Set Theory
some basics and a glimpse of some advanced techniques

Toby Meadows

READ ME:

These notes are still drafts:

- there will be typos;
- there may be errors; and
- I plan to augment them.

That said, they should be complete enough to be useful and I hope you find them so.

I plan to update this document on my website:

https://sites.google.com/site/tobymeadows/.

Unless you've already come from there, it could be worth looking there for a more recent version of this document.

Also, if you do spot any problems or there's something you don't like or understand, I'd like to hear about it. Please drop me an email at:

toby.meadows@gmail.com.

Contents

Preliminary remarks

My goal in these notes is to introduce both some of the basic tools in the foundations of mathematics and gesture toward some interesting philosophical problems that arise out of them. Of necessity, there is only so much we can do in such a short space - so I've made some important restrictions in scope.

First, I am going to concentrate on *set theoretic foundations*. I do this, partly, because it's my background and passion; however, set theory arguably makes the most plausible claim to providing a foundation of mathematics of any approach currently on the market. Moreover, as philosophers, it is quite likely that you have used set theoretic tools for model theoretic purposes in introductory logic courses - although they may have been hidden in the background. As such, I'm hoping there's something to build on.

The second restriction is with regard to how deep we dig. If we were to cover all of the material discussed in these notes in proper detail, it would probably take around 400 pages (and I usually underestimate). With this in mind I've adopted a strategy that I hope will work: we are going to start very slow and then just leap into the abyss. The plan is as follows:

- Session 1: Introduce the axioms of set theory and show how mathematical and semantic structures can be represented there.
- Session 2: Introduce ordinals, transfinite recursion, cardinals, Cantor's theorem and the continuum hypothesis.

These two sessions will be quite detailed and will probably feel a little dry. My goal here is to introduce enough of the language and machinery of set theory that: 1) you'll be familiar with basic set theoretic notation and where it comes from; and 2) you'll gain some familiarity with powerful set theoretic tools which can make light work of logical problems in the truth literature.

We then move on to:

- Session 3: Inner models, outer models, and large cardinals.
- Session 4: The multiverse and indefinite extensibility.

In these sessions, I won't be able to give all the details of the constructions used, but will hopefully be able to give you a *feel* for: what is going on; the overarching strategies involved; and the philosophical problems that emerge. I strongly believe that, while a technical background is helpful, there is a lot that philosophers with different specialties can offer to the foundations of mathematics - just like any other area of philosophy. While Session 3 covers work from the 50s through to the 90s, Session 4 will look a something very new and an old chestnut.

Part 1

Some basics of set theory

CHAPTER 1

Axioms & representations

This week we are going to cover:

- the axioms of ZFC;
- Russell's paradox;
- the language of set theory; and
- representing structures and models using sets.

Most of our time this week will be spent building up the language of set theory and in so doing gaining familiarity with its axioms. We shall work our way up from the sparse foundation to representing structures for a toy language and arithmetic.

1.1. ZFC

1.1.1. The set up. In one sense set theory is very simple.

1.1.1.1. *The language of set theory.* The language of set theory $\mathcal{L}_\in$ consists of just a single two place relation symbol $\in$.

1.1.1.2. *The axioms of ZFC.* Here they are:

1.	Extensionality	$\forall x\forall y\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y)$
2.	Foundation	$\forall x(\exists y\ y \in x\ \rightarrow \exists z(z \in x \wedge \forall w(w \in x \rightarrow w \notin z)))$
3.	Pairing	$\forall x\forall y\exists z\ (x \in z \wedge y \in z)$
4.	Union	$\forall x\exists z\forall y\forall w(y \in w \wedge w \in x \rightarrow y \in z)$
5.	Powerset	$\forall x\exists y\forall z(z \in y \leftrightarrow \forall w(w \in z \rightarrow w \in x))$
6.	Separation	$\forall x_0...\forall x_n\forall w\exists y\forall z(z \in y \leftrightarrow z \in w \wedge \varphi(z, x_0, .., x_n))$
7.	Replacement	$\forall x_0...\forall x_n(\forall y\exists! z(\varphi(y, z, x_0, ..., x_n) \rightarrow \forall u\exists w\forall y(y \in w \leftrightarrow \exists z(z \in u \wedge \varphi(z, y, x_0, ..., x_n)))$
8.	Infinity	$\exists x(\exists y\ y \in x \wedge \forall z(z \in x \rightarrow \{z\} \in x)$
9.	Choice	$\forall x(\forall y(y \in x \rightarrow \exists z(z \in y)) \wedge \forall y\forall z(y \in x \wedge z \in x \rightarrow \neg\exists w(w \in y \wedge w \in z)) \rightarrow \exists u\forall y(y \in u \rightarrow \exists! z(z \in y)))$

With just these tools, we can do anything that people do in the mathematics department. But this is also pretty intimidating if you haven't seen these before. While set theorists tend to think of them as very *natural* and *obvious* - good candidates for axioms - it can take a little while to see this. Moreover, I should add that the naturalness and truth of these axioms is sometimes a point of philosophical contention (see e.g., [Potter, 2004]), although these views are outliers in mathematics.

The plan for this week is to introduce each of the axioms through a kind of narrative which will hopefully provide some understanding of their role and function. My presentation owes a lot to the first chapter of Kunen [2006].

1.1.1.3. *Defined relations and terms.* The first thing we might note is that while it's impressive that set theory can work with just a single relation symbol, $\in$, this makes it very difficult for humans to read and understand. To alleviate this difficulty, we introduce a number of defined:

- relations; and
- terms

into the language.

Here is our first example of this:

DEFINITION 1. Let us say that x is a *subset* of y, abbreviated $x \subseteq y$, if

$$\forall z(z \in x \rightarrow z \in y).$$

This is a *relation* and we've defined it in the sense that it may be eliminated for its definiendum wherever it occurs. For this reason, we don't need to regard it as being a genuine part of the language of set theory, although it is obviously very useful to have around.

We'll define an example of a *term* in a little while. Since a term works like a name (or constant symbol), we need to know that the name actually picks something out. This is a little more technical.

1.2. The *story of the axioms*

1.2.1. Extensionality & Foundation. We start with two axioms which might be described as <u>*regulative*</u> in the sense that they don't tell us what sets are out there; but rather tell us what they must be like.

1. Extensionality $\quad \forall x \forall y \forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y)$

Informally, the axiom of extensionality tell us that if sets x and y have exactly the same members, the x and y are actually the same set.

So it says that sets are extensional; unlike, say, properties which are often treated as intensional. For example, taking Quine's classic (and likely problematic) example, the property of being *renate* (having a kidney) and the property of being *cordate* (having a heart) are distinct properties having distinct intensions; but the respective sets of objects satisfying each these properties are identical. This is usually regard as fundamental to the concept of set: non-extensional set theories are very rarely explored.

The *axiom of foundation* also plays a regulative role. As its name suggests, it says that the sets, in particular the membership relation $\in$, are well-founded.

2. Foundation $\forall x(\exists y\ y \in x \ \rightarrow \exists z(z \in x \wedge \forall w(w \in x \rightarrow w \notin z)))$

A little less formally, it says, that:

- if x is non-empty (i.e., has a member);
- then it has a $\in$-least member (i.e., there is some $z \in x$ such that every member of x is not a member of z).

