff f(a) \leq_B f(b).$$

As usual, two linear orders which are isomorphic can be viewed as essentially the same.

We can prove a few very useful facts about isomorphisms between well-ordered sets. These facts will all be important for using ordinals to "count".

Lemma 2.6. *Let $(A, \leq_A)$ and $(B, \leq_B)$ be well-orders. Suppose that $f, g : A \to B$ are order isomorphisms. Then $f = g$.*

Proof. Let $E \subseteq A$ be the set of all $x \in A$ such that $f(x) \neq h(x)$. We will argue that E is empty. If E was non-empty, then since A is a well-order, E has a least element a. Let $A\restriction_a = \{x \in A : x < a\}$. Then since f is an order isomorphism, $f(a)$ must be the least element b_1 of $B - f(A\restriction_a)$. Similarly, $g(a)$ must be the least element b_2 of $B - g(A\restriction_a)$. Since $f(A\restriction_a) = g(A\restriction_a)$, $b_1 = b_2$ and so $f(a) = g(a)$. □

Corollary 2.7. *Well-orders are rigid: the only isomorphism from a well-order into itself is the identity.*

Proof. Exercise. □

Lemma 2.8. *Let $(A, \leq)$ be a well-order. Then A is not isomorphic to any proper initial segment of itself.*

Proof. Exercise. □

Lemma 2.9. *Let $(A, \leq_A)$ and $(B, \leq_B)$ be well-orders. Then either A and B are isomorphic, or one is isomorphic to a proper initial segment of the other.*

Proof. Given a linear order $(A, \leq_A)$, we write $A\restriction_a$ for $\{x \in A : x <_A a\}$. Note that every initial segment of a well-order is of the form $A\restriction_a$ for some A; indeed, a is the least element not in the initial segment.

Define a partial function $f\colon A \to B$ by $f(a) = b$ if and only if $A\restriction_a$ is isomorphic to $B\restriction_b$. This defines a partial function because by the previous lemma, given $a \in A$, $A\restriction_a$ cannot be isomorphic to both $B\restriction_b$ and $B\restriction_c$ for $b \neq c$.

Suppose that $a_1 < a_2$ and that a_2 is in the domain of f. Let $b_2 = f(a_2)$, so that $A\restriction_{a_2}$ is isomorphic to $B\restriction_{b_2}$. Then $A\restriction_{a_1}$ is isomorphic to an initial segment of $B\restriction_{b_2}$; and this initial segment must be proper. Let $b_1 \in B\restriction_{b_2}$ be such that $A\restriction_{a_1}$ is isomorphic to $B\restriction_{b_1}$. So $f(a_1) = b_1 < b_2 = f(a_2)$. This shows that (1) the domain of f is an initial segment of A, and (2) f is an order isomorphism between its domain and its range. A similar argument shows that the range of f is an initial segment of B.

We will argue that either the domain of f is all of A or the range of f is all of B. This suffices to complete the proof of the lemma. So suppose to the contrary that the domain of f is not all of A, and the range of f is not all of B. Let a be the least element of $A - \operatorname{dom} f$ and let b be the least element of $B - \operatorname{ran} f$. Then f is an isomorphism between $A\restriction_a$ and $B\restriction_b$, and so $f(a) = b$, a contradiction. □

Exercise 2.10. Let $(A, \leq_A)$ and $(B, \leq_B)$ be linear orders. Define the lexicographic order $\leq_{lex}$ on $A \times B$ as follows:

$$(a, b) \leq_{lex} (a', b') \iff a <_A a', \text{ or } a = a' \text{ and } b \leq_B b'.$$

Then $\leq_{lex}$ is a linear order. If A and B are well-orders, then so is $\leq_{lex}$.

Lemma 2.11. *The Axiom of Foundation implies that $\in$ is well-founded.*

Note that the domain of $\in$ is not a set. When we say that it is well-founded, we mean that there is no descending sequence $a_0 \ni a_1 \ni a_2 \ni \cdots$, or equivalently, that each *set* has an $\in$-least element.

Proof. Suppose that we had a descending sequence $a_0 \ni a_1 \ni a_2 \ni \cdots$. Then apply the Axiom of Foundation to the set

$$A = \{a_0, a_1, a_2, \ldots\}.$$

Some a_i must have $a_i \cap A = \varnothing$, but $a_{i+1} \in a_i \cap A$. This is a contradiction. □

2.2 Ordinals

Natural numbers can be used to count sizes (one sheep, two sheep, ...) or order (first sheep, second sheep, ...). We call one, two, three, ... cardinal numbers and first, second, third, ... ordinal numbers. Sometimes it is useful to keep "counting past infinity". For this, we introduce ordinals. The idea is that we want to be able to do induction along ordinals, and for this we will make sure that they share with $\mathbb{N}$ the property that every set of ordinals has a least element.

Definition 2.12. A set x is called transitive if for all $y \in x$, $y \subseteq x$.

Basically the relation $\in$ is transitive, so that if $z \in y \in x$, then $z \in x$.

Definition 2.13. A set is an ordinal if it is transitive and well-ordered by $\in$.

For example, $3 = \{0, 1, 2\}$ is an ordinal. We usually use Greek letters $\alpha, \beta, \gamma, \ldots$ for ordinals.

Lemma 2.14. *Using the Axiom of Foundation, α is an ordinal if and only if α is a transitive set and α is totally ordered by $\in$.*

Proof. We proved that $\in$ is well-founded on $V = \{x : x = x\}$. So it is still well-founded when we restrict to α. □

Because in the definition of an ordinal we ask that $\in$ be well-founded, even without the Axiom of Foundation we get facts like if α is an ordinal, the $\alpha \notin \alpha$. So we can develop the whole theory of ordinals without the Axiom of Foundation.

Lemma 2.15.

1. *$\varnothing$ is an ordinal.*
2. *If α is an ordinal, then $s(\alpha)$ is an ordinal.*
3. *If α is an ordinal and $\beta \in \alpha$, then β is an ordinal.*
4. *If α, β are ordinals and $\alpha \subsetneq \beta$, then $\alpha \in \beta$.*
5. *If α, β are ordinals then $\alpha \cap \beta$ is an ordinal.*
6. *If α, β are ordinals then $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.*
7. *If α, β are ordinals, then $\alpha \cup \beta$ is an ordinal.*
8. *$s(\alpha) \notin \alpha$.*
9. *$\gamma < s(\alpha)$ if and only if $\gamma \leq \alpha$.*

Proof. (1) This is trivial.

(2) We must show that $s(\alpha) = \alpha \cup \{\alpha\}$ is transitive, and that it is well-ordered by $\in$. If $x \in s(\alpha)$, then either $x \in \alpha$ and so $x \subseteq \alpha \subseteq s(\alpha)$, or $x = \alpha \in s(\alpha)$. So α is transitive.

Now $\in$ linearly orders $s(\alpha)$ with α as the largest element; this is because each $x \in s(\alpha)$ other than α is in α. Any $\in$-descending sequence in $s(\alpha)$ must have almost all of its terms in α (all but one), and there are no such descending sequences in α as it is an ordinal. So there $\in$ well-orders $s(\alpha)$.

(3) If $\beta \in \alpha$, then $\beta \subseteq \alpha$. So $\in$ well-orders α (see Lemma 2.3). Also, β is transitive: if $y \in x \in \beta$, then $x, y \in \alpha$. Since $\in$ linearly orders α, either $y \in \beta$ or $\beta \in y$. But we cannot have $\beta \in y \in x \in \beta$, so $y \in \beta$.

(4) Let $\delta \in \beta$ be least such that $\delta \in \beta - \alpha$. We claim that $\delta = \alpha$, so that $\alpha \in \beta$.

Given $x \in \alpha$, either $x \in \delta$ or $\delta \in x$; the latter would imply that $\delta \in \alpha$, which is not the case. So $x \in \delta$. This implies that $\alpha \subseteq \delta$.

Now suppose that there was $x \in \delta - \alpha$. Then $x \in \beta$, and this would contradict the choice of δ as the least element of $\beta - \alpha$. So $\delta \subseteq \alpha$.

(5) $\alpha \cap \beta$ is well-ordered by $\in$ (Lemma 2.3), and it is transitive because α and β are transitive (if $x \in y \in \delta = \alpha \cap \beta$, then $x \in \alpha$ because α is transitive, and $x \in \beta$ because β is transitive).

(6) Let $\delta = \alpha \cap \beta$. This is an ordinal. If $\delta = \alpha$ or $\delta = \beta$, then we are done. So assume that $\delta \neq \alpha$ and $\delta \neq \beta$. Then $\delta \subsetneq \alpha$, and so $\delta \in \alpha$; also $\delta \in \beta$. Thus $\delta \in \alpha \cap \beta = \delta$, a contradiction.

(7) Follows immediately from the previous parts.

(8) If $s(\alpha) \in \alpha$, then $\alpha \in \alpha$. But then ϵ does not well-order α.

(9) This is just saying that $\gamma \in s(\alpha)$ if and only if $\gamma = \alpha$ or $\gamma \in \alpha$. □

So we now have lots of examples of ordinals: 0, 1, 2, 3, ... In fact we can prove (by induction) that every element of $\mathbb{N}$ is an ordinal. But there are lots of other, infinite, ordinals. We will soon prove that $\mathbb{N}$ is an ordinal. But first we will take a slight detour.

The property of being an ordinal is described by a first-order formula $\varphi(x)$. Remember that this does not necessarily mean that there is a set of all x satisfying $\varphi(x)$; in fact, we will see that there is no set of all ordinals. But it is still convenient to think about the collection $\{x : x \text{ is an ordinal}\}$. Recall that we call such a collection a class.

Definition 2.16. Let ON be the class of all ordinals. Define $\alpha < \beta$ if and only if $\alpha \in \beta$, and $\alpha \leq \beta$ if and only if $\alpha \in \beta$ or $\alpha = \beta$.

We will see that ON is a proper class, that is, that there is not a set of all ordinals. Nevertheless, it will be convenient to write things like $\alpha \in ON$, which really means $\varphi(\alpha)$ where φ is the formula expressing that α is an ordinal.

Theorem 2.17. *ON is (strictly) well-ordered by $\in$. That is:*

1. *$\in$ is transitive on the ordinals,*
2. *$\in$ is irreflexive on the ordinals,*
3. *$\in$ satisfies the trichotomy: for all $\alpha, \beta \in ON$, $\alpha \in \beta$ or $\beta \in \alpha$ or $\alpha = \beta$.*

4. $\in$ is well-founded on the ordinals: every non-empty class of ordinals has a least member.

Proof. We've already seen that $\in$ is transitive, irreflexive, and satisfies the trichotomy. Now we will show that it is well-founded. Let C be a non-empty class of ordinals. Pick $\alpha \in C$. Then $\alpha \cap C = \{\beta \in \alpha : \beta \in C\}$ is a set (by the Axiom of Separation), and it has a least element β. Suppose that for some $\gamma \in C$, $\gamma \in \beta$. Then $\gamma \in \alpha$ since α is transitive, and so $\gamma \in \alpha \cap C$. This cannot happen. That means that β is the $\in$-least element of C. □

Corollary 2.18. *If X is a set of ordinals that is transitive, then X is an ordinal.*

Proof. X is well-ordered by $\in$. □

We can use this corollary to easily show that $\mathbb{N}$ is an ordinal.

Lemma 2.19. *The union or intersection of a set of ordinals is an ordinal. Hence $\mathbb{N}$ is an ordinal.*

Proof. Let A be a set of ordinals. It is easy to see that $\bigcup A$ and $\bigcap A$ are transitive. Thus $\bigcup A$ and $\bigcap A$ are ordinals. □

When we refer to $\mathbb{N}$ as an ordinal, we call it ω. Now we also have an ordering on $\mathbb{N}$. This is the first ordinal we know that is not obtained by starting with 0 and applying the successor function. As such, we call it a *limit ordinal*.

