21-602: Introduction to Set Theory I

Lecturer: Ernest Schimmerling

Scribe: Sidharth Hariharan

Carnegie Mellon University - Fall 2026

Contents

Chapter 1

A Recap of Undergraduate Set Theory

Set Theory can be viewed as a field of mathematics that encompasses all other mathematics. We can view set theory as beginning at the empty set $\emptyset$, also denoted V_0. We can construct subsequent sets $V_1, V_2, V_3, \ldots$ by taking power sets, ie, $V_1 = \mathcal{P}(V_0)$, $V_2 = \mathcal{P}(V_1)$, and so on. We can then define V_ω to be the union of all V_n, where $n < \omega$. We can then define $V_{\omega+1} = \mathcal{P}(V_\omega)$, and so on. All of analysis, the field of mathematics Davoud Cheraghi once described as the "rigorous study of infinite constructions", is essentially about $V_{\omega+1}$. In this way, set theory encompasses analysis, for example.

We call the hierarchy

$$V_0 \subset V_1 \subset \cdots \subset V_\omega \subset V_{\omega+1} \subset \cdots$$

the **universe of sets**. Indeed, this is often represented diagrammatically as a V that contains each of these Vs, with V_0 at the bottom, V_1 above V_0, and so on.

There is a problem with the universe of sets: *the power set operation is too strong*. That is, when we take the power set of a set, it "throws too many elements in". This leads to questions like the *continuum hypothesis*, which asks the question of whether $2^{\aleph_0} = |\mathbb{R}| = |V_{\omega+1}|$ is equal to $\aleph_1$.

Something we will study quite closely in this course is a slightly different version of this universe, known as **Gödel's Constructible Universe**. This is a universe of sets that is not built up from $\emptyset$ using the power set, but using a *different* operation $\mathcal{D}$, defined as follows: for a set A, viewing

it as a first-order structure $(A; \in)$,

$$\mathcal{D}(A) := \{x \subseteq A \mid x \text{ is 1st-order definable over } (A; \in) \text{ using parameters from } A\}$$

We can show that for finite sets, $\mathcal{D}$ agrees with $\mathcal{P}$. Moreover, we can show that for a countable set A, $\mathcal{D}(A)$ is countable as well. The point is that if we then define a universe

$$(L_0 = \emptyset) \subset (L_1 = \mathcal{D}(L_0)) \subset \cdots \subset \left(L_\omega = \bigcup_{n<\omega} L_n\right) \subset \cdots$$

we not only have agreement between L_n and V_n for all $n \leq \omega$, but we can build a model of ZFC in which the continuum hypothesis is true!

This is something we will explore in great detail in this course.

We note that the typical background setting for this course is the entirety of ZFC. However, as and when we need to either get rid of an axiom, or introduce a new one, we will mention it explicitly. We also use the following notational conventions.

Notation. We use **ZFC** to mean we have all the ZFC axioms. As and when we need variations, we do the following.

$$\textbf{ZF} := \text{“ZFC without the Axiom of Choice”}$$
$$\textbf{ZFC} - \textbf{F} := \text{“ZFC without the Axiom of Foundation”}$$
$$\textbf{ZF} - \textbf{F} := \text{“ZF without the Axiom of Foundation”}$$

We also fix some notation for functions.

Notation. For a function f, we write $\text{dom}(f)$ for its domain.

We use one further convention throughout.

Convention. When we write sup, we mean “union over”. Just like in Lean.

We will begin by reviewing the ZFC axioms.

1.1 The Zermelo-Fraenkel Axioms and the Axiom of Choice

We begin by reviewing the Zermelo-Fraenkel Axioms of Set Theory. We work in a first-order language with only one relation symbol, denoted $\in$. In principle, we would want to distinguish between the symbol in the language and its interpretation in any model, but we will not do this in practice. In situations where we really want to be pedantic, we will denote the symbol $\dot{\in}$ and the interpretation $\in$.

We begin with the Axiom of Extensionality.

1.1.1 Extensionality

ZFC Axiom 1 (The Axiom of Extensionality)**.**

$$\forall A \forall B \, [(\forall x \, (x \in A \leftrightarrow x \in B)) \rightarrow (A = B)]$$

We define the $\subseteq$ symbol, used infix as $A \subseteq B$, to be shorthand for the formula $\forall x \ (x \in A \rightarrow x \in B)$. This is not a symbol in the language of set theory; it is merely a convenient abbreviation that we will use hereafter without further discussion.

ZFC Axiom 1 essentially tells us that

$$\forall A \forall B \, ((A \subseteq B \wedge B \subseteq A) \rightarrow A = B)$$

(This is provably equivalent to ZFC Axiom 1.)

We now define the Axiom Scheme of Comprehension.

1.1.2 Comprehension

First, some notation.

Notation. We will denote finite tuples using overlines, so that $\bar{y}$ refers to $(y_0, \ldots, y_{n-1})$. We will typically use bars to bundle together variables.

ZFC is an infinite theory. We call the following an Axiom "Scheme" to underscore the fact that it

is really an infinite family of axioms bundled together.

ZFC Axiom 2 (The Axiom (Scheme) of Comprehension)**.** Let $x, \bar{y}$ be free variables. For all formulae $\varphi(x, \bar{y})$, the following is an axiom:

$$\forall A \forall \bar{y} \exists B \forall x \left[x \in B \leftrightarrow (x \in A \wedge \varphi(x, \bar{y})) \right]$$

Intuitively, this means we can define

$$B = \{x \in A \mid \varphi(x, \bar{y})\}$$

This is an axiom scheme because we can increase the arity of φ: we don't know how many variables y_i are bundled together in $\bar{y}$, only that for any given statement of the axiom, there are finitely many y_i.

Note that ZFC Axiom 1 tells us that we can replace the existential quantifier for B in ZFC Axiom 2 with an existence and uniqueness quantifier without changing its meaning.

Before proceeding further, we define the empty set. In first order logic, structures are required to have non-empty universes. Thus, for every structure (A, E) in the language of set theory, with $E \subseteq (A \times A)$ representing membership, we must have

$$(A, E) \vDash \exists x \, (x = x)$$

The Completeness Theorem then tells us that

$$\vdash \exists x \, (x = x)$$

ie, that the formula $\exists x \, (x = x)$ *must be a theorem* in the language of sets. Hence, ZFC Axiom 1 and ZFC Axiom 2 tell us that

$$\vdash \exists! A \, \forall x \ (x \notin A)$$

In other words, it is a theorem in the language of sets that there is a unique set that contains no members whatsoever. This unique set is denoted $\emptyset$, and is called the **empty set**. Its existence (and uniqueness) is not a distinct axiom, but a direct consequences of ZFC Axiom 1 and ZFC Axiom 2.

Next, we give the axiom of pairing.

1.1.3 Pairing

ZFC Axiom 3 (The Axiom of Pairing).

$$\forall x \, \forall y \, \exists A \, (x \in A \wedge y \in A)$$

ZFC Axiom 3 essentially gives us a way of constructing, for any x and y, the unique set $A = \{x, y\}$ with the property that

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

In particular, it allows us to construct, for any x, the set $\{x, x\}$, which, by ZFC Axiom 1, is exactly $\{x\}$. In other words, it tells us that we can stick any x into a set that only contains x.

This allows us to construct pairwise unions, but not arbitrary unions.

1.1.4 Unions

Now, we state the axiom of unions.

ZFC Axiom 4 (The Axiom of Unions). If we denote by $\mathcal{F}$ a family of sets,

$$\forall \mathcal{F} \, \exists B \, \forall A \, [A \in \mathcal{F} \rightarrow A \subseteq B]$$

As usual, we can use ZFC Axiom 1, ZFC Axiom 2 to show that

$$\bigcup \mathcal{F} = \{x \mid \exists A \in \mathcal{F} \, (x \in A)\}$$

is well (and uniquely) defined.

Next comes the power set axiom.

1.1.5 Power Sets

ZFC Axiom 5 (The Axiom of Power Sets).

$$\forall A \exists \mathcal{F} \forall X \ [X \subseteq A \to X \in \mathcal{F}]$$

Again, ZFC Axioms 1 and 2 tell us that this is equivalent to

$$\forall A \exists! \mathcal{F} \forall X \ (X \in \mathcal{F} \leftrightarrow X \subseteq A)$$

This justifies the definition

$$\mathcal{P}(A) := \{X \mid X \subseteq A\}$$

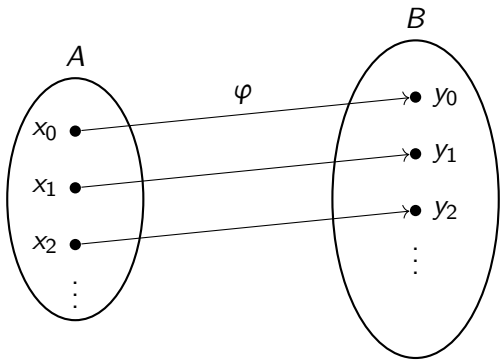
Next comes the second (and last) axiom scheme in ZFC, the axiom scheme of replacement.

