APPENDIX A

First Steps: Axiomatic Set Theory

We shall work in the following first-order language:

$$\mathcal{L}_{\mathsf{set}} = \{ \underline{\in} \},$$

where $\underline{\in}$ is a binary relation symbol. To make the notation somewhat lighter, I will not keep underlining $\underline{\in}$ in expressions. If $\mathcal{U}$ is an $\mathcal{L}_{\mathsf{set}}$-structure ($\mathcal{U}$ for *universe*), then elements of $\mathcal{U}$ should be understood as sets. Thus, we will read $x \in y$ as "x is a *member of* y" or "x is *in* y" (where x is also a set...) – we dealt with this (a bit more informally) in the first chapter of the notes.

The following are useful shorthands:

- $x \notin y$ will stand for $\neg(x \in y)$.
- $(\forall x \in y)\phi$ will stand for $(\forall x)(x \in y \to \phi)$, where ϕ is some $\mathcal{L}_{\mathsf{set}}$-formula.
- $(\exists x \in y)\phi$ will stand for $(\exists x)(x \in y \wedge \phi)$, where, again ϕ is some $\mathcal{L}_{\mathsf{set}}$-formula.

1. Zermelo's Axioms

<u>Axioms (Batch 1)</u>

- **Empty set**: $(\exists x)(\forall y)(y \notin x)$.
- **Extensionality**: $(\forall x)(\forall y)((x \in y \leftrightarrow y \in x) \to x \doteq y)$,
- **Pairs**: $(\forall x)(\forall y)(\exists p)(\forall z)(z \in p \leftrightarrow (z \doteq x \vee z \doteq y))$,
- **Union**: $(\forall x)(\exists u)(\forall y)(y \in u \leftrightarrow (\exists z)(z \in x \wedge y \in z))$.

That's a good point to elaborate. The empty set axiom is self-explanatory. Extensionality, as we've said before, says that sets are determined by their elements. In particular, extensionality implies that there is a unique empty set, which we will (as usual) denote by $\emptyset$.

The axiom of pairs is also rather self-explanatory (they all kinda are, but if we let everyone explain themselves, what would the point be anyway), so let's just say what the symbols mean: For all sets x and y, there is a set p whose *only* elements are x

and y. Since this totally determines p we may well write $\{x, y\}$ for p. Thus, if $\{x, y\}$ and $\{x', y'\}$ are obtained using the axiom of pairs, we have that:

$$\{x, y\} = \{x', y'\} \text{ if and only if } x = x' \text{ and } y = y' \\ \text{or } x = y' \text{ and } y = x'.$$

Naturally, the axiom of pairs does not say that x and y should be distinct. If they so happen to be the same, then the set produced is $\{x\}$ (the *singleton* set containing x).

The axiom of unions says that given a set x there is a set u whose elements are precisely the elements of the elements of x. Our usual notation (in real-life mathematics) is:

$$u = \bigcup_{y \in x} y \quad \left(\text{or } \bigcup x \right).$$

Remark 1.0.1. Our axioms allow us to form the usual union of two sets. Indeed, if x and y are sets, then the usual set $x \cup y$ can be obtained by applying the axiom of unions to the pair $\{x, y\}$.

We can keep going. If x_1, x_2 and x_3 are sets, then there is a set $\{x_1, x_2, x_3\}$ (take the pair of the first two, the singleton of the third and then union them as we did in the remark!).

Axioms (Batch 2)

- **Powerset**: $(\forall x)(\exists p)(\forall y)(y \in p \leftrightarrow (\forall z)(z \in y \to z \in x))$.
- **Comprehension**: Now for a scheme... Let $\phi(x, y_1, \ldots, y_n)$ be an $\mathcal{L}_{\mathsf{set}}$-formula. Then:

  $$(\forall y)(\forall y_1) \cdots (\forall y_n)(\exists c)(\forall x)(x \in c \leftrightarrow (x \in y \wedge \phi(x, y_1, \ldots, y_n)).$$

Okay, powerset is self-explanatory (notation $\mathcal{P}(x)$), but we have finally come across a not-so-self-explanatory axiom (scheme). Comprehension... Let's try to (well) comprehend this. It says that for any "property of sets (possibly with parameters – forget this for now) $\phi(x)$, and any set y, there is a set c whose elements are precisely the elements of y which have property $\phi(x)$". Indeed, in the case of ϕ having a single variable x and y some set, we usually write $\{z \in y : \phi(z)$ for the set given to us by comprehension.