EXAMPLE 2. Suppose x is a set consisting of the elements $y_0, y_1, y_2, ...$; and suppose that[1]

$$y_0 \ni y_1 \ni y_2 \ni ...$$

It is clear that x is non-empty. Then Foundation tell us that there is some z which is the $\in$-least member of x. Fix such a z.

By our definition of x, z must be y_n for some natural number n. But then $y_{n+1} \in z$, which contradicts the fact that z is supposed to be the $\in$-least member of x.

REMARK. This example shows us that Foundation forbids the existence of the kind of set x defined. Such a set is often known as an *infinite descending $\in$-chain.*

EXAMPLE 3. Suppose x is a set whose only member is x; $x \in x$ and for all y, $y \notin x$. Then x is clearly non-empty. Then using Foundation fix some z such that:

- $z \in x$; and
- z is the $\in$-least member of x.

[1]The $\ni$ is a shorthand giving us the inverse relation. Thus we have

$$x \in y \ \leftrightarrow \ y \ni x.$$

Since the only member of x is z, we can see that $z = x$. But this means that x is the $\in$-least member of x and so

$$\forall w(w \in x \to w \notin x).$$

But using x for w and the fact that $x \in x$, we see that $x \notin x$. This a contradiction.

REMARK. The upshot of this example is that Foundation tells us that there is no set x whose only member is x.

EXERCISE 4. Can a set x have more than one $\in$-least member? If so, give an example.

1.2.2. Getting sets, Russell's paradox and Separation. So far we cannot even say that a single set exists. To get things moving, we'll simply add an axiom, which we can remove later on, if we like.

0.	Set existence	$\exists x\ x = x$

This just says that there is a set. It's not very exciting and we're going to need more than just this to get things done.

Why not just say any set we can define exists? For example, we might consider the following axiom.

∞.	Naïve comprehension	$\exists y \forall x(x \in y \leftrightarrow \varphi(x))$ where $\varphi(x)$ is any formula of $\mathcal{L}_\in$ with at most x free.

The idea here is that to any way of describing the necessary and sufficient conditions for membership, we allocate a set to those conditions: the set of things satisfying those conditions. Strictly speaking, this is an *axiom schema* not an axiom. It stands for the infinitely many instances where a formula of $\mathcal{L}_\in$ is substituted in.

For well known reasons, this does not work.

THEOREM 5. *(Russell) Any theory including naïve comprehension is inconsistent.*

PROOF. Assume we are working in a theory that includes the axiom schema of naïve comprehension. We shall prove a contradiction.

Let $\varphi(x) := x \notin x$; so $\varphi(x)$ says that x is not a member of itself.

Using (∞.) fix an r such that for all x

$$x \in r \leftrightarrow r \notin r.$$

But then we have

$$r \in r \leftrightarrow r \notin r$$

which is impossible. □

Fortunately, ZFC provides us with a simple way around this. We still want to be able to use descriptions to pick out sets, but rather than letting any membership conditions pick out a set, we are going to use those conditions to *separate out* those elements satisfying the conditions from a set that we already established to exist.

6. Separation $\forall x_0 ... \forall x_n \forall w \exists y \forall z (z \in y \leftrightarrow z \in w \wedge \varphi(z, x_0, .., x_n))$

It's probably a little easier to take in, without the stream of $\forall x_0 ... \forall x_n$ out the front.[2]

6.' Separation $\forall w \exists y \forall z (z \in y \leftrightarrow z \in w \wedge \varphi(z))$

Informally, this says that given any set w, there is a set y consisting of exactly those elements z of w such that $\varphi(z)$. We *separate out the φ's* from w, to get y.

Unlike naïve comprehension, there is no known proof that systems including separation are inconsistent. For Gödelian reasons, we cannot do much better than this, other than to note that this system has been used extensively for over a hundred years.

We can also tell from this that there is no universal set.

THEOREM 6. *There is no x such that*

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

[2]In fact, with the rest of ZFC in the background, they are equivalent.

PROOF. Suppose for a contradiction that there such an x. Then by the axiom of separation, we get the following set

$$\{z \in x \mid z \notin z\}.$$

But for all w

$$w \in \{z \in x \mid z \notin z\} \leftrightarrow w \notin w$$

since x trivially contains all the sets. Thus, $\{z \in x \mid z \notin z\}$ is Russell's set whose existence is a contradiction. □

We are now in a position to define our first *term*.

THEOREM 7. *There is exactly one set x which has no members.*

PROOF. There are two parts to this claim:

(1) The existence of such a set; and
(2) The uniqueness of such a set.

(Existence) By (0.) fix some set x such that $x = x$. Let $\varphi := \forall y\ y \notin x$. Then by (6.) fix some z such that for all sets u

$$u \in z \leftrightarrow u \in x \wedge \forall y\ y \notin x.$$

Clearly z has no members.

(Uniqueness) Suppose w also has no members. Then for all u we have

$$u \in z \leftrightarrow u \in w$$

and so by extensionality $w = z$: thus, there is only one set with no members. □

Since, Theorem 7 tells us that there is exactly one empty set, we are warranted in introducing a name to denote this element.

DEFINITION 8. Let the *emptyset*, abbreviated $\emptyset$, be the set that is empty.

Theorem 7, tells us that the term $\emptyset$ is *well-defined*.

EXERCISE 9. Let x be an arbitrary set. Verify the following or provide a counter-example:

(1) $\emptyset \subseteq \emptyset$;
(2) $x \subseteq \emptyset$;

(3) $\emptyset \subseteq x$;
(4) $x \subseteq x$.

We can also introduce another useful *term* at this point. This term is more sophisticated than a mere name in that it has - what we might think of as - multiple inputs.

If we consider Axiom (6.) we can use extensionality to see that the set y supplied by the axiom must be unique.

PROPOSITION 10. $\forall x_0 ... \forall x_n \forall w \exists ! y \forall z (z \in y \leftrightarrow z \in w \wedge \varphi(z, x_0, .., x_n)$ for any $\varphi(z, x_0, ..., x_n)$ from $\mathcal{L}_\in$.

PROOF. Exercise. (Hint: *Suppose there were two such sets.*) □

Now Proposition 10, tells that given an inputs of:

(1) sets $x_0, ..., x_n, w$; and a
(2) formula $\varphi(z, x_0, ..., x_n, w)$ from $\mathcal{L}_\in$

there is a unique y containing the elements of w which satisfy $\varphi(z, x_0, ..., x_n, w)$.

We now describe a term which takes these inputs and denotes such a w.

DEFINITION 11. Let

$$\{z \in w \mid \varphi(z, x_0, ..., x_n)\}$$

be the set of those $z \in w$ such that $\varphi(z, x_0, ..., x_n)$.

This term is given to us directly by the axiom of separation. It is used very commonly in set theory as it is - it seems - easier to wield.

However, even with Axioms (0.) and (6.) in place, we still cannot guarantee the existence of kinds of sets we might find useful.