1.1.6 Replacement

ZFC Axiom 6 (The Axiom (Scheme) of Replacement). For each formula $\varphi(x, y, \bar{z})$ that does not contain B as a free variable, the following is an axiom:

$$\forall A \forall \bar{z} \ [[\forall x \in A \exists! y \, \varphi(x, y, \bar{z})] \to [[\exists B \forall x \in A \exists y \in B \, \varphi(x, y, \bar{z})]]]$$

Each $x \in A$ has exactly one witness, and what the axiom buys us is that all of those witnesses fit inside a single set B.



B is drawn with room to spare because the axiom only asks for *some* set containing the witnesses; cutting it down to exactly them is what ZFC Axiom 2 does.

We can unpack this slightly complicated formula using ZFC Axiom 1 and ZFC Axiom 2: it essentially tells us that we can define

$$\{y \mid \exists x \in A\ (\varphi(x, y, \bar{z}))\}$$

Ie, if for every x, there is a unique witness y for which the formula $\varphi(x, y, \bar{z})$ holds, we can collect all these witnesses into a set that is contained B. In particular, "the number of such witnesses is *small enough* that it *can be collected into a set*."

Consider the case where $\varphi(x, y, \bar{z})$ defines a function with domain A. If we can define $A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\}$ (which we have yet to show we can do, but let's say we can), then we can put

$$f = \{(x, y) \in A \times B \mid \varphi(x, y, \bar{z})\}$$

which is a set by ZFC Axiom 2. Effectively, ZFC Axiom 6 tells us that *if the domain of a function is a set, then so is its range*. ZFC Axiom 6 is an essential step in the process of actually *defining* functions (or showing that this is something we can *do*).

One situation in which ZFC Axiom 6 is used (and will be used once sufficiently many axioms are stated) is in the situation we talked about at the start of Chapter 1. We'll see that there is a formula $\varphi(n, S)$ so that $\varphi(n, S) \leftrightarrow (n \in \omega \wedge S = V_n)$. Then, ZFC Axiom 6 would imply that

$$\langle V_n \mid n < \omega \rangle$$

really is a sequence. This allows us to define $V_\omega = \bigcup_{n<\omega} V_n$ and continue.

1.1.7 Foundation

ZFC Axiom 7 (The Axiom of Foundation).

$$\forall A\,(A \neq \emptyset \rightarrow \exists x \in A \forall y \in A\, y \notin x)$$

We now state the most interesting of the ZFC axioms: the Axiom of Choice.

1.1.8 Choice

ZFC Axiom 8 (The Axiom of Choice - (AC)).

$$\forall \mathcal{F}\, \exists c\, [\forall A \in \mathcal{F}\ (A \neq \emptyset \rightarrow A \in \mathrm{dom}(c) \wedge c(A) \in A)]$$

where $\mathcal{F}$ is a family of sets, c is a function, and $\mathrm{dom}(c)$ is the domain of c. We call c a **Choice Function on** $\mathcal{F}$.

The reason we state the Axiom of Choice at this point in our discussion is that modulo the ZFC Axioms that we have seen so far, the Axiom of Choice is equivalent to each of the following.

1. $\forall A\, \exists R \subseteq A \times A\ (A, R)$ is a well-ordering.
2. $\forall A\, \exists \alpha\, \exists f : A \xrightarrow{\sim} \alpha$, where α is an ordinal.

Note that none of the above axioms imply the existence of the natural numbers, without which virtually none of the mathematics we know and love would be possible. We therefore come to the axiom of infinity, which posits the existence of infinite sets, the 'smallest' of which is the natural numbers.

1.1.9 Infinity

ZFC Axiom 9 (The Axiom of Infinity).

$$\exists I\ (\emptyset \in I \wedge \forall x \in I\ (x \cup \{x\} \in I))$$

A set I as above is called an **inductive set**, and it is an axiom of ZFC that such a set exists.

It is possible to prove that given the existence of inductive sets (as posited above), there is a unique inductive set ω such that for all inductive sets J, $\omega \subseteq J$. Indeed, for any inductive set K, ω can be defined as the intersection of all inductive subsets $J \subseteq K$. It can then be shown that this intersection ω lives in all inductive sets. This minimal inductive set ω is known as the **set of natural numbers**. It is easy to see that the famed axioms of Peano are satisfied by ω.

Intuitively, we'd want to say something like

$$\omega = \bigcap \{k \mid k \text{ is inductive}\}$$

The only problem is, a priori, the above intersection is not a set! This is why we need this axiom.

1.2 Transitive Sets

This is a quick recap section on transitive sets and ordinals. We come to cardinals in Section 1.5.

Definition 1.2.1 (Transitive Set). A set A is **transitive** if $\forall y \in A\ (y \subseteq A)$, ie, if $\forall y \in A \forall x \in y\ (x \in A)$

1.2.1 Ordinals

Definition 1.2.2 (Ordinal). A set α is an **ordinal** iff α is a transitive set and $(\alpha, \in)$ is a well-ordering.

Note that writing $(A, \in)$ really means the structure (A, R) where $R = \{(x, y) \in A \times A \mid x \in y\}$. It's slightly problematic to call $\in$ itself a relation because it isn't a set: it's (strictly speaking) just a symbol, which is interpreted as a set (a relation) in a model of set theory. We have to be careful with these sorts of things.

Example 1.2.3 (Examples of Ordinals).

1. $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc.
2. $\omega = \{0, 1, 2, \ldots\}$
3. $\omega + 1 = \omega \cup \{\omega\}$, $\omega + 2$, etc.
4. $\omega + \omega$

and so on.

1.2.2 Mostowski Collapse for Well-Orderings

We will eventually prove this in more generality (Theorem 2.2.4), but here's a useful theorem.

Theorem 1.2.4 (Mostowski Collapse Theorem for Well-Orderings). *For all well-orderings (A, R), there is a unique ordinal α such that there is a unique isomorphism π from (A, R) to $(\alpha, \in)$ given by*

$$\pi(y) = \{\pi(x) \mid xRy\}$$

This is a special case of what's called (surprise surprise) the Mostowski Collapse Theorem. Some things we can do include removing the well/linear requirement on the ordering and replace "ordinal" with just "transitive set".

What the theorem above says is that the ordinals form a complete set of representatives for the class of all well-orderings up to isomorphism.

1.3 The V-Hierarchy

Definition 1.3.1 (The V-Hierarchy). Define $V_0 = \emptyset$, $V_{\alpha+1} = \mathcal{P}(V_\alpha)$ for any ordinal α, and $V_\beta = \bigcup_{\alpha<\beta} V_\alpha$ for any limit ordinal β.

One can show that $\alpha \leq \beta \to V_\alpha \subseteq V_\beta$.

We can ask ourselves what exactly ZFC is describing. The answer is, **the V-hierarchy**.

Interestingly,

$$V_\omega \vDash \textbf{ZFC} - \text{infinity} + \neg\text{infinity}$$

Even more interestingly, ZFC doesn't prove its own consistency (if it could, it would be inconsistent, by Gödel), but it comes exactly one axiom away: it can prove the existence of models of everything except for the axiom of infinity.

1.3.1 The Existence of V_ω

It is worth asking ourselves the following question: *Why does V_ω exist?* Or, more precisely, *why is V_ω a set?*

By induction, we know that each V_n is a set. We want to essentially see that the map $n \mapsto V_n$ is a *set* (ie, a *function*). Then, if we take the union of the range of this function, whose domain is the set ω whose existence is given by ZFC Axiom 9, we can define V_ω to be that union.

We invoke ZFC Axiom 6, the Axiom of Replacement. Let $\varphi(x, y)$ be the formula capturing the following facts:

1. y is a function with domain $\text{dom}(y) = x \cup \{x\}$.
2. If $\emptyset \in \text{dom}(y)$, then $y(\emptyset) = \emptyset$.
3. If $z \cup \{z\} \in \text{dom}(y)$, then $z \in \text{dom}(y)$ and $y(z \cup \{z\}) = \mathcal{P}(y(z))$.

The idea is that $y = \langle V_0, V_1, V_2, \ldots, V_n \rangle$[1] for some $n \in \omega$. Then,

$$\textbf{ZFC} \vdash \forall n \in \omega \; \exists! y \, \varphi(n, y)$$

Applying the Axiom of Replacement (ZFC Axiom 6), we know that the map that takes any $n \in \omega$ to its unique witness y for *n is, indeed, a function*. In other words,

$$\langle \langle V_m \mid m \leq n \rangle \mid n < \omega \rangle$$

is a set. From this, we can read off

$$\langle V_n \mid n < \omega \rangle$$

by taking the union of its range. Each $\langle V_m \mid m \leq n \rangle$ is a function with domain $n + 1$, ie, a set of ordered pairs, and each one is contained in the next, so the union of all of them is a function too: the one with domain $\bigcup_{n<\omega} (n + 1) = \omega$ that sends n to V_n.