Definition 2.20. If an ordinal α is the successor of some ordinal β, we call α a successor ordinal. Otherwise, if $\alpha \neq 0$, we call α a limit ordinal. 0 is neither a successor nor a limit ordinal.

From ω, we can take $s(\omega)$ to get a larger ordinal, $s(s(\omega))$, and so on; and then we can take the limit of these, and keep going. In fact, we can keep going so much that there is no set of all ordinals.

Theorem 2.21. *ON is a proper class: there is no set of all ordinals.*

Proof. If there was a set of all ordinals, say x, then $\bigcup x$ is an ordinal, and so $\bigcup x \in x$. We will argue that then $\bigcup x \in \bigcup x$, a contradiction as $\bigcup x$ is an ordinal.

If $z \in \bigcup x$, then z is an ordinal, and so $z \in x$.

Also, if $z \in x$ then z is an ordinal, so $z \in s(z)$ is an ordinal, and so $s(z) = z \cup \{z\} \in x$. Then $z \in \bigcup x$.

So $x = \bigcup x$, and $\bigcup x \in \bigcup x$. □

Finally, we will show that ordinals are in some sense the canonical well-orders, in the sense that every well-order has the same order type (i.e., is isomorphic to) an ordinal. Moreover, any two ordinals are not order-isomorphic to each other.

Lemma 2.22. *No two distinct ordinals are order-isomorphic to each other.*

Proof. Given two ordinals α, β, either $\alpha \subsetneq \beta$, $\beta \subseteq \alpha$, or $\alpha = \beta$. Suppose that $\alpha \subsetneq \beta$; we will show that α is a proper initial segment of β, and thus cannot be isomorphic to β (by Lemma 2.9). Indeed, if $x \in \beta$ and $y \in \alpha$ have $x \in y$, then since α is transitive, $x \in \alpha$. □

Theorem 2.23. *Every well-ordered set is isomorphic to a unique ordinal.*

Proof. Let $(W, \leq)$ be a well-ordering. We call x "good" if the theorem holds for $W\restriction_x$, i.e., there is an ordinal α such that $(W\restriction_x, \leq) \cong (\alpha, \in)$. Note that by the previous lemma, there is at most one such α for each x.

Let G be the set of good elements. Let $f : G \to ON$ be the class function such that $f(x)$ is the unique α witnessing that x is good. The range of this function exists by the Axiom of Replacement, and then this function exists (as a set) by the Axiom of Separation.

Note that $\beta = \text{ran}(f)$ is an initial segment of ON, and so is itself an ordinal (it is transitive and then we use Corollary 2.18). If $G = W$, then f is an isomorphism between W and β. Otherwise, if $G \neq W$, then let w be the least element of $W - G$. Then $W\restriction_w = G$, and G is isomorphic to β, and so $w \in G$ which is a contradiction. □

Given an ordering $(W, \leq)$, we denote by $ot(W)$ the order type of W: the ordinal α such that $W \cong \alpha$.

2.3 Ordinal Arithmetic

Just like natural numbers, we can add and multiply ordinals.

Definition 2.24. Define addition by $\alpha + \beta = ot((\{0\} \times \alpha) \cup (\{1\} \times \beta), \leq_{lex})$ where $\leq_{lex}$ is the lexicographic order.

Definition 2.25. Define multiplication by $\alpha \cdot \beta = ot(\beta \times \alpha, \leq_{lex})$.

These definitions would not be well-defined if these lexicographic orders were not well-orders. We proved on the homework that $(\beta \times \alpha, \leq_{lex})$ is a well-order. A similar proof shows that $((\{0\} \times \alpha) \cup (\{1\} \times \beta), \leq_{lex})$ is a well-order.

One can think of $\alpha + \beta$ as the order which takes a copy of α, and makes it smaller than a copy β. Think of $\alpha \cdot \beta$ as taking β-many copies of α and putting them one after the other. Note that $+$ and $\cdot$ are not commutative; for example, $1 + \omega = \omega \neq \omega + 1$ and $2 \cdot \omega = \omega \neq \omega \cdot 2$.

Lemma 2.26. *Ordinal addition and multiplication are associative.*

Proof. Exercise. □

Some other easy properties are:

1. $s(\alpha) = \alpha + 1$.
2. Left cancellation for addition: $\alpha + \beta = \alpha + \gamma \to \beta = \gamma$.
3. Left cancellation for multipliation: $\alpha \cdot \beta = \alpha \cdot \gamma \to \alpha = 0$ or $\beta = \gamma$.
4. Subtraction: $\alpha \leq \beta \to \exists! \gamma (\alpha + \gamma = \beta)$.

2.4 Limits

Let X be a set of ordinals. Then

$$\alpha = \bigcup X$$

is an ordinal. Moreover, $\alpha = \sup_{\beta \in X} \beta$. To see this, note that each $\beta \subseteq \alpha$, so that $\beta \leq \alpha$; and each $\gamma \in \alpha$ has some $\beta \in X$ with $\gamma \in \beta$.

We get another characterization of addition:

Lemma 2.27. *If α is an ordinal and γ is a limit ordinal,*

$$\alpha + \gamma = \sup_{\beta < \gamma} \alpha + \beta.$$

Proof. If $\beta < \gamma$, then $(\{0\} \times \alpha) \cup (\{1\} \times \beta)$ is an initial segment of $(\{0\} \times \alpha) \cup (\{1\} \times \gamma)$, so $\alpha + \beta \leq \alpha + \gamma$.

Suppose that $\delta < \alpha + \gamma$. If $\delta < \alpha$, then $\delta < \alpha + \beta$ for any $\beta < \gamma$. Otherwise, by "subtraction", there is β such that $\delta = \alpha + \beta$. We have $\beta < \gamma$, and so $\delta \leq \alpha + \beta$. □

In the next section, we will show how we could actually take this as our definition of addition.

2.5 Transfinite Induction and Recursion

We have already used the following technique in a few proofs:

Lemma 2.28 (Transfinite Induction on ON, version one)**.** *If X is a class of ordinals with the property that whenever α is an ordinal and every $\beta < \alpha$ is in X, then $\alpha \in X$, then $X = ON$.*

We usually think of X being the set of ordinals with some property. Often, the way one applies this is as follows:

Corollary 2.29 (Transfinite Induction on ON, version two)**.** *If X is a class of ordinals with the proprty that*

1. *$0 \in X$,*
2. *whenever $\alpha \in X$, $s(\alpha) \in X$, and*
3. *whenever γ is a limit ordinal and every $\alpha < \gamma$ is in X, $\gamma \in X$,*

Then $X = ON$.

This is because it often turns out to be useful to separate the successor and limit cases as they are often handled differently.

We can also do induction on any well-ordered set:

Lemma 2.30 (Transfinite Induction on well-ordered sets)**.** *Let $(W, \leq)$ be a well-ordered set. If $X \subseteq W$ has the property that whenever $x \in W$ and every $y < x$ is in X, then $x \in X$, then $X = W$.*

However, sometimes we want to make an inductive definition, such as the original definition we had for addition on $\mathbb{N}$. We can prove a metatheorem saying that we can do this. Recall that V is the class of all sets.

Theorem 2.31. *Let $F: V \to V$ be a class function. There is a unique class function $G: ORD \to V$ such that for all ordinals α*

1. *$G{\restriction_\alpha}$, the function G restricted to the domain α, is a set.*
2. *$G(\alpha) = F(G{\restriction_\alpha})$.*

You should think of F as giving a way of extending a function on an initial segment of the ordinals by adding one more value to the domain. The theorem says that there is a class function satisfying this recursive definition, and that this class function is approximated by sets.

Proof. The proof is going to be somewhat similar to when we showed that the definition of addition on $\mathbb{N}$ made sense. Given an ordinal δ, we call a set function h a δ-approximation if $\operatorname{dom} h = \delta$ and h satisfies the recursive definition, i.e., if $h(\alpha) = F(h{\restriction_\alpha})$ for all $\alpha \in \delta$.

First, we show that whenever g is a δ-approximation and h is a γ-approximation, and $\alpha \in \gamma \cap \delta$, $g(\alpha) = h(\alpha)$. In particular, there is only one δ-approximation for each δ. We argue by induction, using the fact that if $g{\restriction_\alpha} = h{\restriction_\alpha}$, then $g(\alpha) = F(g{\restriction_\alpha}) = F(h{\restriction_\alpha}) = h(\alpha)$.

Second, we show that for every δ, there is a δ-approximation. This is again an inductive argument. Here is where it is convenient to split between the limit and successor case. If δ is a successor, let α be such that $\delta = \alpha + 1$ and let g be an α-approximation. Then define $h \supseteq g$ by $h(\alpha) = F(g)$; h is a δ-approximation. If δ is a limit ordinal, then for every $\alpha < \delta$, there is an α-approximation h_α. Then $\bigcup_{\alpha<\delta} h_\alpha$ is a δ-approximation. (We use the Axiom of Replacement here in order to take this union.)

Now define the class G as follows: (α, y) is in G if α is an ordinal and there is $\beta > \alpha$ and a β-approximation h with $h(\alpha) = y$. For any α, $G{\restriction_\alpha}$ is just an (or the) α-approximation. This gives (1). (2) follows because the α-approximations satisfy (2). □

Similarly to induction, sometimes we want to define $G(0) = F(\varnothing)$, $G(\alpha + 1) = F(G{\restriction_{\alpha+1}})$, and $G(\gamma) = F(\bigcup_{\alpha<\gamma} G{\restriction_\alpha})$. We can now definition addition of ordinals recursively.

Definition 2.32. Fix α. Define $\alpha + \beta$ by:

1. $\alpha + 0 = \alpha$,
2. $\alpha + (\beta + 1) = (\alpha + \beta) + 1$,
3. $\alpha + \gamma = \sup_{\beta<\gamma} \alpha + \beta$ for γ a limit ordinal.

Definition 2.33. Fix α. Define $\alpha \cdot \beta$ by:

1. $\alpha \cdot 0 = 0$,
2. $\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$,
3. $\alpha \cdot \gamma = \sup_{\beta<\gamma} \alpha \cdot \beta$ for γ a limit ordinal.

Definition 2.34. Fix α. Define α^β by:

1. $\alpha^0 = 1$,
2. $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$,
3. $\alpha^\gamma = \sup_{\beta<\gamma} \alpha^\beta$ for γ a limit ordinal.

Lemma 2.35. *These definitions of addition and multiplication agree with the ones defined before.*

Proof. This is an easy induction argument. □

2.6 The Well-ordering Theorem

Well-orderings are not very useful if we cannot find them in nature. Often when dealing with a countable set A, we find it helpful to list it out as a sequence $A = \{a_0, a_1, a_2, \ldots\}$. With ordinals, we can do the same thing with any set. This argument uses the Axiom of Choice.

Theorem 2.36. *For ever set A, there is an ordinal α and a bijection $f : \alpha \to E$.*

Proof. Intuitively, what we want to do is choose an ordinal-indexed sequence of elements $a_0, a_1, a_2, \ldots$ such that $a_\alpha \in A - \{a_\beta : \beta < \alpha\}$. Then, finding the least γ such that a_γ is undefined, we can define $f : \gamma \to A$ by $f(\alpha) = a_\alpha$.

Using the Axiom of Choice, let c be a choice function on $\mathcal{P}(A) - \{\varnothing\}$: for each non-empty subset $X \subseteq A$, $c(X) \in X$. Define a class function $G : ON \to A$ by transfinite recursion:

$$G(\alpha) = \begin{cases} c(A - \operatorname{ran} G\restriction_\alpha) & A - \operatorname{ran} G\restriction_\alpha \neq \varnothing \\ \star & \text{otherwise} \end{cases}.$$

where $\star$ is some new symbol (e.g., a set not in A).