You may want to have the simpler scheme that forgoes the set y above, and (in the notation above) looks like:

$$(\forall y_1) \cdots (\forall y_n)(\exists c)(\forall x)(x \in c \leftrightarrow (x \in y \wedge \phi(x, y_1, \ldots, y_n)),$$

but this would be a terrible, terrible idea. Take, for example, as $\phi(x)$ the formula $x \notin x$. Then, the following sentence

$$(\exists c)(\forall x)(x \in c \leftrightarrow x \notin x)$$

would be an axiom. So, we'd have to have a set R such that:

$$x \in R \leftrightarrow x \notin x$$

since R is a set, we'd then have:

$$R \in R \leftrightarrow R \notin R,$$

and that would make Mister (Doctor) Russell rather sad. Let's write Z^- for the axioms we have gathered so far (Z for Zermelo and $^-$ because we don't yet have "infinity").

Exercise 1.0.2. Use Russell's paradox to show that if $\mathcal{U} \models \mathsf{Z}^-$, then $\mathcal{U} \models \neg(\exists x)(\forall y)(y \in x)$.

Sets vs. **Classes** (for people that don't care too much about the distinction, but care a little bit to know that there should be one): If T is some (consistent) set of $\mathcal{L}_{\mathsf{set}}$-sentences (i.e. an $\mathcal{L}_{\mathsf{set}}$-theory), then we understand the elements of some $\mathcal{U} \models T$ to be *sets*. Our axioms guarantee us the existence of certain elements of $\mathcal{U}$ (which represent, in terms of the $\in$ relation what we understand as sets). As we saw above, not every definable[1] subset of $\mathcal{U}$ is "represented" by an element, i.e. there are definable subsets of $\mathcal{U}$ which are not "sets". These collections of elements are what we call classes (e.g. if $\mathcal{U} \models \mathsf{Z}^-$, then there is a class of all sets, it's... you guessed it, U). I will try to be consistent in writing lower-case letters for sets and upper-case letters for classes, and I will not use the $\in$ symbol for classes – but I will say that some element (i.e. set) belongs to some class (which should be understood as meaning that it satisfies the first-order formula that defined said class).

Exercise 1.0.3. Show that Comprehension implies Emptyset.

Let $\phi(x, y)$ be the formula $x \in y$. Suppose that a and b are sets. Then:

$$\mathcal{U} \models (\exists c)(\forall x)(x \in c \leftrightarrow (x \in a \wedge x \in b)),$$

so we have (finite) intersections!

Exercise 1.0.4. Show that if x is a non-empty set, then, there is a set $\bigcap x$ which contains all elements that are contained in all elements of x. What happens if we drop the assertion that x is non-empty?

[1]Here and throughout this appendix, definable means $\mathcal{L}_{\mathsf{set}}$-definable (possibly with parameters).

It is also routine (i.e. go do it) to show that we can also build relative complements and symmetric differences from our axioms. All usual properties of unions and intersections follow from the fact that they hold for $\wedge$ and $\vee$ (whatever this sentence means, if you can interpret it correctly, then it is correct).

Our third axiom batch is what I'd like to call the "mysterious batch" Axioms (Batch 3)

- **Infinity**:
$$(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \to y \cup \{y\} \in x).$$
- **Foundation**:
$$(\forall x)(x \neq \emptyset \to (\exists y \in x)(x \cap y = \emptyset))$$

We have not defined what it means for a set to be infinite, but we have some intuitive understanding. It should be clear that the set whose existence is asserted by the axiom of infinity cannot be finite (in our intuitive understanding of the term).[2] The axiom of foundation tells us that the $\in$ relation is *well-founded* – i.e. we can't find infinite descending $\in$ chains *inside* the same set.

THEOREM 1.0.5. *The axiom of Foundation implies that no set is a member of itself.*

PROOF. Let x be a set. Then $\{x\}$ is a non-empty set, so by foundation, there is some $y \in \{x\}$ such that $y \cap \{x\} = \emptyset$. But, if $y \in \{x\}$, then $y = x$, so $x \cap \{x\} = \emptyset$, which means that $x \notin x$. □

2. Replacement and ZF

The next batch of axioms shouldn't really be called a batch, it's just a single axiom. To state it a bit more concisely, let me introduce some terminology:

Definition 2.0.1. Let $\phi(x, y; z_1, \ldots, z_n)$ be an $\mathcal{L}_{\mathsf{set}}$-formula in two variables from $\mathcal{U}$.[3] Let $Fu_\phi(z_1, \ldots, z_n)$ be the following formula:
$$(\forall x)(\forall y)(\forall y')((\phi(x, y; \bar{z}) \wedge \phi(x, y'; \bar{z})) \to y \doteq y')$$
say that ϕ is *functional in x (in $\mathcal{U}$)* for parameters $a_1, \ldots, a_n \in U$, if $\mathcal{U} \models Fu_\phi(a_1, \ldots, a_n)$.[4]

[2] Go back to Chapter 1 – We defined the *finite* ordinals. A set is finite if it is in bijection with a finite ordinal. There are more notions of finiteness which are sometimes equivalent.