Remark. Recall that $(V_\omega, \in)$ is shorthand for (V_ω, R), where $R = \{(x, y) \in V_\omega \times V_\omega \mid x \in y\}$. It is possible to show not only that

$$(V_\omega, \in) \vDash \text{ZFC without the Axiom of Infinity}$$

[1] By this, we mean it is a *finite sequence*, ie, a function from $n+1$ to something that takes 0 to V_0, 1 to V_1, and so on up to n.

but the stronger fact that

$$(V_\omega, \in) \vDash \text{ZFC without the Axiom of Infinity but with its } \textit{negation}$$

Thus,

$$\mathbf{ZFC} \vdash \text{ZFC without Infinity but with } \neg\text{Infinity is consistent}$$

Ie, from ZFC, we can deduce that the entirety of ZFC, except with the negation of the axiom of infinity instead of the axiom itself, is consistent.

Here, we end our discussion on the V-hierarchy, and proceed to consider a broader type of object than sets, known as classes.

1.4 Classes

Recall Russell's Paradox: there is no set of all sets. One reason for this (which relies on the Axiom of Foundation) is that if V is the set of all sets, then $V \in V$, so $\{V\}$ has no $\in$-minimal member, which contradicts the Axiom of Foundation. Without the Axiom of Foundation, if V is the set of all sets, then if we set

$$W = \{x \in V \mid x \notin x\}$$

(which is a set by the Axiom of Comprehension), then if $W \in W$, then $W \notin W$, and if $W \notin W$, then $W \in W$, which is a contradiction, meaning there can be no set of all sets.

Nevertheless, we really do want to write things like

$$V = \{x \mid x \text{ is a set}\} = \{x \mid x = x\}$$

We would also like to write things like

$$\mathbf{OR} = \{\alpha \mid \alpha \text{ is an ordinal}\} \qquad \text{and} \qquad \{V_\alpha \mid \alpha \in \mathbf{OR}\}$$

The only problem is, none of these are sets. Indeed, the only guarantee of the Axiom of Comprehension is that

$$\{x \in A \mid \varphi(x, \bar{y})\}$$

is a set; there is no guarantee whatsoever that

$$\{x \mid \varphi(x, \bar{y})\}$$

is a set. It sometimes is, but it isn't always.

There are two ways you can prove that $V = \{x \mid x = x\}$ is not a set. The first is that if it were, then $V \in V$ would hold, meaning $\{V\}$ would have no $\in$-minimal element, contradicting foundation. The second is that if it were, $W = \{x \in V \mid x \notin x\}$ would be a valid set, by comprehension. Then, $W \in W \leftrightarrow W \notin W$, a contradiction.

There are two perspectives we can take. The first is the **external perspective**: this is the perspective from which you would look at models of ZFC. There is also the internal perspective, where you are already inside a model of ZFC. Now this internal perspective also has a notion of logic and set theory inside of it. So there's formalisations of set theory inside models of set theory, and these formalisations have their own notions of models of set theory, and so it goes.

At some point, we will study these things in detail.

For the time being, we will use the term **class** to describe something we might want to express as a set, even if it cannot actually be a set for cardinality reasons.

Note that we would want sets to be classes. Indeed, anything of the form $\{x \mid \varphi(x)\}$, with $\varphi(x)$ a formula in the language of set theory, should be a class. Unfortunately, this is not generous enough: we want, for each set A, that

$$A = \{x \mid x \in A\} = \{x \mid \varphi(x, A)\}$$

to be included as well. In other words, we want to have formulae φ that not only allow us to substitute free variables with input sets x but also allow us to substitute more free variables with parameter sets A.

Formulae and parameters determine the class, but the converse is not true. In other words, if

$$\forall x \, (\varphi(x) \leftrightarrow \varphi'(x))$$

then

$$\{x \mid \varphi(x)\} = \{x \mid \varphi'(x)\}$$

For example, we know that

$$\varphi(x) \leftrightarrow (\varphi(x) \wedge x = x)$$

So you can define the same class using more than one formula.

1.4.1 The Class of Ordinals

Definition 1.4.1 (The Class of Ordinals)**.** Define

$$\mathbf{OR} := \{\alpha \mid \alpha \text{ is an ordinal}\}$$

We can show that **OR** is a class.

We can then say that the 'mapping'

$$\alpha \mapsto V_\alpha$$

is a **class function**: it takes elements of a class and maps them to elements of another class.

We can represent this mapping as

$$\{(\alpha, V_\alpha) \mid \alpha \in \mathbf{OR}\}$$

Now we can ask ourselves the following question: for which formulae φ is this set equal to $\{x \mid \varphi(x)\}$?

Let's give two such formulae based on the following.

Theorem 1.4.2. *$\forall \gamma \in \mathbf{OR}$, $\forall S \in V$, TFAE:*

1. $S = V_\gamma$
2. *$\exists g$ a function with the properties that*
   (a) *The domain of g is $\gamma + 1$*
   (b) $g(0) = \emptyset$

(c) $\forall \alpha < \gamma$, $g(\alpha + 1) = \mathcal{P}(g(\alpha))$

(d) $\forall \beta \leq \gamma$ *with* β *a limit ordinal,* $g(\beta) = \bigcup_{\alpha<\beta} g(\alpha)$

(e) $g(\gamma) = S$

3. *Every function* g *exhibiting properties (a)-(d) above also exhibits property (e).*

1.4.2 Well-Foundedness

We now define the class of well-founded sets.

Definition 1.4.3 (The Class of Well-Founded Sets)**.** Define

$$\mathbf{WF} := \bigcup_{\alpha \in \mathbf{OR}} V_\alpha$$

That is, **WF** consists of all x such that there is some ordinal α with $x \in V_\alpha$.

WF is a class.

We will eventually see that the axiom of foundation is equivalent to saying that $V = \mathbf{WF}$ (under **ZF** without foundation).

Proposition 1.4.4. *For all* $\gamma \in \mathbf{OR}$,

$(1)_\gamma$ $\forall \alpha \leq \gamma$, $V_\alpha \subseteq V_\gamma$

$(2)_\gamma$ V_γ *is a transitive set*

Proof. We argue by induction on ordinals γ.

For the base step ($\gamma = 0$), both $(1)_0$ and $(2)_0$ are clear.

For the successor step, if $\gamma = \beta + 1$,

$(1)_\gamma$ Let $\alpha \leq \beta$ and $x \in V_\alpha$. By $(1)_\beta$, $x \in V_\beta$. By $(2)_\beta$, $x \subseteq V_\beta$. So, $x \in \mathcal{P}(V_\beta) = V_{\beta+1} = V_\gamma$.

$(2)_\gamma$ Let $y \in V_\gamma = \mathcal{P}(V_\beta)$. Then, $y \subseteq V_\beta$. By $(1)_\gamma$, $y \subseteq V_\gamma$.

Finally, for the limit step, ie, if γ is a limit ordinal, then $(1)_\gamma$ and $(2)_\gamma$ both follow from the fact

that $\forall \beta < \gamma$, $(1)_\beta$ and $(2)_\beta$ both hold. (This is because in the limit case, constructing V_γ is just done by taking unions.) □

Definition 1.4.5 (Rank). Define, for $x \in \mathbf{WF}$, the **rank** of x, denoted $\text{rank}(x)$, to be the least $\alpha \in \mathbf{OR}$ such that $x \in V_{\alpha+1}$.

Indeed, this is equivalent to $\text{rank}(x)$ being the least $\alpha \in \mathbf{OR}$ such that $x \subseteq V_\alpha$.

Proposition 1.4.6. *The mapping*

$$x \mapsto \text{rank}(x)$$

is a class surjection from **WF** *to* **OR**.

We won't prove this, but we'll say it's easy enough to prove and leave it at that.

Proposition 1.4.7. $\forall \beta \in \mathbf{OR}$,

$$V_\beta = \{x \in \mathbf{WF} \mid \text{rank}(x) < \beta\}$$

Proof sketch. You can just do induction on β and apply the definition of the rank (Theorem 1.4.5). □

It turns out that we can tell whether a set lies in **WF** by looking only at its members.

Proposition 1.4.8. $\forall y \, (y \in \mathbf{WF} \leftrightarrow y \subset \mathbf{WF})$

Proof.

$\rightarrow$ Let $\beta = \text{rank}(y)$. Then $y \subseteq V_\beta \subset \mathbf{WF}$.

$\leftarrow$ Assume $y \subset \mathbf{WF}$. Define $f : y \to \mathbf{OR}$ by $f(x) = \text{rank}(x) + 1$. We're using replacement to know that $f \in V$.

Let $\beta = \sup_{x \in y} f(x)$. Then, $y \subseteq V_\beta$. So $y \in \mathcal{P}(V_\beta) = V_{\beta+1} \subset \mathbf{WF}$.

□

Members have strictly smaller rank than the sets that contain them.