Claim 1. *Whenever $\alpha \neq \beta$ and $G(\alpha), G(\beta) \in A$, $G(\alpha) \neq G(\beta)$.*

Proof. Say $\alpha < \beta$. Then $G(\beta) = c(A - \operatorname{ran} G\restriction_\alpha) \neq G(\alpha)$. □

Claim 2. *$G^{-1}(A)$ is a set.*

Proof. Define a partial class function $H : A \to ON$ by $H(a)$ is the ordinal α, if it exists, such that $G(\alpha) = a$. Then $\operatorname{ran} H = G^{-1}(A)$ is a set by the Axiom of Replacement. □

Claim 3. *$G^{-1}(A)$ is an initial segment of ON.*

Proof. Suppose that $\alpha < \beta \in G^{-1}(A)$. Then $A - \operatorname{ran} G\restriction_\beta$ is non-empty, and is contained in $A - \operatorname{ran} G\restriction_\alpha$. □

Since ON is a proper class, there must be some least δ such that $G(\delta) = \star$. Then $G\restriction_\delta$ is a bijection between δ and A. □

Corollary 2.37 (The Well-ordering Theorem)**.** *Every set can be well-ordered.*

You may have seen that every vector space has a basis. We can give a constructive-looking proof of this fact using the well-ordering theorem. This sort of argument can be used all over the place, whenever we want to do one thing at a time over and over transfinitely.

Theorem 2.38. *Every vector space has a basis.*

Proof. Let V be a vector space. Using the well-ordering theorem, for some ordinal γ there is a bijection between V and γ; so we can write V as $V = \{v_\alpha : \alpha < \gamma\}$.

We will construct a basis for V by transfinite recursion. Define sets $B_\alpha \subseteq V$ for $\alpha \leq \gamma$ as follows. Define $B_0 = \varnothing$ and $B_\lambda = \bigcup_{\delta<\lambda} B_\delta$. The important case is the successor case. Define $B_{\alpha+1} = B_\alpha$ if $v_\alpha \in \text{span}B_\alpha$, and $B_{\alpha+1} = B_\alpha \cup \{v_\alpha\}$ otherwise.

Let $B = B_\gamma$. To argue that B is a basis for V, we can prove the following claims by induction:

Claim 1. *For each $\alpha \leq \gamma$, B_α is linearly independent.*

Claim 2. *Whenever $\beta \leq \alpha \leq \gamma$, $G(\beta) \subseteq G(\alpha)$.*

Claim 3. *For each $\beta < \alpha < \gamma$, $v_\beta \in \text{span}(G(\beta+1))$.*

Together, these claims imply that B is linearly independent and spans V. □

Lots of mathematician do not know about ordinals, so this kind of argument can be nicely packaged into Zorn's Lemma.

Definition 2.39. Let $(P, \leq)$ be a partial order. A chain in P is a set $X \subseteq P$ such that for all $x, y \in X$, $x \leq y$ or $y \leq x$. A chain X has an upper bound if there is $b \in P$ such that for all $x \in X$, $x \leq b$.

Definition 2.40. Let $(P, \leq)$ be a partial order. An element $x \in P$ is maximal if for all $y \in P$, $y \not> x$.

Theorem 2.41 (Zorn's Lemma)**.** *Let $(P, \leq)$ be a partial order such that every chain has an upper bound. Then there is a maximal element of P.*

Proof. Exercise. □

To use Zorn's Lemma to prove that every vector space V has a basis, P is going to be the set of linearly independent subsets of V, ordered by inclusion.

Zorn's Lemma, the Well-ordering Theorem, and the Axiom of Choice are all equivalent over ZF.

Theorem 2.42. *Over ZF, the follwoing are equivalent:*

1. *The Axiom of Choice.*
2. *Any set can be well-ordered.*
3. *Zorn's Lemma.*

Proof. We have already proved (1) implies (2). (1)+(2) imply (3) is the proof of Zorn's Lemma. To see that (2) implies (1), we put a well-ordering on everything and our choice function just chooses the least element from each set.

So what is left is that (3) implies (1). Fix a set X. Let $(P, \leq)$ be the set of all possibly partial functions $f : \mathcal{P}(X) - \{\varnothing\} \to X$ such that $f(z) \in z$ whenever f is defined on z. Write $f \leq g$ if g extends f as a function. Every chain of such functions has an upper bound, namely its union. Moreover, a maximal element in the partial order is a total function (as each non-total function can be extended by adding a single element to its domain). So by Zorn's Lemma, there is a choice function $c : \mathcal{P}(X) - \{\varnothing\} \to X$ □

3 Cardinals

3.1 Comparing Sizes of Sets

In this section we'll talk about counting the sizes of sets. One way to say that one sets has more elements than another is as follows:

Definition 3.1. Let X and Y be sets. We say that the cardinality of X is at most that of Y, and write $X \precsim Y$, if there is an injection from X to Y. We say that the cardinality of X is the same as Y, and write $X \approx Y$, if there is a bijection between X and Y.

Lemma 3.2. *If X is non-empty, then $X \precsim Y$ if and only if there is a surjection from Y onto X.*

Proof. Let $f : X \to Y$ be an injection. Define g on Y by $g(y) = f^{-1}(y)$ when this is defined, and $g(y) = x_0$ otherwise for some fixed x_0. Then g is a surjection.

On the other hand, let $f : Y \to X$ be a surjection. Define $g : X \to Y$ by letting $g(x)$ be some $y \in Y$ with $f(y) = x$. This is injective. □

For this definition to make sense, we would want to know that $X \approx Y$ if and only if $X \precsim Y$ and $Y \precsim X$. This is the case:

Theorem 3.3 (Schröder-Bernstein)**.** *If there is an injection from X to Y, and an injection from Y to X, then there is a bijection between X and Y.*

Proof. Let $f : X \to Y$ and $g : Y \to X$ be injections. Call an element $y \in Y$ lonely if there is no $x \in X$ such that $f(x) = y$. Say that an element $y \in Y$ is a descendent of $y' \in Y$ if there is n such that $y = (f \circ g)^n(y')$.

Define a function $h : X \to Y$ by

$$h(x) = \begin{cases} g^{-1}(x) & \text{if } f(x) \text{ is the descendent of a lonely element,} \\ f(x) & \text{otherwise.} \end{cases}$$

This function is well-defined, as if $f(x)$ is the descendent of a lonely point y, then $f(x) = (f \circ g)^n(y)$ for some lonely point y; and $f(x)$ is not lonely, so $n \geq 1$, and $x = g \circ (f \circ g)^{n-1}(y)$, so that $g^{-1}(x)$ exists.

To see that h is surjective, let $y \in Y$. If y is the descendent of a lonely point, then $f(g(y))$ is the descendent of a lonely point, and so $y = h(g(y))$. If y is not the descendent of a lonely point, then in particular it is not lonely, and there is $x \in X$ with $f(x) = h(x) = y$.

Now we need to argue that h is injective. Suppose that $h(x_1) = h(x_2)$. If $f(x_1)$ is the descendent of a lonely element, then since $h(x_1) = h(x_2)$, we have $g^{-1}(x_1) = f(x_2)$ and so $f(x_1) = f \circ g \circ f(x_2)$. Thus $f(x_2)$ is also the descendent of a lonely element.

If both $f(x_1)$ and $f(x_2)$ are the descendent of a lonely element, then $g^{-1}(x_1) = h(x_1) = h(x_2) = g^{-1}(x_2)$ and so $x_1 = x_2$; and if neither $f(x_1)$ and $f(x_2)$ are the descendent of a lonely element, then $f(x_1) = h(x_1) = h(x_2) = f(x_2)$ and so $x_1 = x_2$. □

We write $X \succ Y$ for $X \gtrsim Y$ but $X \not\approx Y$. One of the first very interesting facts in set theory is Cantor's theorem:

Theorem 3.4 (Cantor)**.** *For every set X, $\mathcal{P}(X) \succ X$.*

Proof. Let $f: X \to \mathcal{P}(X)$ be a map. We will prove that f is not surjective. Define $D = \{a \in X : a \notin f(a)\}$. Then for every a, $D \neq f(a)$ as $a \in D \iff a \notin f(a)$. So D is not in the range of f. □

This type of argument is called a diagonalization argument.

Proposition 3.5.

1. $\mathbb{N} \approx \mathbb{Z} \approx \mathbb{Q} \approx \mathbb{N} \times \mathbb{N}$,
2. $\mathbb{R} \approx \mathcal{P}(\mathbb{N})$.

Proof. Clearly $\mathbb{N} \lesssim \mathbb{Z} \lesssim \mathbb{Q} \lesssim \mathbb{Z}\times\mathbb{Z}$. To see that $\mathbb{N} \approx \mathbb{Z}$, list out $\mathbb{Z}$ as follows: $0, 1, -1, 2, -2, 3, -3, \ldots$. This gives an injection from $\mathbb{N}$ to $\mathbb{Z}$. So we also have $\mathbb{N} \times \mathbb{N} \approx \mathbb{Z} \times \mathbb{Z}$.

Now to finish (1) we just need to find an injection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. We could take, for example, $(x, y) \mapsto 2^x 3^y$. There are also explicit bijections, like

$$(x, y) \mapsto \frac{(x + y)(x + y + 1)}{2 + x}$$

Now for (2), we can find an injection from $\mathbb{R}$ to $\mathcal{P}(\mathbb{Q}) \approx \mathcal{P}(\mathbb{N})$ by mapping r to $\{q \in \mathbb{Q} : q < r\}$. We can find an injection from $\mathcal{P}(\mathbb{N})$ to $\mathbb{R}$ as follows. To each $X \subseteq \mathbb{N}$, associate the sequence (x_i) where $x_i = 0$ if $i \notin X$, and $x_i = 2$ if $i \in X$. Now map X to the real number

$$\sum_{i \in \mathbb{N}} a_i 3^{-i-1}.$$

This is injective. □

We know that every set is in bijection with an ordinal, so every equivalence class under $\approx$ has an ordinal representative. Since the ordinals are well-ordered, we can choose a canonical representative.

Definition 3.6. Let X be a set. Denote by $|X|$, the cardinality of X, the least ordinal α with $X \approx \alpha$. A cardinal is an ordinal α with $|\alpha| = \alpha$.

Lemma 3.7. *$A \precsim B$ if and only if $|A| \leq |B|$.*

Proof. If $|A| \leq |B|$, then A is in bijection with $|A| \subseteq |B|$ which is in bijection with B. So there is an injection from A to B.

On the other hand, suppose that $|A| > |B|$. Then there is no bijection between A and $|B|$; since there is an injection from $|B|$ to $|A|$ and $|A|$ is in bijection with A, by Shroeder-Bernstein there must be no injection from A to $|B|$. Since $|B|$ is in bijection with B, there is not injection from A to B. Thus $A \not\precsim B$. □

Lemma 3.8.

1. *Any natural number is a cardinal.*
2. *ω is a cardinal.*
3. *Every infinite cardinal is a limit ordinal.*

Proof. (1) We argue by induction. Clearly $0 = \varnothing$ is a cardinal. Suppose that n is a cardinal. If $n+1$ is not a cardinal, then there is a bijection between $n+1$ and $m+1$ for some $m < n$ (clearly $n+1$ is not in bijection with $\varnothing$). But then there is a bijection between n and m, contradicting the fact that n is a cardinal.

(2) Clearly $|\omega| \geq n$ for each n, and so it cannot be that $|\omega| = n < n+1 \leq |\omega|$.

(3) We show that $|\alpha + 1| = |\alpha|$. Define an injection $f : \alpha + 1 \to \alpha$ by $f(0) = \alpha$, $f(n+1) = n$ for $n \in \omega$, and $f(\beta) = \beta$ for $\beta \in \alpha - \omega$. □

Definition 3.9. A set X is finite if $|X| < \omega$. X is countable if $|X| \leq \omega$, and uncountable if $|X| > \omega$. X is infinite if $|X| \geq \omega$.

Definition 3.10. Given a cardinal κ, define κ^+ to be the least cardinal greater than κ.

We can give ordinal names to the cardinals.