1.2.3. Pairing, Union and Powerset. These next axioms give us some sets which one would naturally think must exist. As opposed to Extensionality and Foundation, they are not so much regulative in that they tell us that new sets exists. More like Separation, there is a sense in which they take inputs and return new sets: they also have the $\forall\exists$ form. But unlike separation, which has a linguistic/schematic element, the kinds of new sets we get are quite simple to understand. I'll call these *combinatorial axioms* for this reason.

The first of these is the axiom of pairing.

3. Pairing $\forall x \forall y \exists z \, (x \in z \wedge y \in z)$

Informally, this says that given any two sets x and y there is a set which contains both x and y as members.

Moreover, using extensionality, we can show that there is exactly one set z containing exactly x and y.

PROPOSITION 12. *For all x and y, there is a unique w whose members are x and y.*

PROOF. (Existence) Use Pairing to get a set z containing x and y. Use Separation to get

$$\{u \in z \mid u \in x \vee u \in y\}.$$

(Uniqueness) Use Extensionality to show it is unique. □

This fact licenses us to augment our language with another useful term.

DEFINITION 13. For all x and y, let $\{x, y\}$ be the set containing just x and y.

EXERCISE 14. Is $\{x, y\} = \{y, x\}$? Prove it or provide a counterexample.

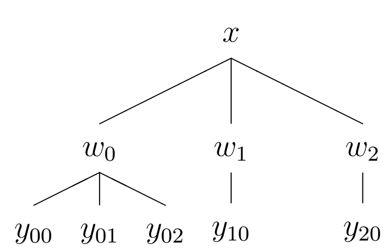
Our next combinatorial axiom is the Axiom of Unions.

4. Union $\forall x \exists z \forall y \forall w (y \in w \wedge w \in x \rightarrow y \in z)$

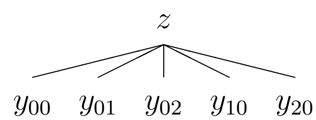
Informally, this tells us that for any set x there is a set z which contains all of the members of members of x.

That's still a bit of a mouthful. It's probably easier to get the idea with a pictorial example.

EXAMPLE 15. Given a set x that has elements w_0, w_1, w_2 each of which has elements described below



there is a set z such that



In a similar fashion to before, we can introduce a term which takes a set x and returns exactly the members of members of x.

PROPOSITION 16. *For all x there is a unique set y containing exactly the members of members of x.*

PROOF. Exercise. (*Similar to previous examples*) □

DEFINITION 17. Let $\bigcup x$ be the set of members of members of x.

EXERCISE 18. Show that for any finite collection of set $x_0, ..., x_n$, there is a set x containing just $x_0, ..., x_n$ as its members, thus showing that the term $\{x_0, ..., x_n\}$ is well-defined. (Hint: *use induction on the size of the collection. You'll need to use both Pairing and Union.*) From what point of view, so to speak, did establish this.

Our final combinatorial axiom is the Axiom of Powersets.

5. Powerset $\forall x \exists y \forall z (z \in y \leftrightarrow \forall w (w \in z \rightarrow w \in x))$

Informally, this axiom tells us that for every set x there is a set y which contains all the subsets of x.

EXAMPLE 19. Let $z = \{x_0, x_1\}$ be a set. Then there exists a set y which contains the following sets:

$$\emptyset, \{x_0\}, \{x_1\}, \{x_0, x_1\}.$$

As with the other combinatorial axioms, we may define a new and useful term.

PROPOSITION 20. *For any set x there is a unique y containing exactly the subsets of x.*

DEFINITION 21. Let the *powerset* of x, abbreviated $\mathcal{P}(x)$ be the set of subsets of x.

1.2.4. Some other useful notation. The following definition contains a number of other common notations used in set theory.

DEFINITION 22. (i) Let $x \cup y$ be the set of members of either x or y.

(ii) Let $x \cap y$ be the set of members of both x and y.

(iii) Let $x \backslash y$ be the set of members of x but not y.

EXERCISE 23. Definition 22 is informal. Show that the definition do provide well-defined sets. (Hint: *Use Separation.*)

1.3. Representing structures

There are only three axioms to go. But we already have enough tools to do quite a lot. Now our story will have a little bit of to-and-fro. We're going to work up to representing semantic models and mathematical structures - we'll build up our language a little more and we'll introduce the remaining axioms as we need them.

1.3.1. Ordered pairs, relations & functions. One of the first things we might note about sets is that they are different from lists. A list has an ordering which allows us to pick, say, the third element of the list - a set is not like this.

However, there is a way that we can construct sets which behave like lists, or as we shall call them sequences. We shall start by making an object which represents a list with just two members. This known as an *ordered pair*. (*Critically, this is exactly the object that mereological systems cannot form, leaving them incapable of providing a mathematical foundation - unless augmented.*)

Again, we shall introduce a term to represent <u>*ordered pairs*</u>. The method is a little different from the previous cases.

THEOREM 24. *There are formulae $\pi_0(v_0, v_1)$ and $\pi_1(v_0, v_1)$ of $\mathcal{L}_\in$ such that for any x_0 and x_1 there is a unique y such that:*

(1) $\pi_0(x_0, y) \wedge \forall z(\pi_0(z, y) \rightarrow z = x_0)$*; and*
(2) $\pi_1(x_1, y) \wedge \forall z(\pi_1(z, y) \rightarrow z = x_1)$.

REMARK. Intuitively, y is the set representing the ordered pair of x_0 and x_1; $\pi_0(x_0, y)$ say that x_0 is *the* first member of the ordered pair; and $\pi_1(x_1, y)$ says that x_1 is *the* second member of the ordered pair.

PROOF. Really, what we want to do here is find an object to do the job and then find formulae that will allow us to *decode* that object. There are many (infinitely many) ways of doing this, but by convention we the standard approach.

Given x_0 and x_1, consider the set

$$\{\{x_0\}, \{x_0, x_1\}\}.$$

Let $\pi_0(x_0, y)$ say that:

- y has no more than two members; and
- there is an element of y having (exactly) one member and this is x_0.

Let $\pi_1(x_1, y)$ say that either:

(1) y has two members, u_0 and u_1 where
  (a) u_0 has one member and u_1 has two members;
  (b) there is a unique x_0 such that $x_0 \in u_0$ and $x_0 \in u_1$; and
  (c) $x_1 \in u_1$ but $x_1 \neq x_0$.
(2) y has one member and $x_1 \in y$.

Given arbitrary x_0 and x_1 is should be clear that there is only one y such that $\pi_1(x_1, y)$ and that for such a y we must have $y = \{\{x_0\}, \{x_0, x_1\}\}$. This gives us the *uniqueness* of y.

Moreover, it should also be clear that given $y = \{\{x_0\}, \{x_0, x_1\}\}$ there can only be one z such that $\pi_0(z, y)$ and such a $z = x_0$. Similarly for π_1. This gives us the uniqueness of x_0 and x_1 as required. □

Now we introduce a term to our the ordered pair of x_0 and x_1.

DEFINITION 25. Let the *ordered pair* of x_0 and x_1 (in that order), abbreviated $\langle x_0, x_1\rangle$ be $\{\{x_0\}, \{x_0, x_1\}\}$.