Proposition 1.4.9.

$$\forall y \in \mathbf{WF},\ \forall x \in y,\ [x \in \mathbf{WF} \wedge \mathrm{rank}(x) < \mathrm{rank}(y)]$$

Proof. We know $x \in y \subseteq V_{\mathrm{rank}(y)} \subset \mathbf{WF}$, so $x \in \mathbf{WF}$. This also shows that $\mathrm{rank}(x)+1 \leq \mathrm{rank}(y)$, so $\mathrm{rank}(x) < \mathrm{rank}(y)$. □

From this, we can compute the rank of a set from the ranks of its members.

Proposition 1.4.10 (The Rank Formula)**.**

$$\forall y \in \mathbf{WF},\ \mathrm{rank}(y) = \sup_{x \in y}(\mathrm{rank}(x) + 1)$$

Proof. Put $\rho = \mathrm{rank}(y)$ and $\sigma = \sup_{x \in y}(\mathrm{rank}(x) + 1)$. By Theorem 1.4.9, if $x \in y$, then $\mathrm{rank}(x) < \mathrm{rank}(y) = \rho$, so $\mathrm{rank}(x) + 1 \leq \rho$. So $\sigma \leq \rho$.

It suffices to show that $y \subseteq V_\sigma$ to conclude that $\rho \leq \sigma$. Now $V_\sigma = \{x \mid \mathrm{rank}(x) < \sigma\}$. Every $x \in y$ has $\mathrm{rank}(x) + 1 \leq \sigma$, by the definition of σ as a supremum, and so $\mathrm{rank}(x) < \sigma$. So y is indeed a subset. □

It turns out that the ordinals are their own ranks.

Proposition 1.4.11.

$$\forall \beta \in \mathbf{OR},\ [\beta \in \mathbf{WF} \wedge \mathrm{rank}(\beta) = \beta]$$

Proof. We argue by induction on β. Base case is trivial. In the inductive case, $\beta \subseteq \mathbf{WF}$, so by previous prop, $\beta \in \mathbf{WF}$. So,

$$\mathrm{rank}(\beta) = \sup_{\alpha < \beta}(\mathrm{rank}(\alpha) + 1)$$

$$= \sup_{\alpha<\beta}(\alpha+1)$$
$$= \beta$$

□

It follows that each V_β contains exactly the ordinals below β.

Proposition 1.4.12.

$$\forall \beta \in \mathbf{OR},\ [V_\beta \cap \mathbf{OR} = \beta]$$

Proof. We know

$$\begin{aligned} V_\beta \cap \mathbf{OR} &= \{x \in \mathbf{WF} \mid \mathrm{rank}(x) < \beta\} \cap \mathbf{OR} \\ &= \{\alpha \in \mathbf{OR} \mid \mathrm{rank}(\alpha) < \beta\} \\ &= \{\alpha \in \mathbf{OR} \mid \alpha < \beta\} \\ &= \beta \end{aligned}$$

□

1.4.3 Well-Foundedness of Membership

We now say what it means for a relation on a class to be well-founded. Say A is a class and R is a class binary relation[2]. When we write (A, R), we really mean the class structure

$$(A, R \cap (A \times A))$$

Definition 1.4.13 (Well-Foundedness). We say that (A, R) is **well-founded** if for every nonempty set $S \subseteq A$, $\exists x \in S$ such that $\forall y \in S$, $\neg (y\, R\, x)$.
We say that (A, R) is **ill-founded** if it is not well-founded.

Membership on a well-founded set is, reassuringly, well-founded.

[2] ie, a binary relation on elements of the class A

Proposition 1.4.14. *$\forall A \in \mathbf{WF}$, $(A, \in)$ is well-founded.*

Proof. Let $S \subseteq A$, $S \neq \emptyset$. Denote

$$\alpha = \min_{x \in S} \operatorname{rank}(x)$$

This makes sense: every $x \in S$ is a member of $A \in \mathbf{WF}$, so $x \in \mathbf{WF}$ by Theorem 1.4.9 and has a rank, and **OR** is well-ordered. Pick $x \in S$ with $\operatorname{rank}(x) = \alpha$. If some $y \in S$ had $y \in x$, then $\operatorname{rank}(y) < \operatorname{rank}(x) = \alpha$ by Theorem 1.4.9 again, contradicting the minimality of α. So x is $\in$-minimal in S. □

Before going further, we give a name to the smallest transitive set containing a given set; that there is one is the content of a lemma below.

Definition 1.4.15 (Transitive Closure)**.** Given a set Y, the **transitive closure** of Y, denoted $\operatorname{trcl}(Y)$, is the smallest transitive set T such that $Y \subseteq T$. In particular, $\operatorname{trcl}(\{x\})$ is the smallest transitive set containing x as a member.

The converse of Theorem 1.4.14 fails, and the transitive closure is exactly what the failure is hiding.

Non-Example 1.4.16 (An Ill-Founded Relation)**.** One way Foundation can fail is if we have two sets $x \neq y$ such that $x = \{y\}$ and $y = \{x\}$.

In this case, $(x, \in)$ and $(y, \in)$ are both well-founded (just because x and y are both singletons). But $x, y \notin \mathbf{WF}$, because if they were, then they would *both* be, so they would both have ranks, and since $x \in y$, $\operatorname{rank}(x) < \operatorname{rank}(y)$ and similarly, since $y \in x$, $\operatorname{rank}(y) < \operatorname{rank}(x)$, which is impossible.

In this example, the transitive closure of x is $\{x, y\}$, which is also the transitive closure of y. Moreover, $(\{x, y\}, \in)$ is ill-founded.

For transitive sets, on the other hand, there is nothing to hide.

Proposition 1.4.17. *For every transitive set A, if $(A, \in)$ is well-founded, then $A \in \mathbf{WF}$.*

Proof. Fix A a transitive set and assume $(A, \in)$ is well-founded. To see that $A \in \mathbf{WF}$, it suffices to show that $A \subset \mathbf{WF}$.

Suppose not. Let $S = A \setminus \mathbf{WF}$. Then, S is a nonempty subset of A. As $(A, \in)$ is well-founded, there is some $x \in S$ that is $\in$-minimal in S. Then $x \notin \mathbf{WF}$, so $x \not\subset \mathbf{WF}$.

Pick $y \in x \setminus \mathbf{WF}$. Then, $y \in x \in A$. Since A is transitive, $y \in A$. Thus, $y \in A \setminus \mathbf{WF} = S$, and $y \in x$. This contradicts our choice of x. □

The transitive closure can be built from below, in ω steps.

Lemma 1.4.18. *Given a set Y, put*

$$Z_0 = Y$$
$$Z_{n+1} = \bigcup Z_n$$
$$Z_\omega = \bigcup_{n<\omega} Z_n$$

Then, $\mathrm{trcl}(Y) = Z_\omega$.

We can now deliver on the promise made earlier in this section: over $\mathbf{ZF} - \mathbf{F}$, Foundation says precisely that $V = \mathbf{WF}$.

Proposition 1.4.19 (Assuming $\mathbf{ZF} - \mathbf{F}$). *TFAE:*

(1) *Foundation: every nonempty S has an $\in$-minimal member*

(2) *For all A, $(A, \in)$ is well-founded*

(3) $V = \mathbf{WF}$

Proof. (1) $\iff$ (2) is clear: taking $A = S$ in (2) gives (1), and conversely a nonempty $S \subseteq A$ is in particular a nonempty set, so (1) hands us an $\in$-minimal member of S.

(3) $\implies$ (2) follows from Theorem 1.4.14.

Let's now assume (2) and prove (3). $\mathbf{WF} \subseteq V$, so we only need $V \subseteq \mathbf{WF}$. Let $x \in V$ and let $A = \mathrm{trcl}(\{x\})$. By Theorem 1.4.17, since A is transitive and $(A, \in)$ is well-founded, by our

assumption (2), we get that $A \in \mathbf{WF}$. So $x \in A \subset \mathbf{WF}$. Thus, $x \in \mathbf{WF}$. □

1.5 Cardinals and the Axiom of Choice

In this section, we discuss how big the levels of the V-hierarchy are. The answer is elementary at the bottom of the hierarchy, but by the time we reach $V_{\omega+1}$ it already depends on the Axiom of Choice. We end by defining two hierarchies of cardinals, only one of which needs choice, and computing the size of every level of the hierarchy in terms of the second.

1.5.1 The Size of V_ω

We begin with the finite levels.

Proposition 1.5.1. *For all $n < \omega$, the set V_n is finite.*

Indeed, their sizes are 0, 2^0, 2^{2^0}, and so on.

So V_ω is the first infinite level. It turns out to be countable, and there are two ways of seeing it.

Proposition 1.5.2. *V_ω is countable.*

Proof using AC. Write

$$V_\omega = \bigcup_{n<\omega} V_n$$

AC tells us that a countable union of countable sets is countable. □

Proof that avoids AC. By recursion on $n < \omega$, we can define $R_n \subseteq V_n \times V_n$ so that

(1) (V_n, R_n) is a well-ordering

(2) R_{n+1} end-extends R_n

Then, we'll be able to put $R_\omega = \bigcup_{n<\omega} R_n$ and show that R_ω is a well-ordering of type ω.