Definition 3.11. Define by transfinite induction:

1. $\aleph_0 = \omega$,
2. $\aleph_{\alpha+1} = \aleph_\alpha^+$,
3. $\aleph_\gamma = \sup\{\aleph_\alpha : \alpha < \gamma\}$ for γ a limit ordinal.

The first infinite cardinal is $\omega = \aleph_0$. The second infinite cardinal is $\aleph_1$. We know that $|\mathbb{R}| \geq \aleph_1$, but we do not know whether $|\mathbb{R}| = \aleph_1$. In fact, we cannot decide this within ZFC.

Definition 3.12. The continuum hypothesis (CH) is the statement that $|\mathbb{R}| = \aleph_1$, i.e., that there is no set X with $|\mathbb{N}| < |X| < |\mathbb{R}|$.

3.2 Cardinal Arithmetic

Adding sizes is different than adding order types. For example, addition with cardinals should be commutative. So we have new definitions of addition, multiplication, and exponentiation for cardinals. Of course, cardinals are actually ordinals, so it is possible for things to be confusing sometimes; but usually it will be clear whether we are using ordinal or cardinal operations. It is common to use $\aleph_0$ instead of ω when dealing with cardinalities, for example, to make this clear. We also tend to use κ, λ for cardinals and α, β, γ for ordinals.

Definition 3.13.

- $\kappa + \lambda = |\{0\} \times \kappa \cup \{1\} \times \lambda|$.
- $\kappa \cdot \lambda = |\kappa \times \lambda|$.
- $\kappa^\lambda = |\{f : \lambda \to \kappa\}|$, the cardinality of the set of functions from $\lambda \to \kappa$. Often we write X^Y or (often in set theory) ${}^Y X$ for this set of function.

Note that $|\mathcal{P}(X)| = 2^{|X|}$.

Theorem 3.14. *If κ is an infinite cardinal, then $\kappa \cdot \kappa = \kappa$.*

Proof. We argue by induction. We already know this for $\kappa = \omega$. So assume $\kappa > \omega$.

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

$$(\alpha, \beta) <_p (\gamma, \delta) \iff \begin{cases} \max(\alpha, \beta) < \max(\gamma, \delta) \\ (\alpha, \beta) <_{lex} (\gamma, \delta) & \text{if } \max(\alpha, \beta) = \max(\gamma, \delta) \end{cases}$$

Let W be this order.

First we show that W is a well-order. Suppose that $E \subseteq W$ is non-empty. First, let γ be the least element of $\{\max(\alpha, \beta) : (\alpha, \beta) \in E\}$. Then let α be the least element of $\{\alpha : \exists \beta\ (\alpha, \beta) \in E \text{ and } \gamma = \max(\alpha, \beta)\}$. Finally, let β be the least element of $\{\beta : (\alpha, \beta) \in E \text{ and } \gamma = \max(\alpha, \beta)\}$. Note that $(\alpha, \beta) \in E$ is the least element of E.

Now we argue that the order type of W is κ. If not, then for some $(\alpha, \beta) \in \kappa \times \kappa$ we have that $W{\restriction}_{(\alpha,\beta)}$ is isomorphic to κ. Let $\delta = \max(\alpha, \beta) + 1$; as κ is a limit ordinal, $\delta < \kappa$. So the order type of $W{\restriction}_{(\delta,\delta)}$ is $> \kappa$. Then

$$\kappa \leq |W{\restriction}_{\delta,\delta}| = |\delta \times \delta| = ||\delta| \times |\delta|| = |\delta|$$

by the induction hypothesis. Since $|\delta| < \kappa$, we have a contradiction. So the order type of W is κ, and $\kappa \cdot \kappa = \kappa$. □

As a result of this theorem, cardinal arithmetic is very simple for infinite cardinals.

Corollary 3.15. *If κ and λ are infinite cardinals, then $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$.*

Proof. Suppose that $\kappa \geq \lambda$. Then $\kappa \leq \kappa + \lambda \leq \kappa \cdot 2 \leq \kappa \cdot \lambda \leq \kappa \cdot \kappa = \kappa$. So these are all equalities. □

Corollary 3.16. *A countable union of countable sets is countable.*

Exponentiation is a little more complicated.

Lemma 3.17.

1. $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$.
2. $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$.
3. $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$.
4. *If* $\kappa \leq \lambda$ *then* $\kappa^\mu \leq \lambda^\mu$.
5. $\kappa^0 = 1$, $1^\kappa = 1$, *and* $0^\kappa = 0$ *if* $\kappa > 0$.

Proof. (1) follows from the fact that there is a correspondence between pairs of functions $f: X \to Y$ and $g: X \to Z$ and the function $(f, g): X \to Y \times Z$.

(2) follows from the fact that, for X and Y disjoint, there is a correspondence between pairs of functions $f: X \to Z$ and $g: Y \to Z$ and the function $f \cup g: X \bigcup Y \to Z$.

(3) follows from "currying": each function $f: X \times Y \to Z$ can be curried to obtain a function $g: X \to Y^Z$ given by $g(x) = \lambda y. f(x, y)$, and vice versa.

(4) and (5) are clear. □

Lemma 3.18. *If* λ *is an infinite cardinal, and* $2 \leq \kappa \leq \lambda$, *then* $\kappa^\lambda = 2^\lambda = |\mathcal{P}(\lambda)|$.

Proof.

$$2^\lambda \leq \kappa^\lambda \leq \lambda^\lambda \leq |\mathcal{P}(\lambda \times \lambda)| = |\mathcal{P}(\lambda)| = 2^\lambda.$$ □

4 Cofinality and Inaccessible Cardinals

Definition 4.1. Let α be an ordinal. The cofinality of α, $\mathrm{cf}(\alpha)$, is the least ordinal β such that there is a function $f: \beta \to \alpha$ which is unbounded: for each $\delta \in \alpha$ there is $\gamma \in \beta$ with $f(\gamma) \geq \alpha$.

We always have $\mathrm{cf}(\alpha) \leq \alpha$ by the identity map.

We can always choose the map witnessing the cofinality of α to be non-decreasing.

Lemma 4.2. *If the map* $f: \beta \to \alpha$ *is cofinal, then there is a non-decreasing cofinal map* $g: \gamma \to \alpha$ *for some* $\gamma \leq \beta$.

Proof. Let $\gamma \leq \beta$ be least such that $\{\sup f(\delta) | \delta < \gamma\} = \alpha$. Define $g(\delta) = \sup\{f(\delta') : \delta' \leq \delta\}$. □

Lemma 4.3. *The cofinality of an ordinal is a cardinal.*

Proof. Suppose that there is a cofinal map $f: \beta \to \alpha$. Let $\kappa = |\beta|$. Then there is a cofinal map $\kappa \to \beta \to \alpha$ by composing f with the bijection between β and κ. □

Lemma 4.4. $\mathrm{cf}(\mathrm{cf}(\alpha)) = \mathrm{cf}(\alpha)$.

Proof. Given non-decreasing cofinal maps $f: \gamma \to \beta$ and $g: \beta \to \alpha$, $g \circ f: \gamma \to \alpha$ is also cofinal. □

Definition 4.5. An infinite cardinal $\aleph_\alpha$ is regular if $\mathrm{cf}(\aleph_\alpha) = \aleph_\alpha$, and otherwise it is singluar.

Lemma 4.6. *Every successor cardinal is regular.*

Proof. If not there is a cofinal function $f\colon \mu \to \kappa^+$ where $\mu \leq \kappa$ is a cardinal. Now for each $\alpha \in \mu$, since $f(\alpha) < \kappa^+$, $|f(\alpha)| \leq \kappa$; so we can use the axiom of choice to pick a surjective map $g_\alpha\colon \kappa \to f(\alpha)$.

Then define $g\colon \mu \times \kappa \to \kappa^+$ by $g(\alpha, \beta) = g_\alpha(\beta)$. This map is onto κ^+. Thus $\kappa = |\mu \times \kappa| = \kappa^+$ which is a contradiction. □

Note that $\aleph_\omega$ is singular, as $(\aleph_n)_{n\in\omega}$ is a cofinal sequence. $\aleph_0$ is regular.

4.1 König's Theorem

We can extend addition and multiplication of cardinals to infinite sums and products, namely the sum of cardinals is the cardinality of the disjoint union, and the product of cardinals is the cardinality of the Cartesian product. The statement that arbitrary Cartesian products are non-empty is esentially a restatement of the Axiom of Choice.

The following theorem is somewhat surprising in that we get a strict inequality, whereas often with infinite sums or products one only expects to get a non-strict inequality.

Theorem 4.7 (König's Theorem)**.** *Let I be a set, and let $(\lambda_i)_{i\in I}$ and $(\kappa_i)_{i\in I}$ be sequence of cardinals such that $\kappa_i < \lambda_i$. Then*

$$\sum_{i\in I} \kappa_i < \prod_{i\in I} \lambda_i.$$

Proof. To prove the non-strict inequality

$$\sum_{i\in I} \kappa_i \leq \prod_{i\in I} \lambda_i$$

fix, for each i, an injection $f_i\colon \kappa_i \to \lambda_i$ and pick $\alpha_i \in \lambda_i - f_i(\kappa_i)$. Then map (j, β) in the disjoint union $\bigcup_i \{(i, \beta) : \beta \in \kappa_i\}$ to the element of $\prod_{i\in I} \lambda_i$ whose jth entry is $f_j(\beta)$, and whose ith entry $(i \neq j)$ is α_i. It is not hard to argue that is an injection, using the fact that $\alpha_i \notin \operatorname{ran} f_i$.

Now to show that it is strict, note (by the Axiom of Choice, and the fact that each $\lambda_i > 0$) that $\prod_{i\in I} \lambda_i$ is non-empty. We will show that there is no surjection from the disjoint union of the κ_i onto $\prod_{i\in I} \lambda_i$. Let f be such a map; we will show that f is not surjective. Let f_i be the composition of f with the projection onto the ith coordinate λ_i. Then f_i induces a map from κ_i to λ_i; since $\kappa_i < \lambda_i$, there is $\alpha_i \in \lambda_i - f_i(\kappa_i)$. Let $\alpha = (\alpha_i)_{i\in I}$. Then α is not the image of any element in the disjoint union of the κ_i, because for each i its ith coordinate is different from the ith coordinate of the image of any element of κ_i under f. □

There are a few interesting consequences of König's Lemma. For example, given a cardinal μ if we take $I = \mu$, $\kappa_i = 1$, and $\lambda_i = 2$, then we get $\mu < 2^\mu$; this is Cantor's Theorem.

We also get:

Corollary 4.8. *If κ is an infinite cardinal, then $\kappa < \kappa^{\mathrm{cf}(\kappa)}$.*

Proof. Choose κ_i, $i \in \mathrm{cf}(\kappa)$, such that $\kappa = \sup_{i \in \mathrm{cf}(\kappa)} \kappa_i$. Then

$$\kappa \leq \kappa \cdot \mathrm{cf}(\kappa) \leq \sum_{i \in \mathrm{cf}(\kappa)} \kappa_i \leq \prod_{i \in \mathrm{cf}(\kappa)} \kappa_i \leq \kappa^{\mathrm{cf}(\kappa)}.$$ □

Corollary 4.9. *Let κ be a regular cardinal. Then $\kappa < \mathrm{cf}(2^\kappa)$.*

Proof. Suppose instead that $\kappa \geq \mathrm{cf}(2^\kappa)$. Then, by the previous corollary,

$$2^\kappa < (2^\kappa)^{\mathrm{cf}(2^\kappa)} \leq (2^\kappa)^\kappa \leq 2^{\kappa \cdot \kappa} \leq 2^\kappa.$$

This is a contradiction. □

Easton's Theorem, which is beyond the scope of the class, says that this and $\lambda \leq \kappa \Rightarrow 2^\lambda \leq 2^\kappa$ are the only two restrictions on the cardinality of the powerset. (I.e., for any class function from cardinals to cardinals satisfying these two conditions, it is consistent with ZFC that that function gives the cardinality of the powerset.)