REMARK. Note that we could have already defined this term before proving Theorem 24. We already have the tools for this in using our pairing notation. We need Theorem 24 rather to prove that $\langle x_0, x_1\rangle$ does what we want it to: i.e., that we can recover x_0 and x_1 from it in their correct order.

EXERCISE 26. Why does the second clause of the definition of $\pi_1(x_1, y)$ in Theorem 24 need two clauses: one for the case where y has two elements; and the other for when y just contains a single element? (Hint: *consider making the ordered pair* $\langle x, x\rangle$.)

EXERCISE 27. Define an ordered triple. Define an ordered n-tuple.

REMARK 28. Given $x_1, ..., x_n$ we write $\langle x_1, ..., x_n \rangle$ to denote the ordered n-tuple of $x_1, ..., x_n$.

We now use this tool to *represent relations*. The ability to do this is crucial to set theory's ability to provide a foundation for semantics and mathematics. We start with an example.

EXAMPLE 29. Let $A = \{a, b, c\}$ be a set. (*Notice that I'm using a capital Roman letter A to represent a set. This is a common convention in set theoretic practice. It's usually done when most of the other sets under consideration are elements of A. But it's also quite a loose convention. When doing set theory, one usually takes it that everything is a set, so it doesn't matter too much what kind of letters we use.*)

Now consider a binary relation R on the objects in A. Suppose that the following list of sentences describes how R relates the object in A:

- aRb, bRc, aRa and cRc.

Then we can just as easily represent this relation using a set of ordered pairs:

- $\{\langle a, b\rangle, \langle b, c\rangle, \langle a, a\rangle, \langle c, c\rangle\}$.

Let us call this set $R^\dagger$. Then it is clear that

$$xRy \ \leftrightarrow \ \langle x, y\rangle \in R^\dagger.$$

The lesson we take from this example is that relations can be represented by sets. Moreover, as set theorists we take this a step further.

DEFINITION 30. A *relation* is a set of ordered n-tuples.

So we are no longer saying that relations are merely *represented* by sets but that they *are* sets. This can be a vexing philosophical issue (see Potter [2004]), but we'll try not to get stuck here at the moment.

We now describe how to *represent functions*. Intuitively, a function is like a machine which takes some inputs and then gives an output.

- A function has a *domain* of inputs which it can take; and
- a *range* of outputs that it returns.

However, it's not too difficult to see that we can represent function using relations. Let's start with another example:

EXAMPLE 31. Suppose we have a function f (f *is everyone's favourite letter for functions*) and suppose it takes a single input of an element of $A = \{a, b, c\}$ and returns a value also from A. Then we can describe the function's behaviour with simple table:

f	
a	a
b	a
c	a

So f is a simple function which just returns a no matter what is given to it. Such functions are known as constant functions.

But we could also represent this situation with a list of sentences as follows:

- $f(a) = a$, $f(b) = a$ and $f(c) = a$.
  And just as we did with relations, we can represent this list using a set
- $\{\langle a, a\rangle, \langle b, a\rangle, \langle c, a\rangle\}$

which we might call $f^\dagger$. Then we clearly have

$$f(x) = y \;\leftrightarrow\; \langle x, y\rangle \in f.$$

Once again, the lesson to be taken here is that any function can be represented by a set. Moreover, such a set is a special kind of relation: for any inputs to the function, there can only be one output. Just as with relations, as set theories we say merely that functions can be represented by sets, but that they are sets.

DEFINITION 32. A n-ary function f (i.e., a function taking n inputs) is a $(n+1)$-ary relation (i.e., set of $(n + 1)$-tuples such that for any $x_1, ..., x_n$ there is at most one y such that

$$\langle x_1, ..., x_n, y\rangle \in f.$$

1.3.2. The axiom of replacement, products and finite sequences. So now with relations and function in place we are just about ready to represent models and structures. However, we're going to pause for a moment and discuss a subtle point with some deep philosophical ramifications.

Some of you will have noticed that my talk about terms in the earlier sections was a lot like talk about functions. As I described them, terms and functions took (possibly multiple) inputs and returned an output. For example, given x and y, the term $\{x, y\}$ was the pair of x and y. So should we understand

$$\{\cdot, \cdot\}$$

as a function taking sets and returning another set.

The problem is that Definition 32 does not allow this since it demands that a function be a set.

LEMMA 33. $\{\cdot, \cdot\}$ *is not a function.*

PROOF. This is trivial once we have the next axiom (the Axiom of Replacement) in place, but it's a little fussy here. (*Essentially, we making another version of the Russell set.*)

To make things a little simpler we'll consider the function $\{\cdot\}$ which takes a set and returns what is known as its *singleton*.

Suppose for a contradiction that $\{\cdot\}$ is a function. Then as a function, it is the set

$$\{\langle x, \{x\}\rangle \mid x = x\}.$$

Call this set P. Then by Separation there must be a set

$$R = \{z \in P \mid \exists x \ \langle x, \{x\}\rangle = z \ \wedge \ z \notin x\}.$$

Consider the set $Q = \langle R, \{R\}\rangle$. Clearly $Q \in P$.

Suppose $Q \in R$. Then by the definition of R, $Q \notin R$.

But then since $Q \notin R$ and $Q \in P$, we have $Q \in R$: contradiction. □

We might say that a term like $\{\cdot, \cdot\}$ is *too big* to be a function. In a similar sense, the universal set was *too big* to be a set.

Nonetheless, it's also handy to treat the universal set somewhat like a set and treat $\{\cdot, \cdot\}$ somewhat like a function. To this end, we introduce what we shall call *classes*. Classes are like sets and indeed can have the same extension as sets, but they can also be larger. We shall allow that any set of conditions determines a class (although we know that it may not determine a set.

Thus if we write $\{x \mid \varphi(x)\}$ we know there is a class corresponding to that, although there may not be a set. For example,

$$\{x \mid x \notin x\}$$

is a class but not a set. Such classes are known as *proper classes*. The existence of these entities are philosophically vexed. Are they out there? Maybe they're just pluralities Boolos [1984]? We shall try to adopt an agnostic attitude to them by noting that we could always remove talk of proper classes and just make do with the formula that defines them.

Thus, whenever we write

$$y \in \{x \mid \varphi(x)\}$$

we could also just write $\varphi(y)$.

However, we shall also find it convenient to have a notation for referring to these *entities*. Thus, we shall refer to proper classes by bold roman capitals, e.g., $\mathbf{A}, \mathbf{B}, \mathbf{F} \dots$.

Note that we can also represent terms using this notational convention. Thus there is a proper class function $\mathbf{F}$ such that

$$\mathbf{F} = \{z \mid \exists x \exists y \ \langle x, y, \{x, y\} \rangle\}$$

exactly corresponding to the term $\{\cdot, \cdot\}$. To apply this function, we shall write

$$\mathbf{F}(x, y) = \{x, y\}.$$

This now brings us to the *Axiom of Replacement*.