Since (V_n, R_n) is a well-ordering of a finite set, we may enumerate V_n as

$$a_0 \, R_n \, a_1 \, R_n \, \cdots \, R_n \, a_{k-1}$$

where $k = |V_n|$. Every member of $V_{n+1} = \mathcal{P}(V_n)$ is a subset of V_n, so we may write any $x, y \in V_{n+1}$ as

$$x = \{a_i \mid i \in I\} \qquad \text{and} \qquad y = \{a_i \mid i \in J\}$$

for unique $I, J \subseteq k$. Now, put $x \, R_{n+1} \, y$ iff

(i) $x, y \in V_n$ and $x \, R_n \, y$.

(ii) $x \in V_n$ and $y \in V_{n+1} \setminus V_n$.

(iii) $x, y \in V_{n+1} \setminus V_n$, and $i \in J$, where i is the largest natural number lying in exactly one of I and J.

Case (iii) makes sense because $x \neq y$ forces $I \neq J$, so some natural number lies in exactly one of them and there are only finitely many candidates. The relation it defines is the ordering of the subsets of V_n by binary numerals, reading x as the string whose ith digit records whether $i \in I$, so it is a linear ordering.

So R_{n+1} is R_n followed by a linear ordering of the finitely many sets in $V_{n+1} \setminus V_n$, which makes it a well-ordering and gives us (1). It agrees with R_n on V_n and places all of V_n before all of $V_{n+1} \setminus V_n$, so V_n is an initial segment of (V_{n+1}, R_{n+1}), which gives us (2).

The point of the construction is that R_{n+1} is defined outright from R_n, with no choices made along the way, and the recursion starts from $R_0 = \emptyset$ on $V_0 = \emptyset$. The R_n increase, so R_ω linearly orders V_ω. Chaining (2) makes V_n an initial segment of (V_m, R_m) for every $m \geq n$, so each $x \in V_\omega$ has all of its R_ω-predecessors inside the finite V_n it came from, and every proper initial segment of R_ω is therefore finite. Since $\omega \subseteq V_\omega$, the ordering is infinite, so sending each x to its number of predecessors is an isomorphism onto ω. $\square$

1.5.2 Cardinality Without Choice

The reason the second proof is worth having is that without AC, countable unions are badly behaved.

Remark. If ZF is consistent, then so is ZF + "there's a countable union of pairs that has no cardinality".

The same problem arises one level further up.

Remark. Without AC, we don't know whether the set $V_{\omega+1} = \mathcal{P}(V_\omega)$ even has a *cardinality*. With AC,

$$|V_{\omega+1}| = |\mathcal{P}(V_\omega)| = 2^{\aleph_0}$$

where $\aleph_0 = \omega = |V_\omega|$.

1.5.3 Cardinal Successors

It turns out that the successor of a cardinal can still be defined without choice, provided we go about it the right way.

Given AC, we know that $2^\kappa > \kappa$ and we can define κ^+ to be the least cardinal strictly greater than κ. Without AC, we still know, given an infinite cardinal κ, that

$$\{R \subseteq \kappa \times \kappa \mid (\kappa, R) \text{ is a well-ordering}\}$$

is a set. We can still have Mostowski Collapse for well-orderings. We can then conclude that

$$\{\text{types}(\kappa, R) \mid (\kappa, R) \text{ is a well-ordering}\}$$

is a set of ordinals. Then, we can set κ^+ to simply be the supremum of this set.

1.5.4 The Alephs and the Beths

We can now define the two standard hierarchies of cardinals, of which only the second needs the Axiom of Choice.

Without AC, define

$$\begin{aligned}\aleph_0 &:= \omega \\ \aleph_{\alpha+1} &:= (\aleph_\alpha)^+ \\ \aleph_\beta &:= \sup_{\alpha<\beta}(\aleph_\alpha) \text{ if } \beta \text{ is a limit ordinal}\end{aligned}$$

With AC, define

$$\begin{aligned}\beth_0 &:= \omega \\ \beth_{\alpha+1} &:= 2^{\beth_\alpha} \\ \beth_\beta &:= \sup_{\alpha<\beta}(\beth_\alpha) \text{ if } \beta \text{ is a limit ordinal}\end{aligned}$$

It's clear `sorry`

With the $\beth$s in hand, we can say exactly how big every level of the hierarchy is, from V_ω onwards.

Proposition 1.5.3.

$$\forall \eta \in \mathbf{OR},\ |V_{\omega+\eta}| = \beth_\eta$$

Proof using AC. Do induction on η. For the base case, $|V_\omega| = \aleph_0 = \beth_0$, as V_ω is countable.

For the successor step, if $|V_{\omega+\eta}| = \beth_\eta$, then

$$|V_{\omega+\eta+1}| = |\mathcal{P}(V_{\omega+\eta})| = 2^{|V_{\omega+\eta}|} = 2^{\beth_\eta} = \beth_{\eta+1}$$

Finally, let η be a limit ordinal. Then $\omega + \eta$ is a limit ordinal too, and every ordinal below it is either below ω or of the form $\omega + \alpha$ for some $\alpha < \eta$, so

$$V_{\omega+\eta} = \bigcup_{\alpha<\eta} V_{\omega+\alpha}$$

Each $V_{\omega+\alpha}$ is a subset of $V_{\omega+\eta}$, so $\beth_\alpha \leq |V_{\omega+\eta}|$ for every $\alpha < \eta$, and hence $\beth_\eta = \sup_{\alpha<\eta} \beth_\alpha \leq |V_{\omega+\eta}|$. Conversely, $V_{\omega+\eta}$ is a union of $|\eta|$ sets, each of cardinality at most $\beth_\eta$, so (using AC) $|V_{\omega+\eta}| \leq |\eta| \cdot \beth_\eta$. It remains to see that $|\eta| \leq \beth_\eta$, and in fact $\alpha \leq \beth_\alpha$ for every ordinal α: this is clear for $\alpha = 0$, it passes through limits by taking suprema, and it passes through successors

because $\beth_{\alpha+1} = 2^{\beth_\alpha} > \beth_\alpha \geq \alpha$ forces $\beth_{\alpha+1} \geq \alpha + 1$. So $|V_{\omega+\eta}| \leq \beth_\eta \cdot \beth_\eta = \beth_\eta$. □

A word of caution before we go on.

Warning. If we continue to naïvely indulge ourselves the idealistic notion that we can use $\aleph$s to denote cardinals and ωs to denote the corresponding ordinals, we will soon encounter a world of confusion and pain.
The two notations will often be used interchangeably, just as they are in practice. We will try our best, but we might not always succeed.

The proposition indexes the levels from V_ω rather than from V_0, but the shift by ω is invisible from ω^2 onwards.

Proposition 1.5.4 (without AC).

$$\forall \eta \geq \omega^2,\ (\omega + \eta = \eta)$$

Note that $\omega^2 = \omega \cdot \omega$; by ordinal arithmetic, it's a countable ordinal.

Proof. First, write $\eta = \omega^2$. Then,

$$\omega + \omega^2 = \omega\,(1 + \omega) = \omega \cdot \omega = \omega^2$$

Now, in general, for $\eta = \omega^2 + \delta$, then $\omega + (\omega^2 + \delta) = (\omega + \omega^2) + \delta = \omega^2 + \delta$. □

Chapter 2

Induction, Recursion and Collapse

In Chapter 1, we saw that membership is a well-founded relation on **WF**, and that Foundation says precisely that this is true everywhere. In this chapter, we study well-founded relations on classes in general, with a view to doing induction and recursion along them. We begin with a second property a class relation may or may not have, which will be needed alongside well-foundedness.

2.1 Induction and Recursion

In what follows, let A be a class and R be a class relation on A. We want to prove things about the members of A by induction along R and define things on them by recursion along R, and this section says what R has to look like for that to work, and then does it.

2.1.1 Set-Like Relations

The property we want is that looking below any one point of A only ever sees a set's worth of A.

Definition 2.1.1 (Set-Like). We say (A, R) is **set-like** iff

$$\forall y \in A, \{x \in A \mid x \, R \, y\}$$

is a set.

The example we care about most has this property for the most basic of reasons.

Example 2.1.2 (Membership). $(A, \in)$ is set-like for every class A.

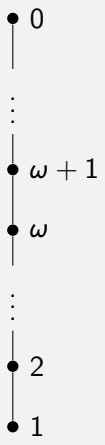
In fact, the fact in the above example is **equivalent** to the axiom scheme of comprehension (ZFC Axiom 2)! Indeed, if $A = \{x \mid \varphi(x, \bar{z})\}$ and $y \in A$, then

$$\{x \in A \mid x \in y\} = \{x \in y \mid \varphi(x, \bar{z})\}$$

So "membership is set-like" is a fancy way of saying "Comprehension".

But not all relations are set-like.

Non-Example 2.1.3. Let $A = \mathbf{OR}$. Define R by the following picture:



That is, it looks a bit like $(\mathbf{OR} \setminus \{0\}, <)$.

2.1.2 Proofs by Induction

We now have another *scheme* of facts. In what follows, start with A a class and $R \subseteq A \times A$ a class binary relation. Assume (A, R) is well-founded and set-like.