Kónig's Lemma also says something about why we should care about why we should care about regular cardinals:

Corollary 4.10. *A cardinal κ is regular if and only if it is not the union of $< \kappa$-many cardinals, each of which is $< \kappa$.*

This motivates the following definition:

Definition 4.11.

1. A weakly inaccessible cardinal is an uncountable regular limit cardinal.
2. A strongly inaccessible cardinal is an uncountable regular limit cardinal κ such that for all $\lambda < \kappa$, $2^\lambda < \kappa$.

The existence of weakly inaccessible cardinals is not provable in ZFC, but it is consistent with ZFC.

5 Ultrafilters and Applications

5.1 Ultrafilters

Definition 5.1. Let X be a set and $\mathcal{F}$ a collection of subsets of X. We call $\mathcal{F}$ a filter if:

1. $\varnothing \notin \mathcal{F}$,
2. If $A \in \mathcal{F}$ and $A \subseteq B$, then $B \in \mathcal{F}$,
3. If $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

The intuition is that sets in the filter are large. For example, the cofinite subsets of X form a filter, or every $\{X\}$ is a filter on X. We want to add a condition that puts more sets into the filter.

Definition 5.2. An ultrafilter on X is a filter $\mathcal{F}$ on X such that for every set $A \subseteq X$, either $X \in \mathcal{F}$ or $X - A \in \mathcal{F}$.

Essentially, an ultrafilter decides, for each set, whether that set or its compliment is large. It is still easy to find trivial examples of ultrafilters, for example, $\{X \subseteq \omega : 7 \in X\}$ is an ultrafilter. But it is not very useful: it says that the "large" sets are exactly those which contain 7. We call such an ultrafilter principal.

Definition 5.3. An ultrafilter $\mathcal{U}$ on X is principal if there is $x \in X$ such that $\mathcal{U} = \{Y \subseteq X : x \in Y\}$; otherwise, $\mathcal{U}$ is non-principal.

If a set is finite, then any ultrafilter on it is principal. Otherwise, we can use the Axiom of Choice to show the existence of a non-principal ultrafilter.

Lemma 5.4. *Let X be a finite set. Every ultrafilter on X is principal.*

Proof. Let $\mathcal{U}$ be an ultrafilter on X. If $\mathcal{U}$ is not principal, then for each $x \in X$, either $\{x\}$ or $X - \{x\}$ is in $\mathcal{U}$. If $\{x\} \in \mathcal{U}$, then $\mathcal{U}$ is the principal ultrafilter generated by x as each superset of X must be contained in $\mathcal{U}$, and each set Y disjoint from x must not be in $\mathcal{U}$ as $Y \cap \{x\} = \varnothing \notin \mathcal{U}$. Since this cannot happen, for each x, $X - \{x\}$ is in $\mathcal{U}$; then the (finite) intersection of all of these, which is $\varnothing$, is in $\mathcal{U}$. This is a contradiction. □

Theorem 5.5. *Every filter can be extended to an ultrafilter.*

Proof. We use Zorn's Lemma. Let P be the partial order of filters on X extending a given filter $\mathcal{G}$, ordered by inclusion.

Let $\mathcal{C}$ be a chain in P. We argue that $\bigcup \mathcal{C}$ is a filter on X extending $\mathcal{G}$, so that by Zorn's Lemma, P has a maximal element. Clearly $\bigcup \mathcal{C}$ extends the $\mathcal{G}$, and it does not contain $\varnothing$ as no filter in C contains $\varnothing$. Given $X \in \bigcup \mathcal{C}$ and $X \subseteq Y$, there is some filter $\mathcal{F} \in \mathcal{C}$ with $X \in \mathcal{F}$; so $Y \in \mathcal{F} \subseteq \bigcup \mathcal{C}$. Finally, given $X, Y \in \bigcup \mathcal{C}$, there are $\mathcal{F}_1 \ni X$ and $\mathcal{F}_2 \ni Y$. Since $\mathcal{C}$ is a chain, for $i \in \{1, 2\}$, $X, Y \in \mathcal{F}_i$, so that $X \cap Y \in \mathcal{F}_i$. Thus $X \cap Y \in \bigcup \mathcal{C}$.

Let $\mathcal{F}$ be a maximal element of P. We argue that $\mathcal{F}$ is an ultrafilter. Fix $Y \subseteq X$, and suppose to the contrary that $Y \notin \mathcal{F}$ and $X - Y \notin \mathcal{F}$. Then either every set in $\mathcal{F}$ intersects Y, or every set in $\mathcal{F}$ intersects $X - Y$; otherwise, we would have sets $Z_1 \subseteq Y$ and $Z_2 \subseteq X - Y$ in $\mathcal{F}$, and hence their intersection, $Z_1 \cap Z_2 = \varnothing$, would be in $\mathcal{F}$. Suppose without loss of generality that every set in $\mathcal{F}$ intersects Y.

Let $\mathcal{F}' = \{Z : \exists Z' \in \mathcal{F} \quad Z' \cap Y \subseteq Z\}$. Then $\mathcal{F}' \supsetneq \mathcal{F}$. We argue that $\mathcal{F}'$ is a filter. It does not contain $\varnothing$ as every set in $\mathcal{F}$ intersects Y. The other properties of a filter are easy to see. This is a contradiction. Hence $\mathcal{F}$ is an ultrafilter. □

Corollary 5.6. *Let X be an infinite set. There is a non-principal ultrafilter on X.*

Proof. Recall that the collection Cof of all cofinite subsets of X is a filter. So there is an ultrafilter extending Cof. This cannot be principal, as for each $x \in X$, $X - \{x\}$ is in Cof, and so $\{x\}$ cannot be in the ultrafilter. □

One can think of an ultrafilter as a way of deciding votes. Suppose that X is the set of voters, and that the voters Y vote for some option (a) while the voters $X - Y$ vote for some other option (b). Then an ultrafilter on X decides whether the voters Y or $X - Y$ win the vote.

5.2 Ramsey's Theorem

We will show how to apply ultrafilters to obtain a proof of Ramsey's theorem. We denote by $[X]^n$ all of the n-element subsets of X.

Theorem 5.7 (Ramsey's Theorem for infinite sets)**.** *Let n, k be natural numbers, and let $c\colon [\omega]^n \to k$ be a function. Then there is an infinite subset $H \subseteq \omega$ such that c is constant on subsets of H.*

We think of the function c as colouring subsets of ω of size n with one of k colours. Then the set H is one all of whose subsets receive the same colour; we call H *homogeneous*.

Proof. We will prove the case $n = k = 2$. Let $\mathcal{U}$ be a non-principal ultrafilter on ω. For each $a \in \omega$, consider partition of $\omega - \{a\}$ into sets $\{b : c(a,b) = 0\}$ and $\{b : c(a,b) = 1\}$. One of these two sets must be in the ultrafilter. If it is the former, define $d(a) = 0$, and if the later, $d(a) = 1$.

Now consider the sets $\{a : d(a) = 0\}$ and $\{a : d(a) = 1\}$. These sets partition ω, and so one of them is in $\mathcal{U}$. Let $i \in \{0, 1\}$ be such that $\{a : d(a) = i\} \in \mathcal{U}$. We will now construct a set H which is homogeneous for the colour i.

Let $U = \{a : d(a) = i\}$. Given $a \in U$, let $U_a = \{b \in \omega : c(a,b) = i\}$. These sets are all in $\mathcal{U}$. Then define a_0 to be the least element of U. Define a_{n+1} recursively to be the least element of

$$U \cap U_{a_0} \cap U_{a_1} \cap \cdots \cap U_{a_n} \cap \{b : b > a_n\}.$$

Note that this intersection is in $\mathcal{U}$, hence non-empty. Let $H = \{a_n : n \in \omega\}$. We claim that H is homogeneous for the colour i. Indeed, given $m < n$, $a_n \in U_{a_m}$ and so $c(a_m, a_n) = i$. □

One consequence of Ramsey's Theorem is the following:

Theorem 5.8. *Let $(L, \leq)$ be an infinite linear order. Then L has either an infinite strictly ascending sequence or an infinite strictly descending sequence.*

Proof. Take an injection $f\colon \omega \to L$. Define a colouring c of $[\omega]^2$ by letting $c(\{m, n\})$, for $m < n$, be 0 if $f(m) < f(n)$ and 1 if $f(m) > f(n)$.

Let H be an infinite homogeneous set. Suppose that $H$is homogeneous for the colour 0. Let $a_0, a_1, a_2, \ldots$ list the elements of H. Then for $i < j$, $c(a_i, a_j) = 0$ and so $f(a_i) < f(a_j)$. Thus $f(a_0) < f(a_1) < f(a_2) < \cdots$.

A similar argument works if H is homogeneous for the colour 1, except that we get a descending sequence. □

There is also a version of Ramsey's theorem for finite sets. Often this is proved using the Compactness Theorem, but we will prove it directly using ultrafilters. (Indeed, one popular proof of the compactness theorem which we will see soon uses ultrafilters.)

Theorem 5.9 (Ramsey's Theorem for finite sets)**.** *For any $k \in \omega$ there is an $n \in \omega$ such that for any colouring $c\colon [n]^2 \to 2$ there is a homogenous set $H \subseteq n$ of size k.*

Proof. Assume that the finite version of Ramsey's theorem fails for some k. We will use compactness / ultrafilters to show that Ramsey's theorem for infinite sets must fail, a contradiction. Fix an ultrafilter $\mathcal{U}$ on ω.

For each n, let $c_n : [n]^2 \to 2$ be a colouring which does not admit a homogeneous set of size k. Define a colouring $d : [\omega]^2 \to 2$ by $d(i,j) = 0$ if $\{n : c_n(i,j) = 0\} \in \mathcal{U}$ and $d(i,j) = 1$ if $\{n : c_n(i,j) = 1\} \in \mathcal{U}$. (Note that $c_n(i,j)$ is undefined for only finitely many n, and ultrafilters ignore finite sets.) Let $H \subseteq \omega$ be an infinite set homogeneous for d, say without loss of generality for the colour 0.

Let $a_1, \ldots, a_k$ be distinct elements of H. For each $i < j$, let $U_{i,j} = \{n : c_n(a_i, a_j) = 0\}$; then $U_{i,j} \in \mathcal{U}$ for each i, j. So there is

$$n \in \bigcap i < j \leq k U_{i,j}.$$

The set $\{a_1, \ldots, a_k\}$ is homogeneous for c_n. □

5.3 Ultraproducts and Łoś Theorem

Definition 5.10. Let $\mathcal{U}$ be an ultrafilter on I, let $\mathcal{L}$ be a first-order language, and let $(\mathcal{A}_i)_{i \in I}$ be a sequence of $\mathcal{L}$-structures. The ultraproduct

$$\prod_{i \in I} \mathcal{A}_i / \mathcal{U}$$

is the structure defined as follows.

The domain of $\prod_{i \in I} \mathcal{A}_i / \mathcal{U}$ is the set $\prod_{i \in I} A_i / \mathcal{U}$ of equivalence classes of the following equivalence relation

$$(a_i)_{i \in I} \sim (b_i)_{i \in I} \iff \{i \in I : a_i = b_i\} \in \mathcal{U}$$

on $\prod_{i \in I} A_i$. Essentially, we set two tuples to be equal if they are equal on a large set of entries. Then we define the relations and functions as follows:

$$R([a^1], \ldots, [a^n]) \iff \{i \in I : R^{\mathcal{A}_i}(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$$

and

$$f([a^1], \ldots, [a^n]) = (f^{\mathcal{A}_i}(a_i^1, \ldots, a_i^n))_{i \in I}.$$

We have to note that these definitions are well-defined, i.e., if $[a^1] = [b^1], \ldots, [a^n] = [b^n]$, then

$$R([a^1], \ldots, [a^n]) \iff R([b^1], \ldots, [b^n]).$$

This is because the set $\{i \in I : a_i^1 = b_i^1, \ldots, a_i^n = b_i^n\}$ is in $\mathcal{U}$.