7. Replacement $\forall x_0 ... \forall x_n (\forall y \exists ! z (\varphi(y, z, x_0, ..., x_n) \rightarrow \forall u \exists w \forall z (z \in w \leftrightarrow \exists y (y \in u \wedge \varphi(y, z, x_0, ..., x_n)))$

PROPOSITION 34. *Given sets $x_0, ..., x_n$, a (class) function $\mathbf{F}(v_0, ..., v_n)$ and a set u there is a unique set w which is the set of values of $\mathbf{F}$ for inputs $x_1, ..., x_n$ and y from u.*

DEFINITION 35. Given sets $x_0, ..., x_n$ a function $\mathbf{F}(v_0, ..., v_n)$ and a set u let

$$\{\mathbf{F}(y, x_1, ..., x_n) \mid y \in u\}.$$

THEOREM 36. *For all sets A and B, there is a unique set C consisting of of all the ordered pairs $\langle a, b \rangle$ where $a \in A$ and $b \in B$.*

PROOF. We make two appeals to the Axiom of Replacement. First let $\mathbf{F}(v_0, v_1)$ be the function such that for x, y

$$\mathbf{F}(x, y) = \langle x, y \rangle.$$

Then by Replacement, we see that for any a, there is a set

$$\{\mathbf{F}(a, b) \mid b \in B\} = \{\langle a, b \rangle \mid b \in B\}.$$

First let $\mathbf{G}(v_0)$ be a function such that for $a \in A$

$$\mathbf{G}(a) = \{\langle a, b \rangle \mid b \in B\}.$$

Then by the Replacement again, there is a set

$$\{\mathbf{G}(a) \mid a \in A\}$$

and by union there is a set

$$\begin{aligned} \bigcup\{\mathbf{G}(a) \mid a \in A\} &= \bigcup\{\{\langle a, b \rangle \mid b \in B\} \mid a \in A\} \\ &= \{\langle a, b \rangle \mid a \in A \wedge b \in B\} \end{aligned}$$

which is what we wanted. □

DEFINITION 37. Given sets A and B, the *Cartesian product of A and B*, denoted $A \times B$ is the set of ordered pairs $\langle a, b \rangle$ with $a \in A$ and $b \in B$.

REMARK. Regarding why this is called a product, consider the sets $A = \{a_1, a_2\}$ and $B = \{b_1, b_2, b_3\}$. Verify that the size of $A \times B$ is 6; i.e., the product of the sizes of A and B.

1.3.2.1. *Notational conventions, proving relations exist, defining functions.* Whenever we define a relation R we remark on what R is a relation *on*. Thus, we write

$$R \subseteq A \times B.$$

For a function f, we remark that it is a function from its domain to some other set, its codomain. Thus, we write

$$f : A \to B.$$

Note that the codomain of a function is not necessarily its range. It can be any superset of the range. Strictly this information is not contained in the function itself.

However, it is extremely useful to follow these conventions. This is known as *typing*. They can allow us to follow the gist of complicated definitions without following all the detail.

For example, let $Sent_{\mathcal{L}}$ be the sentences (coded as sets) of some language $\mathcal{L}$ and let $2 = \{0, 1\}$. Then if we write

$$val : Sent_{\mathcal{L}} \to 2$$

we see that this function takes sentences and returns either 0 or 1 as output. A semantic valuation functions is an example of a function that does this. Even without seeing the full detail of a definition, we can learn something interesting about a function. Moreover, this makes checking that complicated working actually makes sense much easier.

It is also often useful to be able to take a function and *restrict* it. For example, suppose we have a function f with domain A and a set $C \subseteq B$. Then it's pretty obvious that there is a another function $g : C \to B$, which does exactly the same thing for input from A and is not defined for elements in A but not in C. More formally we introduce the following definition.

DEFINITION 38. Given a function f with domain A and a set $C \subseteq A$, let

$$f \upharpoonright C = \{\langle c, f(c)\rangle \mid c \in C\}.$$

EXERCISE 39. What sort of "thing" is $\upharpoonright$? Why is it well-defined.

REMARK 40. It is often useful to use these typing conventions with regard to class functions and relations. Thus we shall often write $\mathbf{F} : \mathbf{V} \to \mathbf{V}$ and $\mathbf{R} \subseteq \omega \times \omega$.

DEFINITION. Let $\mathbf{F} : \mathbf{V} \to \mathbf{V}$ be a class function and let A be a set. Show that $\mathbf{F} \upharpoonright A$ is an ordinary (set) function.

The following definitions describe some other common tools.

DEFINITION 41. Let A and B be sets, then

$$B^A = \{f \in \mathcal{P}(A \times B) \mid f \text{ is a function.}\}$$

Thus, informally, B^A is the set of functions from A to B. This means that another way of writing $f : A \to B$ is to say that $f \in B^A$.

DEFINITION 42. Let f be a function with domain A and let $B \subseteq A$. Then the *pointwise image of* B *via* f, abbreviated $f``B$, is such that

$$f``B = \{f(b) \mid b \in B\}.$$

EXERCISE 43. How is Definition 42 justified?

1.3.3. Models & Structures. A *language* $\mathcal{L} = \{P, R, ...f, g, ...c, d\}$ consists of relation symbols (or predicates), function symbols and constant symbols (or names).

Rather than being fully general, let us consider a simple example of a language which contains one example of each type of vocabulary. Let $\mathcal{L} = \{R, f, c\}$ where R is a 2-place relation symbol, f is a 2-place function symbol and c is a constant symbol.

A *model* $\mathcal{M}$ has a domain M which we want the language to talk about. $\mathcal{M}$ provides an interpretation for the the vocabulary of $\mathcal{L}$. We denote the interpretation of:

- R by $R^{\mathcal{M}}$;
- f by $f^{\mathcal{M}}$; and
- c by $c^{\mathcal{M}}$.

The *interpretation* is constrained such that:

- $R^{\mathcal{M}} \subseteq M \times M$;
- $f^{\mathcal{M}} : M \times M \to M$; and
- $c^{\mathcal{M}} \in M$.

Any model which provides interpretations satisfying these constraints is a model of $\mathcal{L}$. If we wanted to go further down this path, we would then define a satisfaction relation $\models$, that could tell us when sentences φ of $\mathcal{L}$ are true in some model $\mathcal{M}$.

REMARK 44. Instead of writing $R^{\mathcal{M}}$ we could write $\mathcal{I}(R)$ to highlight the fact that interpretation is a kind of function. This is often done in modal logic books, but - I think - is more cumbersome and is not the standard practice in mathematical logic.

1.3.4. Infinity. We now turn to one of the more infamous axioms of set theory.

8. Infinity $\exists x(\exists y\ y \in x \wedge \forall z(z \in x \rightarrow \{z\} \in x)$

Intuitively, it might be easier to run through a sketch of Dedekind's "proof" of the Axiom of Infinity in order to get a better feel for it.

EXAMPLE 45. Let's forget about sets for a moment and give a different interpretation to the membership relation $\in$. Let us say that

$$x \in y \;\leftrightarrow\; y \text{ is a thought about } x.$$

Then under this interpretation, $\{x\}$ is the thought that is just about x.