The first fact is that well-foundedness, which speaks of subsets of A, extends to subclasses.

Proposition 2.1.4 (Minimal Members of Subclasses). *If S is a nonempty subclass of A, then S has an R-minimal member.*

Proof. We may assume that $R \subseteq A \times A$. Since $S \neq \emptyset$, we have some $z \in S$. Let T be the

R-closure of $\{z\}$. That means that we can define, by recursion on $n < \omega$, the sets

$$T_0 = \{z\} \qquad T_{n+1} = \bigcup_{y \in T_n} \{x \mid x\, R\, y\}$$

The fact that (A, R) is set-like tells us that the union above is a union of sets, and therefore, a set.

Next, put $T = \bigcup_{n<\omega} T_n$. This is the smallest R-closed set that has z as a member.

Now, $S \cap T$ is a nonempty subset of A. So it has an R-minimal member, say x. Here, using the fact that T is R-closed, it is easy to show that x is R-minimal among members of S. Otherwise, we'd have $y\, R\, x$, with $y \in S$. Since $x \in T$ and T is R-closed, we have $y \in T$. This contradicts R-minimality of x in $S \cap T$. □

From this we get induction along R.

Theorem 2.1.5 (Theorem Scheme about Proofs by Induction). *Let $\varphi(u, w)$ be a formula in the language of set theory, and assume z is a set.*
Assume that for all y, if y satisfies the formula that defines A (that is, $\forall y \in A$),

$$[(\forall x\, R\, y,\ \varphi(x, z)) \to \varphi(y, z)]$$

We can then conclude that $\forall x \in A,\ \varphi(x, z)$.

Proof. Deny for contradiction. Let $S = \{x \in A \mid \neg\varphi(x, z)\}$. By the previous proposition, as S is a nonempty subclass of A, S has an R-minimal member, say y. Then, $\forall x\, R\, y,\ \varphi(x, z)$.

By assumption, $\varphi(y, z)$, ie, $y \notin S$. This is a contradiction. □

2.1.3 Definitions by Recursion

Notice that in mathematics, we define things by recursion and prove things by induction. But when we explain how everything works, we usually go backwards!

Theorem 2.1.6 (Theorem Scheme about Definitions by Recursion). *Let $F : A \times V \to V$ be a class function. Then, there is a class function $G : A \to V$ such that $\forall y \in A$,*

$$G(y) = F(y, G \restriction \{x \mid x \, R \, y\})$$

Moreover, G is unique in the sense that if $H : A \to V$ is any class function such that

$$\forall y \in A,\ H(y) = F(y, H \restriction \{x \mid x \, R \, y\})$$

then $\forall x \in A,\ H(x) = G(x)$.

Warning. Important subtlety: the statement "there exists a class function" is not a valid statement in the language of set theory, as we cannot quantify over *classes*. What *is* valid, though, is saying something like "such and such a G satisfies such and such a property", where the definition shall come from the proof below.

The proof builds G out of set-sized pieces, which deserve a name.

Definition 2.1.7 (Approximation). We call g an **approximation** iff g is a set function whose domain is an R-closed subset of A (ie, if $y \in \mathrm{dom}(g)$ and $x \, R \, y$, then $x \in \mathrm{dom}(g)$) and $\forall y \in \mathrm{dom}(g)$, $g(y) = F(y, g \restriction \{x \mid x \, R \, y\})$.

Proof. We'll prove uniqueness and then existence.

Uniqueness. Suppose g and h are approximations. Then, $\forall y \in \mathrm{dom}(g) \cap \mathrm{dom}(h)$, $g(y) = h(y)$.

You can prove this using R-induction.

Existence. We'll show that for all $z \in A$, there is some approximation g_z with $\mathrm{dom}(g_z)$ being the R-closure of $\{z\}$. These g_z, if we can show them to exist, are unique (by above).

Use R-induction. The induction hypothesis says that we have g_y as required for each $y \, R \, z$. Let $h = \bigcup_{y \, R \, z} g_y$. We can see that $\mathrm{dom}(h)$ is the R-closure of $\{y \mid y \, R \, z\}$.

Put $g_z = h \cup \{(z, F(z, h \restriction \{x \mid x \, R \, z\}))\}$.

This works! Then, put $G = \bigcup_{z \in A} g_z$.

Indeed, h is a set, by Replacement, since $y \mapsto g_y$ is a class function on the set $\{y \mid y\, R\, z\}$; it is a function, because any two of the g_y agree on their common domain by uniqueness; and it is an approximation, because a union of R-closed sets is R-closed and each g_y satisfies the defining equation on its own domain. Moreover, $z \notin \mathrm{dom}(h)$, as otherwise z would lie on an R-cycle, which well-foundedness forbids. So g_z is a function, its domain $\mathrm{dom}(h) \cup \{z\}$ is the R-closure of $\{z\}$, and it satisfies the defining equation at z by construction and below z because h does.

Finally, G is a class function because the g_z agree wherever their domains overlap, again by uniqueness, and every $y \in A$ lies in $\mathrm{dom}(g_y)$, so $G(y) = g_y(y) = F(y, g_y \restriction \{x \mid x\, R\, y\}) = F(y, G \restriction \{x \mid x\, R\, y\})$, since G agrees with g_y on $\mathrm{dom}(g_y) \supseteq \{x \mid x\, R\, y\}$. Uniqueness of G in the sense of the theorem is R-induction once more: if H satisfies the same equation, then $H(x) = G(x)$ for all $x \in A$.

□

The proof also tells us how to compute G, which is what makes "$G(x) = s$" a statement in the language of set theory after all.

Lemma 2.1.8. *Given any $x \in A$ and set s, TFAE:*

1. $G(x) = s$
2. *There exists an approximation g with $x \in \mathrm{dom}(g)$ and $g(x) = s$.*
3. *For all approximations g, if $x \in \mathrm{dom}(g)$ then $g(x) = s$.*

2.2 Mostowski Collapse

In Theorem 1.2.4 we saw that every well-ordering is isomorphic to a unique ordinal, by a map that sends each point to the set of images of its predecessors. In this section we prove the general version promised there: the ordering hypothesis is dropped, and "ordinal" becomes "transitive class". One more property of the relation is needed to make the map injective, and we begin with it.

2.2.1 Extensional Relations

Definition 2.2.1 (Extensionality). (A, R) is **extensional** iff $\forall y, z \in A$,

$$\{x \in A \mid x \, R \, y\} = \{x \in A \mid x \, R \, z\} \to y = z$$

This is just like the axiom of extensionality, but worded as a property of sets and relations.

Not all relations are extensional.

Non-Example 2.2.2 (Non-Extensional Relations). If $A = \{\{1\}, \{2\}\}$, then $(A, \in)$ is not extensional. The reason is that

$$\{x \in A \mid x \in \{1\}\} = \emptyset = \{x \in A \mid x \in \{2\}\}$$

but obviously $\{1\} \neq \{2\}$.

But many are.

Example 2.2.3 (Extensional Relations).

1. If M is a transitive class, then $(M, \in)$ is extensional. Indeed, suppose $y, z \in M$ and $y \neq z$. Since M is transitive, $y, z \subseteq M$, so $y = y \cap M$ and $z = z \cap M$. Thus

$$y \cap M = \{x \in M \mid x \in y\} \neq \{x \in M \mid x \in z\} = z \cap M$$

2. If (A, R) is a class linear ordering (ie, a linear ordering on the class A), then (A, R) is extensional.
   For instance, if $A \subseteq \mathbf{OR}$ and R is the usual ordering of ordinals (write R as $\in$ or $<$), then (A, R) is extensional.

2.2.2 The Collapse Theorem

We are now ready to state and prove Mostowski Collapse. Note that it will use neither Choice nor Foundation.

Theorem 2.2.4 (Mostowski Collapse Theorem Scheme). *Let (A, R) be a class structure that's well-founded, set-like and extensional. Define, using the theorem scheme on definitions by recursion, the mapping $G : A \to V$ by*

$$\forall y \in A, G(y) = \{G(x) \mid x \in A \text{ and } x \mathrel{R} y\}$$

Write the range of G as M. Then,

1. *M is a transitive class*
2. *G is a class isomorphism from (A, R) to $(M, \in)$*
3. *If M' is a transitive class and $G' : (A, R) \simeq (M', \in)$, then $M' = M$ and $G' = G$.*

*Note that G is referred to as the **Mostowski Collapse Isomorphism**.*

Remark. When we say that $M' = M$ and $G' = G$ above, we are not saying that their defining formulae are the same. In other words, their equality is not definitional.

Proof. We may assume $R \subseteq A \times A$.

1. Say $a \in b \in M$. We must see $a \in M$. Then, $b \in M$. As M is the range of G, there's some $y \in A$ such that $b = G(y)$. As $G(y) = \{G(x) \mid x \mathrel{R} y\}$, there's some $x \mathrel{R} y$ such that $a = G(x)$. Now $a = G(x) \in M$.