If we fix a single structure $\mathcal{A}$ and take $\mathcal{A}_i = \mathcal{A}$ for every i, then we write $\prod_{i \in I} \mathcal{A} / \mathcal{U}$ and say that this is an ultraproduct.

When something happens on a set in the ultrafilter, we say that it happens almost everywhere; for example, if $a \sim b$, we say that a and b are equal almost everywhere.

Ultraproducts are only interesting when the ultrafilter is non-principal; otherwise, the ultraproduct is just isomorphic to one of the facts.

Lemma 5.11. *$\mathcal{A}$ is a substructure of $\prod_{i \in I} \mathcal{A} / \mathcal{U}$ via the embedding $a \mapsto [(a)_{i \in I}]$.*

Proof. If $a, b \in \mathcal{A}$, $a \neq b$, then $[(a)_{i\in I}] \neq [(b)_{i\in I}]$ because they are not equal almost everywhere. Similarly, we can show that the same functions and relations hold of such elements. □

Example 5.12. Let $\mathcal{U}$ be a non-principal ultrafilter on $\mathbb{N}$. Consider $\prod_{i\in\mathbb{N}} \mathbb{R}/\mathcal{U}$ as an ordered field. The previous lemma says that this contains a substructure isomorphism to $\mathbb{R}$; the domain of this substructure is the set of equivalence classes of the constant sequences $[(r)_{i\in\mathbb{N}}]$ for each $r \in \mathbb{R}$, which we identify with r. But there are other new elements. For example, there are infinitesimal elements such as $[(\frac{1}{n})_{n\in\mathbb{N}}]$, which satisfies for each $r \in R$, $r > 0$,

$$0 < [(\frac{1}{n})_{n\in\mathbb{N}}] < r.$$

(For each fixed r, $0 < \frac{1}{n} < r$ for all but finitely many $n \in \mathbb{N}$.) We also have infinite elements such as $[(n)_{n\in\mathbb{N}}]$ which satisfies

$$[(n)_{n\in\mathbb{N}}] > r$$

for all $r \in \mathbb{R}$.

The power of ultraproducts is that what is true in an ultraproduct is exactly what is true in almost all of the coordinates. This is Łoś Theorem (pronounced "wash"):

Theorem 5.13 (Łoś Theorem)**.** *Let $\mathcal{L}$ be a first-order language, U an ultrafilter on I, and for each $i \in I$, $\mathcal{A}_i$ an $\mathcal{L}$-structure. Then for each $a^1, \ldots, a^n \in \prod_{i\in I} A_i$,*

$$\prod_{i\in I} \mathcal{A}_i/\mathcal{U} \vDash \varphi([a^1], \ldots, [a^n]) \iff \{i \in I : \mathcal{A}_i \vDash \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}.$$

Proof. The proof is by induction on formulas. We will assume that the language is relational by replacing n-ary functions by $n+1$-ary relations.

- The theorem is true for equality and for atomic formulas by definition of the ultrapower.
- Suppose that the theorem holds for φ and ψ. It is easy to see that the theorem holds for $\varphi \wedge \psi$ by noting that if and only if

  $$\{i \in I : \mathcal{A}_i \vDash \varphi(a_i^1, \ldots, a_i^n)\} \cap \{i \in I : \mathcal{A}_i \vDash \psi(a_i^1, \ldots, a_i^n)\} = \{i \in I : \mathcal{A}_i \vDash [\varphi \wedge \psi](a_i^1, \ldots, a_i^n)\}$$

  and that the right hand side is in $\mathcal{U}$ if and only if both the sets on the left hand side are in $\mathcal{U}$.
- Suppose that the theorem holds for φ. Then

  $$\begin{aligned} \prod_{i\in I} \mathcal{A}_i/\mathcal{U} \vDash \neg\varphi([a^1], \ldots, [a^n]) &\iff \prod_{i\in I} \mathcal{A}_i/\mathcal{U} \nvDash \varphi([a^1], \ldots, [a^n]) \\ &\iff \{i \in I : \mathcal{A}_i \vDash \varphi(a_i^1, \ldots, a_i^n)\} \notin \mathcal{U} \\ &\iff \{i \in I : \mathcal{A}_i \vDash \neg\varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U} \end{aligned}$$

  where the last line uses the fact that $\mathcal{U}$ is an ultrafilter.

- Suppose that the theorem holds for φ. Then suppose that

$$\prod_{i \in I} \mathcal{A}_i/\mathcal{U} \vDash \exists x \varphi(x, [a^1], \ldots, [a^n]).$$

Pick a representative $b \in \prod A_i$ such that

$$\prod_{i \in I} \mathcal{A}_i/\mathcal{U} \vDash \varphi([b], [a^1], \ldots, [a^n]).$$

Note that this uses the Axiom of Choice to choose a particular representative b of the equivalence class. Then

$$\{i \in I : \mathcal{A}_i \vDash \varphi(b_i, a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$$

and this set is a subset of the set

$$\{i \in I : \mathcal{A}_i \vDash \exists x \varphi(x, a_i^1, \ldots, a_i^n)\}$$

which is thus also in $\mathcal{U}$.

On the other hand, suppose that

$$\{i \in I : \mathcal{A}_i \vDash \exists x \varphi(x, a_i^1, \ldots, a_i^n)\} \in \mathcal{U}.$$

For each i in this set, pick b_i with $\mathcal{A}_i \vDash \varphi(b_i, a_i^1, \ldots, a_i^n)\}$. For each other i, pick any b_i. Let $b = (b_i)_{i \in I}$; note that $[b]$ is determined only by the former choices but not by the latter. Then

$$\{i \in I : \mathcal{A}_i \vDash \varphi(b_i, a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$$

and so

$$\prod_{i \in I} \mathcal{A}_i/\mathcal{U} \vDash \varphi([b], [a^1], \ldots, [a^n]).$$

This proves the theorem. □

One application is a proof of the Compactness theorem without passing through the Completeness theorem.

Theorem 5.14 (Compactness theorem)**.** *Let $\mathcal{L}$ be a first-order language. A set Φ of $\mathcal{L}$-sentences is satisfiable if and only if each finite subset is satisfiable.*

Proof. One direction is easy. For the other, suppose that each finite subset of Φ is satisfiable. Let I be the collection of all finite subsets of Φ. For each $i \in I$, pick a model $\mathcal{A}_i$ of that finite set of sentences.

For each $i \in I$, let $i^* = \{j \in I : i \subseteq j\}$. Let I^* be the set of all subsets of I containing i^* for some $i \in I$. It is easy to see that I^* is a filter; the only nontrivial property to check is that intersections of elements of I^* are also in I^*, and this is due to the fact that $i_0^* \cap i_1^* = (i_0 \cup i_1)^*$. So there is an ultrafilter $\mathcal{U}$ on I extending I^*.

Let $\mathcal{A}_I/\mathcal{U} = \prod_{i \in I} \mathcal{A}_i/\mathcal{U}$. We claim that $\mathcal{A}_I/\mathcal{U} \vDash \Phi$. Given $\varphi \in \Phi$, $\{\varphi\} \in I$ and $\mathcal{A}_i \vDash \varphi$ for every i with $i \ni \varphi$. So

$$\{\varphi\}^* = \{i \in I : \varphi \in i\} \subseteq \{i \in I : \mathcal{A}_i \vDash \varphi\}.$$

Since $\{\varphi\}^* \subseteq I^* \subseteq \mathcal{U}$, by Łoś Theorem, $\mathcal{A}_I/\mathcal{U} \vDash \varphi$. □

6 The Cumulative Hierarchy

Definition 6.1. Define by transfinite recursion:

- $V_0 = \varnothing$,
- $V_{\alpha+1} = \mathcal{P}(V_\alpha)$,
- $V_\beta = \bigcup_{\alpha<\beta} V_\alpha$.

This gives a class function $\alpha \mapsto V_\alpha$. Let $V = \bigcup_{\alpha \in ORD} V_\alpha$. This is called the Von Neumann universe.

Recall that a set x is transitive if whenever $y \in x$, $y \subseteq x$.

Lemma 6.2. *The power set of a transitive set is transitive.*

Proof. Let x be a transitive set. If $z \in y \in \mathcal{P}(x)$, then $y \subseteq x$, and so $z \in x$. Since x is transitive, $z \subseteq x$ and so $z \in \mathcal{P}(x)$. □

Lemma 6.3. *For each α, V_α is transitive.*

Proof. By transfinite induction. $V_0 = \varnothing$ is transitive. If V_α is transitive, then $V_{\alpha+1} = \mathcal{P}(V_\alpha)$ is transitive by the previous lemma. Finally, the union of transitive sets is transitive. □

Lemma 6.4. *If $\alpha \leq \beta$, then $V_\alpha \subseteq V_\beta$.*

Proof. Fix α for which we will show that for all $\beta \geq \alpha$, $V_\alpha \subseteq V_\beta$. We argue by transfinite induction on β. For $\beta = \alpha$ this is clear. If $V_\alpha \subseteq V_\beta$, then $V_\alpha \in \mathcal{P}(V_\beta) = V_{\beta+1}$. Since $V_{\beta+1}$ is transitive, $V_\alpha \subseteq V_{\beta+1}$. Finally, if β is a limit ordinal and $V_\alpha \subseteq V_\gamma$ for some $\gamma < \beta$, then $V_\alpha \subseteq V_\beta$. □

We want to show that V contains every set. First, we need to introduce the transitive closure of a set.

Definition 6.5. Let x be a set. The transitive closure of x is defined as follows:

$$x_0 = x, \qquad x_{n+1} = \bigcup x_n, \qquad TC(x) = \bigcup x_n.$$

Note that for every set x, $TC(x)$ is transitive.

Lemma 6.6. *Every non-empty class has an $\in$-minimal element.*

Proof. Let C be a class and pick $x \in C$. If $\{y \in x | y \in C\}$ is empty, then x is an $\in$-minimal element of C. Otherwise, $TC(x) \cap C$ is non-empty, and by the Axiom of Foundation, has an $\in$-minimal element y. Then y is an $\in$-minimal element of C; indeed, if $z \in y \cap C$, then $z \in TC(x)$ which cannot be the case as y is an $\in$-minimal element of $TC(x) \cap C$. □

Theorem 6.7. *Every set x is in V.*

Proof. Suppose that there is some set not in V. Then let x be an $\in$-minimal element of the complement of V. For all $y \in x$, $y \in V$. Define a class function F such that $F(y)$ is the least ordinal α such that $y \in V_\alpha$. Then $\{f(y) : y \in x\}$ is a set of ordinals, and so must be bounded above by some ordinal β (otherwise, the downward closure of this set is a set containing every ordinal). So if $y \in x$, then $y \in V_\beta$. But then $x \in V_{\beta+1}$. □

We can think of this as giving a rank function on sets.

Definition 6.8. Define $\text{rank}(x)$ to be the least ordinal α such that $x \in V_{\alpha+1}$.

Lemma 6.9.

1. *If $x \in y$, then* $\text{rank}(x) < \text{rank}(y)$.
2. *If α is an ordinal, then* $\text{rank}(\alpha) = \alpha$.

Proof. (i) Let $\alpha = \text{rank}(y)$, so $y \in V_{\alpha+1}$. So $y \subseteq V_\alpha$, and $x \in V_\alpha$. If α is a successor ordinal $\alpha = \gamma + 1$, then $\text{rank}(x) = \gamma$. Otherwise, if α is a limit ordinal, $x \in V_\alpha = \bigcup_{\gamma<\alpha} V_\gamma$ and so $x \in V_{\gamma+1}$ for some $\gamma < \alpha$. Then $\text{rank}(x) = \gamma$ for the least such γ.