Now consider the Axiom of Infinity under this interpretation. It says that:

- there is a thought x such that:
  - x is a thought about something y; and
  - for every thought z that x is about, x is also about $\{z\}$ (i.e., the thought that us just about x).

Informally, we might defend the existence of such a thought as follows.

(1) Take any thought, z.
(2) Then I can think about just that thought; thus, $\{z\}$ exists.
(3) Moreover, I can think about thinking about just that thought; so $\{\{z\}\}$ also exists.
(4) Similarly, for any number n, I can think about, thinking about, thinking about, ... thinking about z.
(5) But I can also think about all the thoughts referred to in (4.). Call such a thought, x.
(6) Then x is an example of a thought satisfying the Axiom of Infinity.

EXERCISE 46. Prove the existence of $\emptyset$ using the Axiom of Infinity.

1.3.5. Modeling arithmetic. Now we'll provide a model for the theory of arithmetic. We'll keep things simply and assume our language, $\mathcal{L}_{Ar}$, consists of:

- a constant symbol, 0;
- a relation symbol, $<$; and
- a function symbol, s (the successor function).

The rest of the cast can easily be defined from this. Our goal is to provide a domain and interpretation for this vocabulary.

But first, I want to note that I've chosen a slightly unusual form of infinity above. It is close in spirit to Dedekind's *proof* of the axiom of infinity. It gives us what is known as a Zermelo ordinal rather than a von Neumann ordinal. In ZFC we can prove one axiom from the other (although this is not always possible in weaker set theories).

We now provide a more common formulation of the Axiom of Infinity, for which the following class function is useful.

DEFINITION 47. Let $\mathrm{s} : V \to V$ be such that $\mathrm{s}(x) = x \cup \{x\}$ for all $x \in \omega$.

Our alternative Axiom of Infinity is as follows:

8'. Infinity $\exists x(\emptyset \in x \wedge \forall z(z \in x \to \mathrm{s}(z) \in x)$

FACT 48. $(8) \leftrightarrow (8')$.

REMARK. We need to develop some more powerful tools (induction), which we'll introduce next time, before we can easily prove this.

In order to provide a model of arithmetic, we first need a domain. For this, we use the smallest infinite object given to us by Axiom $8'$.

PROPOSITION 49. *There is a unique set y such that for every object $z \in y$ either:*

- $z = \emptyset$*; or*
- $\exists u\ z = \mathrm{s}(u)$.

REMARK. Any such z will be finite.

PROOF. Let x be given by Infinity, then use separation to get the set

$$y = \{z \in x \mid z = \emptyset \ \vee \ \exists u(z = \mathrm{s}(u)\}.$$

Clearly, y has the same extension as the class

$$\{z \mid z = \emptyset \ \vee \ \exists u(z = \mathrm{s}(u)\}$$

so uniqueness follows. □

This allows us to define a term ω.

DEFINITION 50. Let ω be the set y given by Proposition 49.

We now define our standard model $\mathbb{N}$ for $\mathcal{L}_{Ar}$ as follows:

DEFINITION 51. Let $\mathbb{N} = \langle \omega, \emptyset \in \cap(\omega \times \omega), \mathrm{s} \restriction \omega \rangle$.

Thus, the domain of $\mathbb{N}$ is ω and:

- $0^{\mathbb{N}} = \emptyset$;
- $<^{\mathbb{N}} = \in \cap(\omega \times \omega)$; and
- $s^{\mathbb{N}} = \mathrm{s} \restriction \omega$.

REMARK 52. Note that we have restricted $\in$ is a relation with a larger field that ω and s is a function with larger domain and range than ω. Thus we, strictly, need to restrict these to the appropriate domain in order to define the model. However, it is common practice to not bother as the restriction required is always obvious. Thus we more commonly write $\mathbb{N} = \langle \omega, \emptyset, \in, \mathrm{s} \rangle$.

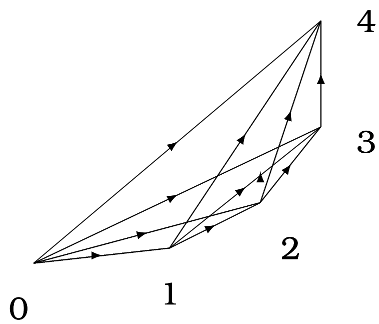
We think of the sets in ω as natural numbers.

EXAMPLE 53. Here are the first few natural numbers:

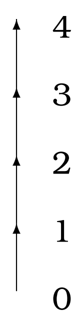
0	$\emptyset$
1	$\{\emptyset\}$
2	$\{\emptyset, \{\emptyset\}\}$
3	$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
4	$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$

This may seem a little strange or artificial, but the following diagram may make it a little easier to see why this is a very natural way to represent the natural numbers.

EXAMPLE 54. The following diagram is an illustration of the natural number 4 as canonically represented by a set. An arrow from one node of the diagram is intended to mean that the first node is a member of the second node.



Notice there is a sense in which some of the arrows here are redundant. Since the membership relation is transitive, in the context of representing finite ordinals, we can omit the *composed* arrows and represent the situation as follows:



This is known as a *Hasse diagram*. The idea is that the fact that we don't need to actually draw an arrow between 0 and 2; this is already conveyed by the fact that there is an arrow between 0 and 1 and another arrow between 1 and 2.

1.3.5.1. *Finite sequences.* Before we finish here, I want to mention a convention regarding n-tuples. Above we described them as a *generalisation* of an ordered pair. However, usually we do something a little different.

If we want to make a *list* of n sets, we use a function $f : n \to x$. This is a set of ordered pairs indexed by a natural number. We call such functions *finite sequences.* Moreover, we usually write $\langle x_0, ..., x_{n-1}\rangle$ to denote such a sequence. (Thus, we have used the same notation to denote a finite sequences and n-tuples. This almost never causes any problems unless you are interested in weak set theories Devlin [1984].)

EXAMPLE 55. Given a list x_0, x_1, x_2 we may represent this list as a finite sequence $\langle x_0, x_1, x_2 \rangle$ which is a function $f : 3 \to \{x_0, x_1, x_2\}$ such that:

$$\begin{aligned} f(0) &= x_0 \\ f(1) &= x_1 \\ f(2) &= x_2. \end{aligned}$$

Or even more fundamentally,

$$\begin{aligned} \langle x_0, x_1, x_2 \rangle = \{ &\{\{0\}, \{0, x_0\}\}, \\ &\{\{1\}, \{1, x_1\}\}, \\ &\{\{2\}, \{2, x_2\}\} \} \end{aligned}$$

1.4. Some remarks

1.4.1. Reduction? I claimed at the beginning that using the nine axioms of ZFC, we could do everything that people do in the mathematics department. This sounds like a *reductive claim*. I don't propose to defend or reject this, but we might note a few points that have emerged through this chapter.

1.4.1.1. *Ontological reduction?* While the axiom of pairs appears to give us something which just *is* the pair of two sets, the ordered pair construction seemed more contrived. Instead of getting some primitive list of two objects, we used sets to construct an object from which we could always *decode* the first element of the pair and the second element of the pair. As such, we'd be uncomfortable saying that this object *is* an ordered pair; perhaps it merely *represents* that thing.