[3] When discussing formulas without parameters, we can sometimes write them in a *partitioned* way, that is as $\phi(x_1, \ldots, x_n; y_1, \ldots, y_m)$, where, we understand the x_i's as the *object* variables and the y_j's as the *parameter* variables.

[4] Recall that this is shorthand for $Fu_\phi[\bar{a}/\bar{z}]$.

So, if $\phi(x, y)$ (either parameter-free or with the parameters supressed) is functional in $\mathcal{U}$, it defines a partial function $F_\phi : U \to U$ in the obvious way: If $b \in U$ and $\mathcal{U} \vDash (\forall y)\neg\phi(b, y)$, then $F_\phi(b)$ is undefined, and otherwise, there is a unique c such that $\mathcal{U} \vDash \phi(b, c)$, since $\mathcal{U} \vDash Fu_\phi$, and then we set $F_\phi(b) = c$.

Axioms (Batch 3)

- **Replacement**: For each formula $\phi(x, y; z_1, \ldots, z_n)$:

$$(\forall w)(\forall z_1)\cdots(\forall z_n)(Fu_\phi(z_1, \ldots, z_n) \to (\exists r)(\forall u)(u \in r \leftrightarrow (\exists x)(x \in w \wedge u \doteq F_\phi(x))))$$

So, this just says that if ϕ is a functional formula and w is any set, then the image of w under F_ϕ is also a set. This is usually denoted by:

$$\{u : (\exists x \in w)(u = F_\phi(x))\}.$$

Putting together all the axioms we've listed so far, we get the theory ZF.

3. The Axiom of Choice

First, some terminology. Let $(a_i)_{i\in I}$ be a family of sets... This means something precise now. In particular, this means that I *is* a set and the family in question is precisely a mapping (i.e. a set of ordered pairs (i.e. elements of the form $\{\{a\}, \{a, b\}\}$ for this class) which satisfies the functional property that if (x, y) and (x, y') are elements of the mapping then $y = y'$) whose domain (i.e. the set of all elements x for which there is some y such that (x, y) is in the mapping) is I. After you've digested all of these parenthetical definitions,[5] if a is a family of sets indexed by I (i.e. a mapping whose domain is I), we also denote this by $(a_i)_{i\in I}$, where a_i is just $a(i)$ (i.e. the second coordinate of the ordered pair in the mapping whose first coordinate is i). Anyway... let $(a_i)_{i\in I}$ be a family of sets. The product of this family is the set of mappings $f : I \to \bigcup_{i\in I} a_i$ (here this union is the union of the elements in the range of I) such that for all $i \in I$ we have that $f(i) \in a_i$. Then:

- **Choice**: Let $(a_i)_{i\in I}$ be a family of sets, such that for all $i \in I$, $a_i \neq \emptyset$. Then, the product of $(a_i)_{i\in I}$ is non-empty.

Now that you know first-order logic, you should convince yourselves that the axiom of choice is a first-order axiom:

Exercise 3.0.1. Write a sentence that expresses the axiom of choice.

[5]It would also be great to check that everything here is a set.

At this point, you should go back to the beginning of the course and read all about ordinals and cardinals.

We have now expressed ZFC as a first-order theory. If it has a model, then by Löwenhéim-Skolem it has a countable model. If you go back to the first chapter of this course (or just do Cantor's diagonal argument again or whatever) you can see that in any model of ZFC there is some uncountable set. The fact that there exists an uncountable set can be expressed in first-order logic ($\mathcal{U}$ thinks that there is an infinite set and this set has a powerset, that set will not be countable). What gives?

APPENDIX B

Łoś's Theorem

In this section, we assume the axiom of choice.

1. Filters and Ultrafilters

Let I be a set. A **filter** on I is a collection $p \subseteq \mathcal{P}(I)$ satisfying:

- $I \in p$, $\emptyset \notin p$.
- If $A, B \in p$ then $A \cap B \in p$.
- If $A \in p$ and $B \subseteq I$ is such that $A \subseteq B$, then $B \in p$.