(ii) First, we show that $\text{rank}(\alpha) \leq \alpha$. We argue by induction. First, $0 \in \mathcal{P}(\varnothing) = V_1$ and so $\text{rank}(0) \leq 0$. If $\alpha \in V_{\alpha+1}$, then $\alpha + 1 = \alpha \cup \{\alpha\} \subseteq V_{\alpha+1}$ by transitivity of $V_{\alpha+1}$, and so $\alpha + 1 \in V_{\alpha+2}$. If α is a limit ordinal, then if $\gamma < \alpha$, $\gamma \in V_{\gamma+1}$. So $\alpha \subseteq V_\alpha = \bigcup V_\gamma$. So we have shown that $\text{rank}(\alpha) \leq \alpha$.

Now suppose that $\text{rank}(\alpha) < \alpha$, and let α be least with this property. Let $\beta = \text{rank}(\alpha) < \alpha$. Then by (i), $\text{rank}(\beta) < \text{rank}(\alpha) = \beta$, contradicting the minimality of α. □

7 Relativizing Formulas

Let C be a class and E a binary class relation (so E is defined by a formula $\varphi(x, y)$, possibly with parameters). Let φ be a formula in the language of set theory with n free variables, and let $a_1, \ldots, a_n \in C$. Then we define $C, E \vDash \varphi(a_1, \ldots, a_n)$ inductively as follows:

1. $C, E \vDash a_i \in a_j$ if $a_i E a_j$.
2. $C, E \vDash a_i = a_j$ if $a_i = a_j$.
3. $C, E \vDash \varphi \wedge \psi$ if $C, E \vDash \varphi$ and $C, E \vDash \psi$.
4. $C, E \vDash \neg\varphi$ if $C, E \nvDash \varphi$.
5. $C, E \vDash \exists x \varphi(x)$ if there is $a \in C$ such that $C, E \vDash \varphi(a)$.

This is a definition in the metalanguage. Usually we will have $E = \in$, in which case we just write $C \vDash \varphi$ for $(C, \in) \vDash \varphi$. We call (C, E) a model of set theory.

Definition 7.1. Let $M \subseteq N$ be models of set theory. Let φ be a formula with n free variables. We say that φ is absolute between M and N if for all $a_1, \ldots, a_n \in M$,

$$M \vDash \varphi(a_1, \ldots, a_n) \iff N \vDash \varphi(a_1, \ldots, a_n).$$

If φ is absolute between M and V, we just say that φ is absolute for M.

If M is a transitive class, then $(M, \in)$ (or just M) is called a transitive model. If $M \subseteq N$ are transitive models, then there is a large class of formulas that are absolute between them.

Definition 7.2. A formula φ is called a Δ_0 formula if:

- every atomic formula is Δ_0,
- if φ is Δ_0, then $\neg\varphi$ is Δ_0,
- if φ and ψ are Δ_0, then $\varphi \wedge \psi$ is Δ_0,
- if φ is Δ_0, then $(\exists x \in y)\varphi$ and $(\forall x \in y)\varphi$ are Δ_0.

Again this is a definition in the metalanguage.

Lemma 7.3. *If M is a transitive model and φ is absolute, then φ is absolute for M.*

Proof. We argue by induction on formulas.

- every atomic formula clearly absolute for M.
- if φ is absolute for M, then

  $$M \vDash \neg\varphi(a) \iff M \nvDash \varphi(a) \iff V \nvDash \varphi(a) \iff V \vDash \neg\varphi(a).$$

  and so $\neg\varphi$ is absolute for M,
- if φ and ψ are absolute for M, then so is $\varphi \wedge \psi$ by a similar argument as the previous case.
- if $\varphi(x, y, \bar{z})$ is absolute for M, and $b, \bar{c} \in M$, then if

  $$M \vDash (\exists x \in b)\varphi(x, b, \bar{c})$$

  there is $a \in b$ with

  $$M \vDash \varphi(a, b, \bar{c}).$$

  Since φ is absolute for M,

  $$V \vDash \varphi(a, b, \bar{c})$$

  and so

  $$V \vDash (\exists x \in b)\varphi(x, b, \bar{c}).$$

  On the other hand, if

  $$V \vDash (\exists x \in b)\varphi(x, b, \bar{c})$$

  then there is $a \in b$ such that

  $$V \vDash \varphi(a, b, \bar{c}).$$

  Since $a \in b \in M$ and M is transitive, $a \in M$. Then since φ is absolute for M,

  $$M \vDash \varphi(a, b, \bar{c})$$

  and so

  $$M \vDash (\exists x \in b)\varphi(x, b, \bar{c}).$$

□

In particular, many formulas such as $x \subseteq y$ $(\forall u \in x(u \in y))$ are absolute for transitive models.

Lemma 7.4. *If $\vdash (\forall x)\varphi(x) \leftrightarrow \psi(x)$, and $\psi(x)$ is absolute for M, then $\varphi(x)$ is absolute for M.*

Proof. Let $a \in M$. We have

$$M \vDash \varphi(a) \iff M \vDash \psi(a) \iff V \vDash \psi(a) \iff V \vDash \varphi(a).$$ □

Theorem 7.5. *If α is a limit ordinal, then V_α models all of the axioms of ZFC except possibly replacement and infinity.*

If $\alpha > \omega$, then V_α models the Axiom of Infinity.

Proof.

Extensionality: The formula which expresses Extensionality for two sets x and y is equivalent to a Δ_0 formula:

$$\big((\forall z \in x)[z \in y] \wedge (\forall z \in y)[z \in x]\big) \longleftrightarrow x = y.$$

Hence it is absolute for V_α.

Pairing: Given $x, y \in V_\alpha$, let $z = \{x, y\}$. Since α is a limit ordinal, we can choose $\beta < \alpha$ such that $x, y \in V_\beta$. Then $z \subseteq V_\beta$, and so $z \in V_{\beta+1}$. Also, the formula expressing the required property of z is Δ_0 and hence holds in V_α:

$$V_\alpha \vDash \forall u \in z(u = x \vee u = y) \ \wedge \ x \in z \ \wedge \ y \in z.$$

Union: Homework.

Foundation: Take $x \in V_\alpha$, $x \neq \varnothing$. By Foundation in V, let $y \in x$ be such that $y \cap x = \varnothing$. Since V_α is transitive, $y \in V_\alpha$. That $x \cap y = \varnothing$ is expressible by a Δ_0 formula $\forall u \in y(u \notin x)$. So

$$V_\alpha \vDash x \cap y = \varnothing$$

and Foundation holds in V_α.

Infinity: Here we use that $\alpha > \omega$ and so $\omega \in V_\alpha$. The formula which says that ω is inductive is Δ_0:

$$V \vDash \varnothing \in \omega \wedge \forall x \in \omega(x \cup \{x\} \in \omega)$$

where $x \cup \{x\} \in \omega$ is just the Δ_0 formula

$$\exists y \in \omega(x \in y \ \wedge \ \forall u \in x(u \in y) \ \wedge \ \forall u \in y(u \in x \vee u = x)).$$

So ω is inductive in V_α.

Powerset: Let $x \in V_\alpha$. Choose $\beta < \alpha$ such that $x \in V_\beta$, and so $x \subseteq V_\beta$ by transitivity. Let $u = \mathcal{P}(x)$. So $u \in V_{\beta+1}$. We claim that u is the powerset of x in V_α. We have that $V_\alpha \vDash \forall z \in$

$u(z \subseteq x)$ since this is Δ_0 hence absolute. Now we must argue that $V_\alpha \vDash \forall z(z \subseteq x \to z \in u)$. If $z \in V_\alpha$ and $z \subseteq x$, this is also true in V and so $z \in u$.

Separation: Take $x, \bar{p} \in V_\alpha$. Let $\varphi(u, \bar{p})$ be a formula. Let $\beta < \alpha$ be such that $x, \bar{p} \in V_\beta$. Define

$$y = \{u \in x : V_\alpha \vDash \varphi(u, \bar{p})\}.$$

We will show that y is in V, from which it follows that as $y \in \mathcal{P}(x)$ that $y \in V_{\beta+1}$.

To see that y is a set, we use Separation in V. Note that we must replace φ by its relativization to V_α, i.e., replacing each quantifier $\exists x$ with $\exists x \in V_\alpha$ to obtain a new formula $\hat{\varphi}$, and noting that

$$y = \{u \in x : \hat{\varphi}(u, \bar{p})\}.$$

So $y \in V_\alpha$. □

Recall that a strongly inaccessible cardinal is an uncountable regular limit cardinal κ such that for all $\lambda < \kappa$, $2^\lambda < \kappa$.

Theorem 7.6. *If κ is a strongly inaccessible cardinal, V_κ is a model of ZFC.*

Proof. We use the following fact from the homework:

Lemma 7.7. *If κ is a strongly inaccessible cardinal, then for all $\alpha < \kappa$, $|V_\alpha| < \kappa$.*

The only missing axiom is replacement. Given a formula $\varphi(x, y)$ with other suppressed parameters, suppose that

$$V_\kappa \vDash \forall x \forall y \forall z(\varphi(x, y) \wedge \varphi(x, z) \to y = z).$$

We must show that

$$V_\kappa \vDash \forall a \exists b \forall y \, (y \in b \longleftrightarrow \exists x \in a \, \varphi(x, y)).$$

Fix $a \in V_\kappa$. Choose $\beta < \kappa$ such that $a \in V_\beta$, and by transitivity $a \subseteq V_\beta$. We have $|a| \leq |V_\beta| < \kappa$. Define a map $f : a \to \kappa$ by

$$f(x) = \begin{cases} 0 & V_\kappa \vDash \neg \exists y \varphi(x, y) \\ \alpha & \alpha \text{ is least such that there is } y \in V_\alpha \text{ with } V_\kappa \vDash \varphi(x, y) \end{cases}.$$

The image of a under f cannot be cofinal in κ because κ is regular, so there is $\lambda < \kappa$ an upper bound on the image. This means that if $x \in a$ and $V_\kappa \vDash \exists y \, \varphi(x, y)$, then there is $y \in V_\lambda$ with $V_\kappa \vDash \varphi(x, y)$.

Let

$$b = \{y \in V_\lambda : V_\kappa \vDash (\exists x \in a)\varphi(x, y)\} = \{y \in V_\kappa : V_\kappa \vDash (\exists x \in a)\varphi(x, y)\}.$$

By Separation in V_κ, $b \in V_\kappa$. □

Corollary 7.8. *ZFC cannot prove the existence of a strongly inaccessible cardinal.*

Proof. Suppose that ZFC could prove the existence of a strongly inaccessible cardinal κ. Then ZFC could prove that there is a set model V_κ of ZFC. By the Soundness theorem, this proves $Con(ZFC)$, the statement that ZFC is consistent. But Gödels incompleteness theorem says that ZFC cannot prove $Con(ZFC)$, a contradiction. □

8 The Contructible Universe

Our next goal is to prove that the continuum hypothesis is consistent with ZFC. We do this by building a class model of ZFC where this continuum hypothesis holds. This model is called Gödel's constructible universe and is denoted L.

The idea is that we want to find a model L where the power set of ω (in L) is small enough to be $\aleph_1$. This means that we want to include in L as few subsets of ω as possible. Some subsets of ω must be included in L in order to satisfy the axioms of ZFC, for example the Axiom of Separation.

We will define $L = \bigcup L_\alpha$ similar to the way in which we defined the hierarchy V_α, except that instead of taking the power set at successor stages, we will take something called the definable power set which includes fewer subsets.

Definition 8.1. A set $y \subseteq x$ is a definable subset of x if there is a formula φ and parameters $p_1, \ldots, p_n \in x$ such that

$$y = \{z \in x : x \vDash \varphi(z, p_1, \ldots, p_n)\}.$$

The definable power set $\mathcal{P}_{def}(x)$ of x is the collection of all definable subsets of x.

Note that since we are asking that φ holds in x, this is essentially Δ_0 about x since (if x is transitive) all the quantifiers are over elements of x.

We do not yet know that $\mathcal{P}_{def}(x)$ is a set, because we cannot quantify over all formulas φ. To see that it is a set, we need to prove the following theorem in the metalanguage, which implies that $\mathcal{P}_{def}(x)$ exists by the Axiom of Separation.