This might make us think that there was no ontological reduction going on. However, we might make a further move. Perhaps the idea of an ordered pair is merely derivative on our concept of set. Perhaps we've gained some kind of insight here. Perhaps we've learnt that sets are more fundamental than lists.

But there's another problem. There isn't just one way of coding up a object to represent an ordered object. There are infinitely many ways to do this. So it seems a little odd to say that we've told you that list are just sets that are like *blah*, when they could also be like *blerg* and *blorg*. This is a variation of the basic problem for structuralism.

1.4.2. No ordinary objects. Note that we are not making sets of objects in the real world. For most purposes we can make do without them.

CHAPTER 2

Backbones & problems

This week we are going to look at:

- ordinals;
- transfinite induction and recursion;
- cardinals;
- Cantor's theorem; and
- the continuum hypothesis (CH).

2.1. Ordinals

2.1.1. Well orderings. We start with some basic definitions and examples of ordering and work out way to the notion of well-ordering. This will all be quite abstract for a while, but there are some interesting results at the end of this seminar.

2.1.1.1. *Orderings.*

DEFINITION 56. Let $\langle A, < \rangle$ be a structure.

(1) $<$ is *reflexive* if $a < a$ for all $a \in A$;
(2) $<$ is *symmetric* if for all $a, b \in A$, $a < b \leftrightarrow b < a$;
(3) $<$ is *transitive* if for all $a, b, c \in A$, $a < b \wedge b < c \rightarrow a < c$;
(4) $<$ satisfies *trichotomy* if for all $a, b \in A$, $a < b \vee a = b \vee b < a$.

EXERCISE. Let $A = \{a_1, a_2, a_3\}$. Describe examples of each of the relations in Definition 56 on A.

DEFINITION 57. Let $\langle A, < \rangle$ be a structure.

(1) $<$ is *irreflexive* if $a \not< a$ for all $a \in A$;
(2) $<$ is *asymmetric* if for all $a, b \in A$, $a < b \leftrightarrow b \not< a$;
(3) $<$ is *anti-symmetric* if for all $a, b \in A$, $a < b \wedge b < a \leftrightarrow a = b$.

EXERCISE. Let $A = \{a_1, a_2, a_3\}$. Describe examples of each of the relations in Definition 57.

DEFINITION 58. Let $\langle A, < \rangle$ be a structure.

(1) $\langle A, < \rangle$ is a (strict) partial ordering if $<$ is transitive, irreflexive and asymmetric.
(2) $\langle A, < \rangle$ is a total ordering if $<$ is a partial ordering that satisfies trichotomy.

EXAMPLE 59. The natural numbers are a total ordering. So are the rational numbers and the real numbers.

DEFINITION 60. Let $\langle A, < \rangle$ be a structure. $<$ is well-founded if for all $X \subseteq A$ with $X \neq \emptyset$, there is some $x \in X$ such that for all $y \in X$, $y \not< x$ (we call x a $<$*-minimal element* of X).

EXERCISE 61. Which of the following are well-orderings.

(1) the natural numbers;

(2) the integers;
(3) the rational numbers;
(4) the real numbers.

DEFINITION 62. Let $\langle A, < \rangle$ be a structure. $\langle A, < \rangle$ is a well-ordering if it is a well-founded linear ordering.

EXAMPLE 63. We'll work with the natural numbers, ω, but rather than using their natural ordering, we define some alternative well-orderings, which we denote by $\prec$.

(1) $1, 2, 3, ...0$;
(2) $0, 2, 4, ..., 1, 3, 5, ...$.

The idea with (1) is that we put 0 at the end of the ordering. So for every n such that $n \neq 0$, we have $n \prec 0$ and we keep the ordinary ordering for the rest.

To verify that (2) gives us a well-ordering, let $X \subseteq \omega$. Then either X contains an even number or it doesn't. Suppose it does. Then there is some $<$-least even number that it contains, say m. Then is clear that m is the $\prec$-least even number; and m is $\prec$ all of the odd numbers by definition. Suppose X doesn't contain any even numbers; i.e., it only contains odd numbers. Then there is a $<$-least odd odd number, n, in X which is also the $\prec$-least odd number. And since the odd numbers are all $\succ$ the even numbers, n is the $\prec$-least of any of the numbers in X. Thus in either, case X has a $\prec$-minimal element. Hence, $\prec$ is well-founded.

EXERCISE 64. Why is the ordering $\prec$ of (2) in Example 63 a total ordering?

2.1.2. Ordinals. Generalising on our method of representing natural numbers, we also represent well-orderings using sets.

DEFINITION 65. x is *transitive* if $\forall y(y \in x \to y \subseteq x)$.

PROPOSITION 66. *x is transitive iff* $\forall y \forall z(z \in y \wedge y \in x \to z \in x)$.

Thus a transitive set x is one such that any member of a member of x is a member of x. Transitive sets are extremely important in the model theory of set theory.

EXERCISE 67. Show that x being transitive is not the same as $\in$ being transitive on $\langle x, \in \rangle$.

DEFINITION 68. x is an *ordinal* if:

- x is transitive; and
- x is well-ordered by $\in$.

We denote the class of ordinals by $\mathbf{On}$.

REMARK 69. By convention, we use lower case Greek letters $\alpha, \beta, \gamma...$ to represent ordinals. So rather than writing $x \in \mathbf{On} \wedge \varphi(x)$, we just write $\varphi(\alpha)$. If we wanted to be particularly formal about this, we are introducing a new *sort* of variable into our language.

PROPOSITION 70. *For all α, $\langle \alpha, \in \cap (\alpha \times \alpha) \rangle$ is a well-ordering.*

To simplify the notation, we shall mostly write $\langle \alpha, \in \rangle$ to mean $\langle \alpha, \in \cap (\alpha \times \alpha) \rangle$.

LEMMA 71. *(i) $\emptyset$ is an ordinal.*
(ii) If α is an ordinal and $x \in \alpha$, then x is an ordinal.
(iii) If $\alpha \neq \beta$ are ordinals and $\alpha \subseteq \beta$, then $\alpha \in \beta$.
(iv) If α, β are ordinals, then either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.

PROOF. (i) $\emptyset$ is trivially transitive, trivially a total order and trivially well-founded since $\emptyset$ is empty.
(ii) Let $x \in \alpha$. We show that x is transitive. Let $y \in x$. Since $y \in x \in \alpha$ and α is transitive, we have $y \in \alpha$ and $y \subseteq \alpha$.
Since $\in$ is a well-ordering, it is transitive on $\langle \alpha, \in \rangle$ and so we have

$$\forall x, y, z \in \alpha (z \in y \wedge y \in x \to z \in x).$$

Putting this together, we have $\forall z (z \in y \to z \in x)$; i.e., $y \subseteq x$ and so x is transitive.
We show that $\in$ is a well-ordering on $\langle x, \in \rangle$. First we show that $\in$ is transitive on $\langle x, \in \rangle$. Let $w, u, v \in x$. Then by transitivity of α we see that $w, u, v \in \alpha$ and so by the transitivity of $\in$ on $\langle \alpha, \in \rangle$ we have

$$w \in u \wedge u \in v \to w \in v$$

which is what we want. Similar for trichotomy.
Finally we show well-foundedness. Suppose $A \subseteq x$ with $A \neq \emptyset$. Now suppose $a \in A$; then $a \in x \in \alpha$ and so $a \in A$ by transitivity of α. This means that $A \subseteq \alpha$ and since $\in$-well-orders $\langle \alpha, \in \rangle$, there exists $y \in A$ such that y is $\in$-least in

$A \cap \alpha$. Then since $A \subseteq x \subseteq \alpha$, we have $A \cap \alpha = A \cap x$, so y is $\in$-least in $A \cap x$; i.e., y is $\in$-least in A for the structure $\langle x, \in \rangle$.