2. We need to show G is an order-preserving bijection.

   Bijectivity. It's clear that G is surjective, because M is the range of G. To show G is injective, it is enough to show that for every $x \in A$,

   $$\forall x' \in A\,[G(x) = G(x') \to x = x'] \tag{2.2.1}$$

   Let's prove (2.2.1) by induction on R. Let $y \in A$ and assume (2.2.1) holds for all $x \mathrel{R} y$. We'll now prove that (2.2.1) holds for y as well.

   Fix any $y' \in A$ and assume that $G(y) = G(y')$. We need to prove $y = y'$. Because (A, R) is extensional, it is enough to show that

   $$\{x \mid x \mathrel{R} y\} = \{x' \mid x' \mathrel{R} y'\}$$

($\subseteq$) Suppose $x\,R\,y$. Then

$$G(x) \in G(y) = G(y') = \{G(x') \mid x'\,R\,y'\}$$

So there's some $x'\,R\,y'$ with $G(x) = G(x')$. By (2.2.1) for x, we know that $x = x'$. Thus $x\,R\,y'$.

($\supseteq$) Suppose $x'\,R\,y'$. Then,

$$G(x') \in G(y') = G(y) = \{G(x) \mid x\,R\,y\}$$

So there's some $x\,R\,y$ such that $G(x') = G(x)$. By (2.2.1) for x, $x'\,R\,y$.

Monotonicity. Suppose $G(x) \in G(y)$. By definition of $G(y)$, there is some $x'\,R\,y$ such that $G(x) = G(x')$. As G is injective, we know that $x = x'$. Thus, $x\,R\,y$.

3. Suppose M' is a transitive class and $G' : (A, R) \simeq (M', \in)$. We need $M = M'$ and $G = G'$.

   Since M is the range of G and M' is the range of G', it is enough to show that $G = G'$. By the uniqueness of recursive definitions (Theorem 2.1.6), it is enough to see that G' satisfies the same defining equation as G. That is, $G'(y) = \{G'(x) \mid x\,R\,y\}$.

   ($\supseteq$) $x\,R\,y \to G'(x) \in G'(y)$ just because G' preserves order (by assumption).

   ($\subseteq$) Suppose $a \in G'(y)$. Now $G'(y) \in M'$ trivially because M' is the range of G'. Since M' is a transitive class, $G'(y) \subseteq M'$. Hence, $a \in M'$. So we have some $x \in A$ such that $a = G'(x)$. Now $G'(x) \in G'(y)$. Since G' is an isomorphism, this implies $x\,R\,y$.

□

2.2.3 Examples

Let's study some examples. The first collapses a proper class, and what it collapses onto is a class we have had since Chapter 1.

Example 2.2.5 (Sequences of Ordinals). Consider the following classes:

$$^{<\omega}\mathbf{OR} = \{f \mid \exists n < \omega,\ f : n \to \mathbf{OR}\}$$

$$\mathbf{OR}^{<\omega} = \{t \mid \exists n < \omega,\ t \text{ is an } n\text{-tuple of ordinals}\}$$

$$[\mathbf{OR}]^{<\omega} = \{s \mid \exists n < \omega, |s| = n \text{ and } s \subset \mathbf{OR}\}$$

The above are pairwise in bijection. No parameters are necessary for anything here.

Let's consider $[\mathbf{OR}]^{<\omega}$. Say $s \in [\mathbf{OR}]^{<\omega}$. Say $n < \omega$ and $s \in [\mathbf{OR}]^n$. List s in decreasing order as $s_0 > s_1 > \cdots > s_{n-1}$, where $s = \emptyset$ if $n = 0$. Identify s with the n-tuple $(s_0, \ldots, s_{n-1})$. Let $<_{\text{lex}}$ denote the lexicographic ordering on members of $[\mathbf{OR}]^{<\omega}$ under this identification, where a tuple counts as smaller than any tuple properly extending it. We see that

- $<_{\text{lex}}$ is a class well-ordering of $[\mathbf{OR}]^{<\omega}$ that is definable without parameters
- $\left([\mathbf{OR}]^{<\omega}, <_{\text{lex}}\right)$ is well-founded and extensional
- $<_{\text{lex}}$ is set-like: given $s \in [\mathbf{OR}]^{<\omega}$,

$$\left\{t \in [\mathbf{OR}]^{<\omega} \;\middle|\; t <_{\text{lex}} s\right\} \subseteq [\max(s) + 1]^{<\omega}$$

We can apply the Mostowski Collapse Theorem Scheme (Theorem 2.2.4) to get a map

$$G : \left([\mathbf{OR}]^{<\omega}, <_{\text{lex}}\right) \simeq (\mathbf{OR}, <)$$

In particular, G is a class bijection between $[\mathbf{OR}]^{<\omega}$ and $\mathbf{OR}$ that is definable without parameters.

The second example is less a computation than a technique, one we will keep coming back to.

Example 2.2.6 (ZFC). Consider any $\alpha \in \mathbf{OR}$. Recall that V_α is a transitive set, so $(V_\alpha, \in)$ is extensional, set-like and well-founded.

Now, you can formalize all of mathematics in ZFC, including Model Theory—in particular, the downwards Löwenheim-Skolem Theorem (DLS). (We will do this eventually in this course.)

Indeed, within the ZFC formalization, we can define the notion of an **elementary sub-**

structure as follows: we say $X \preceq Y$ iff for every "formula"[a] $\varphi(\bar{u})$ and every matching $\bar{a}$ from X, $X \vDash \varphi(\bar{a})$ iff $Y \vDash \varphi(\bar{a})$. You can imagine that with similar handwaving, we can formalize other notions, like DLS.

Once we have done all of this, we can prove that there is some countable X such that $(X, \in) \preceq (V_\alpha, \in)$. Moreover, note that

- if $\alpha \leq \omega$ then $X = V_\alpha$
- if $\alpha \geq \omega + 1$, then $X \subsetneq V_\alpha$ (X is countable but V_α is uncountable).

Observe that $(X, \in)$ is

well-founded: $X \subseteq V_\alpha \to X \in V_{\alpha+1} \subset \mathbf{WF}$

set-like: clear

extensional: we know that V_α is transitive, hence extensional. This is equivalent to the following fact (*within the ZFC-formalism of ZFC*):

$$V_\alpha \vDash \text{"Extensionality Axiom"}$$

Then, since $X \preceq V_\alpha$,

$$X \vDash \text{"Extensionality Axiom"}$$

as well. This is equivalent to $(X, \in)$ being extensional.

We can also argue more directly for extensionality, which asks that any two distinct members of X be told apart by a member of X. Say $b, c \in X$ with $b \neq c$. Since V_α is extensional, some $a \in V_\alpha$ lies in the symmetric difference $b \triangle c$, so that

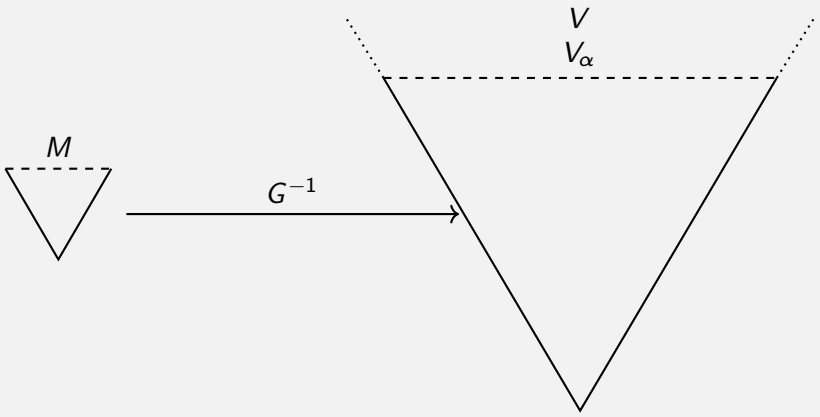
$$V_\alpha \vDash \exists a\,(a \in b \triangle c)$$

This is a "formula" whose parameters b and c both come from X, and $X \preceq V_\alpha$, so X satisfies it too. The witness it supplies is a member of X lying in $b \triangle c$, which is exactly what we wanted.

Let $G : (X, \in) \simeq (M, \in)$ be the Mostowski Collapse Isomorphism (ie, if G is the Mostowski Collapse Isomorphism, then $(M, \in)$ is the Mostowski Collapse—the range of G). As this is an isomorphism, M is transitive. Note that $G^{-1} : M \to V_\alpha$ is an elementary embedding (we compose $G^{-1} : M \simeq X$ with the inclusion $X \hookrightarrow V_\alpha$). In particular, M and V_α satisfy

the same "sentences".

Intuitively, what's going on is that we have built some countable microcosm—some tiny, countable analog—of the entire V-universe, *inside of the ZFC formalism*. This is one of the most important subroutines in set theory: the idea that you can take a big initial piece of V, take some elementary substructure via DLS, take its Mostowski Collapse, and do something with it. We will see a lot of this later on.