Example 1.0.1. If you've seen some measure theory, then the collection of all subsets of $\mathbb{R}$ whose complement has Lebesgue measure 0 is a filter (indeed, letting μ denote Lebesgue measure, we have $\mu(\emptyset) = 0$, if $\mu(\mathbb{R} \setminus A), \mu(\mathbb{R} \setminus B) = 0$, then $\mu(\mathbb{R} \setminus (A \cap B)) = \mu((\mathbb{R} \setminus A) \cup (\mathbb{R} \setminus B)) = 0$ and if $A \subseteq B$ then $\mathbb{R} \setminus A \supseteq \mathbb{R} \setminus B$ so $\mu(\mathbb{R} \setminus A) = 0$ then $\mu(\mathbb{R} \setminus B) = 0$).

Example 1.0.2. Let I be a set and $x \in I$. Then $p_i = \{A \subseteq I : i \in A\}$ is a filter ($i \in I$, if $i \in A$ and $i \in B$ then $i \in A \cap B$, if $i \in A$ and $B \supseteq A$ then $i \in B$). We call this the *principal filter generated by i*.

Example 1.0.3. Let κ be an infinite cardinal and I a set with $|I| \geq \kappa$. Then $\{X \subseteq I : |I \setminus X| < \kappa\}$ is a filter (the argument is similar). If $\kappa = \aleph_0$ we call this the *Fréceht filter on I* – the filter of all *cofinite* subsets of I.

Exercise 1.0.4. Suppose that p is a filter on I and let $X \subseteq I$. If $X \notin D$ show that:

(1) $q = \{Y \subseteq I : (\exists Z \in p) Z \setminus X \subseteq Y\}$ is a filter.

(2) $q \supseteq p$.

(3) $I \setminus X \in q$.

A filter p on I is an *ultrafilter* if for all $X \subseteq I$ we have that either $X \in p$ or $I \setminus X \in p$. (What does this remind you of?) We sometimes write βI for the collection of all filters on I. This is called the *Stone–Čech compactification of* I.[1]

Example 1.0.5. Clearly, every principal filter is an ultrafilter (every subset of I either contains or does not contain a given element). In particular we can identify I with a subset of βI.

What about non-principal ultrafilters?

THEOREM 1.0.6. *Let p be a filter on I. Then, there is an ultrafilter $q \supseteq p$*

PROOF. By Zorn's lemma (see the first chapter) and the previous exercise (we'd love to use induction here, but we can't – this is the sticking point also for the proof of say completeness when the language is uncountable). To say a bit more, look at the set of all subsets of $\mathcal{P}(I)$ which are filters on I and contain p. Ordered by inclusion, this is a poset, and since p is in there, it's non-empty. Every chain in this poset has a maximal element (it's union, which is still a filter), so by Zorn's lemma this set has a maximal element. This element must be an ultrafilter, otherwise using the previous exercise we could extend it to a bigger filter. □

In particular, the Fréchet filter extends to a non-principal ultrafilter.

2. Ultraproducts

Let I be a set $p \in \beta I$ an ultrafilter on I and $(\mathcal{M}_i)_{i\in I}$ a family of $\mathcal{L}$-structures. We define an equivalence relation $\sim_p$ on $\prod_{i\in I} M_i$ by taking:

$$(a_i)_{i\in I} \sim_p (b_i)_{i\in I} :\iff \{i \in I : a_i = b_i\} \in p.$$

The ultraproduct of $(\mathcal{M}_i)_{i\in I}$ is the quotient $\widehat{\mathcal{M}} = \prod_{i\in I} M_i / \sim_p$ with the "obvious" structure. We usually denote the elements of $\widehat{M}$ by $[(a_i)_{i\in I}]_{\sim_p}$. The fundamental theorem here is the following:

THEOREM 2.0.1 (Łoś's Theorem). *Let $\phi(x_1, \dots, x_n)$ be an $\mathcal{L}$-formula. For any $[(a_i^1)_{i\in I}]_{\sim_p}, \dots [(a_i^n)_{i\in I}]_{\sim_p} \in \widehat{M}$, the following are equivalent:*

(1) $\widehat{\mathcal{M}} \vDash \phi\left([(a_i^1)_{i\in I}]_{\sim_p}, \dots, [(a_i^n)_{i\in I}]_{\sim_p}\right)$.

[1]For the word compactification to make sense I should give you a topology. But topology is not a prerequisite for this course. Anyway the (cl)opens are $\{p \in \beta I : A \subseteq p\}$, for each $A \subseteq X$. This space is compact – our goal is to prove compactness anyway.

(2) $\{i \in I : \mathcal{M}_i \models \phi(a_i^1, \ldots, a_i^n)\} \in p$.

PROOF. Induction on the complexity of ϕ! The base case is by the definition of the ultraproduct, while the $\neg$ case is by the fact that we are working with an ultrafilter. The $\wedge$ case is by the fact that ultrafilters are, well, filters.

For the $\exists$ case we need to worry a little bit, but the axiom of choice makes the argument work. □