Theorem 8.2. *There is a Δ_0 formula $\theta(x, y, z)$ such that for all sets a and b, with a transitive, $b = \mathcal{P}_{def}(a)$ if and only if $\exists z \theta(a, b, z)$.*

Proof sketch. Define the *basic Gödel functions* as follows:

1. $G_1(x, y) = \{x, y\}$.
2. $G_2(x, y) = x \times y$.
3. $G_3(x, y) = \{(u, v) \in x \times y : u \in v\}$.
4. $G_4(x, y) = x - y$.
5. $G_5(x, y) = x \cap y$.
6. $G_6(x) = \bigcup x$.
7. $G_7(x) = \mathrm{dom}(x)$.
8. $G_8(x) = \{(u, v) : (v, u) \in x\}$.
9. $G_9(x) = \{(u, v, w) : (u, w, v) \in x\}$.
10. $G_{10}(x) = \{(u, v, w) : (v, w, u) \in x\}$.

A *Gödel operation* is a composition of basic Gödel operations. Define $g(x) : V \to V$ to map x to the closure of x under all Gödel operations, i.e., the set we obtain by repeatedly applying the basic Gödel operations to elements of x. Then we must prove:

$$\mathcal{P}_{def}(M) = \{y \subseteq M : y \in g(M \cup \{M\})\}.$$

Then we need to write out a Δ_0 formula expressing this. □

Lemma 8.3. *If $|a| = \kappa \geq \omega$, then $|\mathcal{P}_{def}(a)| = \kappa$.*

Proof. There are countably many formulas in the language of set theory, and κ-many choice of finitely many parameters from a. So there are κ-many definable subsets of a. □

Lemma 8.4. *If x is transitive, then $\mathcal{P}_{def}(x)$ is transitive.*

Proof. Suppose that $z \in y \in \mathcal{P}_{def}(x)$. Then $y \subseteq x$ and so $z \in x$. Since x is transitive, $z \subseteq x$. Then

$$z = \{u \in x : u \in z\}$$

and so z is a definable set in x (with parameter z). □

Now we can carry through the same definition as the von Neumann hierarchy V_α, but with $\mathcal{P}_{def}$ at the successor stages.

Definition 8.5. Define the constructive hierarchy as follows:

$$\begin{aligned} L_0 &= \varnothing \\ L_{\alpha+1} &= \mathcal{P}_{def}(L_\alpha) \\ L_\lambda &= \bigcup_{\alpha<\lambda} L_\alpha \end{aligned}$$

Define $L = \bigcup L_\alpha$.

Each L_α is a set, but L is a proper class. By similar arguments as for V_α, we can prove that each L_α is a transitive set and that $L_\alpha \subseteq L_\beta$.

If M is a transitive model and $X \in M$, X being a well-order is not absolute for M; it could be that there is a set $Y \subseteq X$ in V which has no least element, but $Y \notin M$. However, if the linear order is $\in$, then we know by foundation that any subset of $(X, \in)$ has an $\in$-least element, so to say that $(X, \in)$ is a well-order, we just have to say that it is a linear order. This is Δ_0. So being an ordinal is absolute for transitive models.

Lemma 8.6. *If M is a transitive set, $y = \{\alpha \in M : \alpha$ is an ordinal$\}$ is a definable set in M.*

Proof. As described above, there is a Δ_0 formula $\varphi(x)$ such that $M \vDash \varphi(x)$ if and only if $V \vDash \varphi(x)$ if and only if x is an ordinal. □

Lemma 8.7. *For any ordinal α, $\alpha \in L_{\alpha+1}$ but not in L_α.*

Proof. Since $L_\alpha \subseteq V_\alpha$, $\alpha \notin L_\alpha$.

Note that $0 = \varnothing$ is definable in L_0, and hence is in L_1. Now argue by induction:

$$\alpha = \{x \in L_\alpha : x \text{ is an ordinal}\}.$$

By the induction hypothesis, the ordinals in L_α are exactly the ordinals $\beta < \alpha$. □

Theorem 8.8. *L is a model of ZF.*

Proof. We start with the easier axioms, and will need another theorem for separation and replacement.

Extensionality: This is the same as for V_α.

Pairing: Let $x, y \in L$, say $x, y \in L_\lambda$. Then $\{x, y\}$ is definable in L_λ by the formula

$$\varphi(z) \coloneqq \exists u \in z(u = x) \ \wedge \ \exists u \in z(u = y) \ \wedge \ \forall u \in z(u = x \vee u = y).$$

Union: Let $a \in L$, so that $a \in L_\lambda$ for some λ. Let $u \in \mathcal{P}_{def}(L_\lambda)$ be the set of $x \in L_\lambda$ such that $L_\alpha \vDash \exists y(z \in y \wedge y \in x)$. That is, u is the union of the collection x in L_λ, and $u \in L_{\lambda+1}$. We claim that it is the union of the collection x in L. Indeed, saying that u is the union of the collection x is a Δ_0 formula about u and x, and so it is absolute between $L_{\lambda+1}$ and L.

Foundation: This is the same as for V_α.

Infinity: By Lemma 8.7, $\omega \in L$. Again, expressing that ω is an inductive set is Δ_0, hence absolute for L.

Powerset: Let $x \in L$. Let $y = \mathcal{P}(x) \cap L$. First, we have to show that y is a set; indeed, there is a class function $y \to ON$ mapping each $a \subseteq x$ to the least ordinal α with $a \in L_\alpha$, or 0 if $a \notin L$. By the axiom of replacement, the image under this map is a set, and hence has an upper bound λ as ON is a proper class. So $y \subseteq L_\lambda$. Increasing λ, we may assume that $x \in L_\lambda$.

Now $y = \{u \in L_\lambda : u \subseteq x\}$ and so $y \in L_{\lambda+1} \subseteq L$ (recall that $u \subseteq x$ is absolute). The set y is the powerset of x in L.

Now we need the following theorem whose proof will follow:

Theorem 8.9 (Reflection theorem)**.** *For any formula φ and ordinal α, there is an ordinal $\beta > \alpha$ such that φ is absolute between L and L_β.*

Separation: Let $x, p_1, \ldots, p_n \in L$ and let φ be a formula. Let α be such that $x, p_1, \ldots, p_n \in L_\alpha$, and choose $\beta > \alpha$ such that φ is absolute between L and L_β. Note that $x \subseteq L_\beta$. Let

$$y = \{u \in x : L_\beta \vDash \varphi(u, p_1, \ldots, p_n)\} = \{u \in x : L \vDash \varphi(u, p_1, \ldots, p_n)\}$$

by absoluteness of φ. The first expression for y shows that y is definable in L_β, and hence in $L_{\beta+1}$. The second expression for y shows that y witnesses the Axiom of Separation in L.

Replacement: Let φ define a class function in L with parameters $p_1, \ldots, p_n \in L$:

$$L \vDash \forall x, y, z \ (\varphi(x, y, \bar{p}) \wedge \varphi(x, z, \bar{p}) \to y = z).$$

Let $a \in L$. By the Axiom of Replacement in V, there is α such that for all $x \in a$, if $L \vDash \exists y \varphi(x, y, \bar{p})$, then the unique witness $y \in L$ is in L_α (see the proof for Powerset). Increasing α, we may assume that $x, p_1, \ldots, p_n \in L_\alpha$. Let

$$b = \{y \in L_\alpha : L \vDash \exists x \in a \ \varphi(x, y, \bar{p})\}.$$

By Separation in L, $b \in L$. By choice of α,

$$b = \{y \in L : L \vDash \exists x \in a \ \varphi(x, y, \bar{p})\}.$$

□

Now we have to go back and prove the Reflection theorem:

Proof of the Reflection theorem. Let $\{\varphi_1, \ldots, \varphi_n\}$ list the subformulas of φ, including φ itself. The subformulas of φ are just the formulas that show up when building φ inductively according to the rules for building formulas. Define a class function $f : ON \to ON$ as follows. Given α, define $f(\delta)$ to be the least ordinal β such that for each $i = 1, \ldots, n$, if $\varphi_i(x_1, \ldots, x_k)$ is a formula with k free variables, and $a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_k \in L_\delta$, and

$$L \vDash \exists x_j \;\; \varphi(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_k),$$

then

$$L \vDash \exists x_j \in L_\beta \;\; \varphi(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_k).$$

Let $\alpha_0 = \alpha$, and $\alpha_{i+1} = f(\alpha_i)$. Let $\beta = \sup \alpha_i$. We claim, by induction on subformulas that each φ_i is absolute between L_β and L. This is clear for atomic formulas, and the absolute formulas are closed under $\neg$ and $\wedge$.

So suppose that φ_j is $\exists x_\ell \varphi_i(x_1, \ldots, x_k)$ and that φ_i is absolute between L_β and L. We must show that φ_j is absolute between L_β and L. (For simplicity, assume that $\ell = 1$.)

On the one hand, suppose that $a_2, \ldots, a_k \in L_\beta$ and

$$L_\beta \vDash \exists u \; \varphi_i(u, a_2, \ldots, a_k).$$

Then for some $a_1 \in L_\beta$,

$$L_\beta \vDash \varphi_i(a_1, a_2, \ldots, a_k).$$

Since φ_i is absolute between L_β and L,

$$L \vDash \varphi_i(a_1, a_2, \ldots, a_k)$$

and hence

$$L \vDash \exists u \; \varphi_i(u, a_2, \ldots, a_k).$$

On the other hand, suppose that $a_2, \ldots, a_k \in L_\beta$ and

$$L \vDash \exists u \; \varphi_i(u, a_2, \ldots, a_k).$$

Let α_ℓ be such that $a_2, \ldots, a_k \in L_{\alpha_\ell}$. Then

$$L \vDash (\exists u \in L_{\alpha_{\ell+1}}) \; \varphi_i(u, a_2, \ldots, a_k).$$

So there is $a_1 \in L_{\alpha_{\ell+1}} \subseteq L_\beta$ with

$$L \vDash \varphi_i(a_1, a_2, \ldots, a_k).$$

Since φ_i is absolute between L_β and L,

$$L_\beta \vDash \varphi_i(a_1, a_2, \ldots, a_k)$$

and so

$$L_\beta \vDash \exists u \; \varphi_i(u, a_2, \ldots, a_k). \qquad \square$$

Now our next goal is to show that L satisfies the axiom of choice. We will do this by giving a well-ordering of each L_α; since every set $x \in L$ has $x \in L_\alpha$ for some α, and so $x \subseteq L_\alpha$ by transitivity, this induces a well-ordering on every set in L.

We denote the ordering on L_α by $\leq_\alpha$. We want these orderings to fit together well for different values of α:

1. If $\alpha < \beta$, then $\leq_\beta$ extends $\leq_\alpha$.

2. If $\alpha < \beta$, $x \in L_\alpha$, and $y \in L_\beta - L_\alpha$, then $x <_\beta y$.

We build the orders $\leq_\alpha$ by induction on α. $L_0 = \varnothing$ and so has a trivial well-ordering.

At limit stages, $L_\lambda = \bigcup_{\alpha<\lambda} L_\alpha$ and so is well-ordered by $\bigcup_{\alpha<\lambda} \leq_\alpha$; it is here that we must use properties (1) and (2) above to argue that if each $\leq_\alpha$ is well-ordered, then so is $\leq_\lambda$. (This is on the homework.)

The real work comes at successor stages. Recall that we proved that

$$L_{\alpha+1} = \mathcal{P}_{def}(L_\alpha) = \{x \subseteq L_\alpha : x \in g(L_\alpha \cup \{L_\alpha\})\}$$

where $g(M \cup \{M\})$ is the closure of $M \cup \{M\}$ under the Gödel operations. Let $g_0, g_1, g_2, \ldots$ list the Gödel operations. So each $x \in L_{\alpha+1}$ is the image under a Gödel operation g_i of

Part II
Determinacy