[a] There are no formulae because we're in a formalism where everything's a set! But there are "formulae": just make air quotes and change your tone of voice and everything is ok :D

Chapter 3

Relativization

The last two chapters were about sets, classes and the relations on them. This one is about formulae: what it takes for one to be true *in* a class rather than in a set, and how much of the universe a formula has to look at before we can tell.

We built **WF** in Chapter 1 and saw that Foundation says exactly that every set lies in it. To ask whether the rest of the axioms are true in **WF**, we need a notion of truth in a class, and satisfaction will not do: it is a notion about sets. Relativization is what replaces it.

3.1 Relativization to a Class

Relativization is essentially like satisfaction, just for classes. Our ultimate goal is the following: given (A, R) a class structure and a formula ψ, we'd want to define $\psi^{(A,R)}$ to be the relativization (think "interpretation") of ψ to (A, R).

Here are some complications with this technique:

- R might not be $\in$
- There might be parameters involved in the definitions of A and R. For instance, if A is a set, then "$A = \{x \mid x \in A\}$" takes in a set A as a parameter.

Look at $A = \{x \mid \varphi(x)\}$. As a scheme, define $\psi^A(\bar{v})$ (with $\bar{v}$ denoting free variables) by recursion

on the complexity of $\psi(\bar{v})$ as follows. We will have atomic, propositional and quantifier steps.

Before starting, rename the variables of φ other than x so that none of them occurs in ψ. Write φ' for the result, so that $A = \{x \mid \varphi'(x)\}$ still, and use φ' in place of φ throughout. Both directions of the clash matter. The quantifier step below substitutes a variable u of ψ for the free variable x, so a u bound in φ would be captured and $\varphi(u)$ would say something other than "$u \in A$"; and what that substitution produces is then dropped underneath the quantifiers of ψ, so a parameter of φ sharing its name with a variable that ψ binds would be captured the other way.

- Atomic: We interpret equality and membership symbols as equality and membership. Ie,

$$(v_0 = v_1)^A \text{ is } v_0 = v_1$$
$$(v_0 \in v_1)^A \text{ is } v_0 \in v_1$$

- Propositional: If ψ is a propositional combination of ψ_0 and ψ_1, then ψ^A is the same propositional combination of ψ_0^A and ψ_1^A.

- Quantifiers: We relativize a quantifier by restricting it to A. Ie,

$$(\forall u\, \psi)^A \text{ is } \forall u \left[\varphi'(u) \to \psi^A\right]$$
$$(\exists u\, \psi)^A \text{ is } \exists u \left[\varphi'(u) \wedge \psi^A\right]$$

The atomic step above is written for the case $R = \in$. For a general R it is that step which changes, because "$\in$" is a symbol of the language we are interpreting, and what it is interpreted as in (A, R) is R. So if $R = \{(x, y) \mid \varphi_R(x, y, q)\}$, then $(v_0 \in v_1)^{(A,R)}$ is $\varphi_R(v_0, v_1, q)$, and since the recursion bottoms out at the atomic subformulae, this replaces every occurrence of $\in$ in ψ. The parameter q simply comes along for the ride, which is exactly why φ_R needs the same renaming as φ did, in both directions: in particular q must not be a variable that ψ binds.

3.2 The Relative Consistency of ZF

Why are we doing all of this? Here's one reason.

Theorem 3.2.1 (working in $\mathbf{ZF} - \mathbf{F}$). *For each axiom $\sigma \in \mathbf{ZF}$ (**including** foundation), $\sigma^{\mathbf{WF}}$.*

The above reads, "The relativization of each **ZF** axiom to **WF** is true, working in $\mathbf{ZF} - \mathbf{F}$." Another equivalent formulation is that for all axioms σ of **ZF**,

$$\mathbf{ZF} - \mathbf{F} \vdash \sigma^{\mathbf{WF}} \tag{3.2.1}$$

Before drawing the consequence we are after, we isolate the fact about relativization that the argument will run on. We built ψ^A to be the interpretation of ψ in A, and an interpretation ought to carry proofs and not just statements: reading every quantifier of a proof as ranging over A turns each of its steps into a step about A. The one thing that can go wrong is A being empty, since first-order logic assumes a non-empty domain and proves $\forall u\, \psi \to \exists u\, \psi$; so non-emptiness is carried as a hypothesis rather than as a side condition, being a sentence of the object language like any other.

Lemma 3.2.2 (Scheme). *Let $A = \{x \mid \varphi(x)\}$, say with no parameters, and let σ and τ be sentences. Then, in the meta-theory, whenever $\sigma \vdash \tau$, we have σ^A, $\exists x\, \varphi(x) \vdash \tau^A$.*

Theorem 3.2.1 and Theorem 3.2.2 together yield the following corollary.

Corollary 3.2.3. *If $\mathbf{ZF} - \mathbf{F}$ is consistent, then so is **ZF**.*

Proof. We argue by contraposition. Suppose **ZF** is inconsistent. Say $\sigma_1, \ldots, \sigma_n$ are axioms of $\mathbf{ZF} - \mathbf{F}$ and

$$\{\sigma_1, \ldots, \sigma_n\} \cup \{F\} \vdash \varphi \wedge \neg\varphi$$

Apply Theorem 3.2.2 with σ the conjunction $\sigma_1 \wedge \cdots \wedge \sigma_n \wedge F$ and τ the contradiction. Relativization commutes with the propositional connectives, by the propositional step of the definition, so what comes out is

$$\left\{\sigma_1^{\mathbf{WF}}, \ldots, \sigma_n^{\mathbf{WF}}\right\} \cup \left\{F^{\mathbf{WF}}\right\} \cup \{\exists x\, (x \in \mathbf{WF})\} \vdash \varphi^{\mathbf{WF}} \wedge \neg\varphi^{\mathbf{WF}}$$

□

Together with the above theorem scheme, we get that $\mathbf{ZF} - \mathbf{F}$ is inconsistent—still under the supposition that $\mathbf{ZF}$ is. Indeed, Theorem 3.2.1 gives $\mathbf{ZF} - \mathbf{F} \vdash \sigma_i^{\mathbf{WF}}$ for each i, and also $\mathbf{ZF} - \mathbf{F} \vdash F^{\mathbf{WF}}$; and $\mathbf{ZF} - \mathbf{F} \vdash \exists x\,(x \in \mathbf{WF})$, since $\emptyset \in V_1$. So every hypothesis on the left of that display is a theorem of $\mathbf{ZF} - \mathbf{F}$, and hence $\mathbf{ZF} - \mathbf{F} \vdash \varphi^{\mathbf{WF}} \wedge \neg\varphi^{\mathbf{WF}}$.

3.3 Complexity of Formulae

Whether a formula says the same thing inside a class as it does in V depends on how its quantifiers range, so we need a way of measuring that. The first step is to give the quantifiers that range over a set, rather than over the whole universe, notation of their own.

Notation (Bounded Quantifiers).

- $\forall u \in v\,\varphi$ stands for $\forall u\,[u \in v \to \varphi]$
- $\exists u \in v\,\varphi$ stands for $\exists u\,[u \in v \wedge \varphi]$

Quantifiers written this way are said to be **bounded**.

A formula all of whose quantifiers are bounded can only ever look inside the sets it is handed, and such formulae sit at the bottom of the hierarchy we are about to build.

Definition 3.3.1 (Δ_0 Formulae). A formula is Δ_0 iff every quantifier occurring in it is bounded.

Everything above the bottom level is obtained by putting a block of unbounded quantifiers, all of the same kind, in front of a Δ_0 formula.

Definition 3.3.2 (Σ_1 and Π_1 Formulae).

- Σ_1 formulae look like

$$\exists u_0 \exists u_1 \cdots \exists u_n\,\varphi$$

  where φ is Δ_0.

- Π_1 formulae look like

$$\forall u_0 \forall u_1 \cdots \forall u_n \, \psi$$

where ψ is Δ_0.

There are no Δ_1 formulae. On the shapes above, a formula cannot begin with both an unbounded $\exists$ and an unbounded $\forall$; and allowing either block to be empty only puts the Δ_0 formulae into both classes, which buys us nothing we did not already have. What there are is Δ_1^T formulae, and to get them we first have to fix a theory to do the translating for us.

Definition 3.3.3 (Complexity Relative to a Theory)**.** Let T be a theory and let $\psi(\bar{u})$ be a formula. We say $\psi(\bar{u})$ is Δ_0^T iff there is a Δ_0 formula $\varphi(\bar{u})$ such that

$$T \vdash \forall \bar{u} \, [\psi(\bar{u}) \leftrightarrow \varphi(\bar{u})]$$

Similarly for Σ_1^T and Π_1^T. We say $\psi(\bar{u})$ is Δ_1^T iff it is both Σ_1^T and Π_1^T.

Keep going to get Σ_n and Π_n formulae, and also Σ_n^T, Π_n^T and Δ_n^T formulae.

Visit https://thefundamentaltheor3m.github.io/SetTheoryNotes/main.pdf for the latest version of these notes. If you have any suggestions or corrections, please feel free to fork and make a pull request to the associated GitHub repository.