(iii) Let γ be the $\in$-least element of $\beta \backslash \alpha$. We claim that $\gamma = \alpha$, which suffices since $\gamma \in \beta$. First we note that since β is transitive

$$\{\xi \in \beta \mid \xi \in \gamma\} = \gamma.$$

Suppose $\xi \in \alpha$. Then since $\alpha \subseteq \beta$, $\xi \in \beta$. Then by trichotomy, we have either

$$\xi \in \gamma \ \vee \ \xi = \gamma \ \vee \ \gamma \in \xi.$$

If $\gamma = \xi$, then $\gamma \in \alpha$, contrary to its definition.

If $\gamma \in \xi$, then $\gamma \in \xi \in \alpha$, so $\gamma \in \alpha$: contradiction.

Thus $\xi \in \gamma$ and so $\alpha \subseteq \gamma$.

Suppose $\xi \in \beta$ and $\xi \in \gamma$ and for a contradiction, suppose, $\xi \notin \alpha$. But then $\xi \in \beta \backslash \alpha$ and $\xi \in \gamma$ contradicting the minimality of γ.

(iv) Let α and β be ordinals. It can be seen that $\alpha \cap \beta$ is an ordinal. We claim that $\alpha \cap \beta = \alpha$ or $\alpha \cap \beta = \beta$. Suppose not. Then by (iii) $\alpha \cap \beta \in \alpha$ and $\alpha \cap \beta \in \beta$, which means that $\alpha \cap \beta \in \alpha \cap \beta$, which contradicts the definition of ordinal since it is a (strict) total ordering. □

EXERCISE 72. Show that if α and β are ordinals, then $\alpha \cap \beta$ is an ordinal.

It's probably easier to get a feel for ordinals by looking at some examples of ordinals.

EXAMPLE 73. $\{\} = \emptyset$.

$1 = \{\emptyset\}$.

$2 = \{\emptyset, 1\}$

$\omega = \{\emptyset, 1, 2, ...\}$

$\omega + 1 = \{0, 1, 2, ...\omega\}$

$\omega + 2 = \{0, 1, 2, ..., \omega, \omega + 1\}$

$\omega + \omega = \{0, 1, 2, ...\omega, \omega + 1, \omega + 2, ...\}$

2.1.2.1. *Successor and limit ordinals.*

THEOREM 74. *(i) Let α be an ordinal, then $\mathrm{s}(\alpha)$ is the least ordinal greater than α.*

PROOF. Clearly $\alpha \in \alpha \cup \{\alpha\} = \mathrm{s}(\alpha)$, so $\alpha < \mathrm{s}(\alpha)$. It will suffice to show that it is the least such ordinal. Suppose not. Then fix β with $\alpha < \beta < \mathrm{s}(\alpha)$. Then since $\beta \in \alpha \cup \{\alpha\}$, either $\beta \in \alpha$ or $\beta = \alpha$. It either case we cannot then have $\alpha \in \beta$, which contradicts our assumption. □

DEFINITION 75. (i) A *successor ordinal* is an ordinal α for which there is an ordinal β such that $\alpha = \mathrm{s}(\beta)$.
(ii) A *limit ordinal* α is such that for all $\beta < \alpha$, $\mathrm{s}(\beta) < \alpha$.

EXERCISE 76. What sort of ordinal is $\emptyset$?

EXERCISE 77. Suppose α is a limit ordinal. Show that $\bigcup \alpha = \alpha$. If α is a successor ordinal, what is $\bigcup \alpha$.

EXERCISE 78. Show that $\mathbf{On}$ is a proper class. (Hint: *Suppose* $\mathbf{On}$ *is a set and then show that it is transitive and well-ordered.*)

2.2. Transfinite induction and transfinite recursion

These two principles can be regarded as two sides of the same coin.

THEOREM 79. *(Transfinite Induction) Let* $\varphi(v)$ *be a formula of* $\mathcal{L}_\in$ *and be such that:*

(1) $\varphi(0)$*;*
(2) *if* $\varphi(\alpha)$*, then* $\varphi(\alpha + 1)$*; and*
(3) *if* β *is a limit ordinal and* $\forall \alpha < \beta$, $\varphi(\alpha)$*, then* $\varphi(\beta)$*.*

Then $\varphi(\gamma)$ *for all* $\gamma \in \mathbf{On}$*.*

REMARK. Transfinite induction is, essentially, a kind of least number principle, except here were are using the ordinals rather than the natural numbers. The least number principle says that if a property is not true of all the natural numbers, then there is some least natural number for which the property fails. Similarly, transfinite induction tells us that if there is a property not true of all the ordinals, there is some least ordinal at which the property fails.

PROOF. Suppose not; i.e., there is some $\delta \in \mathbf{On}$ such that $\neg\varphi(\delta)$. We want to find some β which will be the $\in$-least member of the class $\{\alpha \mid \neg\varphi(\alpha)\}$. Let $\gamma \in \{\alpha \mid \neg\varphi(\alpha)\}$ be arbitrary. If γ is the the $\in$-least member of $\{\alpha \mid \neg\varphi(\alpha)\}$

then we just let $\beta = \gamma$. But if γ is not the $\in$-least member of $\{\alpha \mid \neg\varphi(\alpha)\}$ then by Separation there is a non-empty set

$$\{\alpha \in \gamma \mid \neg\varphi(\alpha)\}.$$

Then since $\langle \gamma, \in \rangle$ is a well-ordering, it has a least $\in$-element, β, which is clearly also the $\in$-least element of $\{\alpha \mid \neg\varphi(\alpha)\}$.

REMARK. This opening is a little laboured as we needed to get a minimal element of a class rather than a set.

With β in hand, we now consider the different kinds of ordinal that β could be and show that each of them lead to a contradiction.

Suppose $\beta = 0$. Then we have $\neg\varphi(0)$ contradicting (1).

Suppose $\beta = \alpha + 1$. Then we have $\neg\varphi(\alpha + 1)$ and (by contraposition of) (2), we have $\neg\varphi(\alpha)$, contradicting the fact that β was supposed to be the least β such that $\neg\varphi(\beta)$.

Suppose β is a limit. Then we have $\varphi(\beta)$ and by (3), there is some $\alpha < \beta$ such that $\neg\varphi(\alpha)$ contradicting the minimality of β. $\square$

THEOREM 80. *(Transfinite Recursion) Let x be a set $\mathbf{G} : \mathbf{V} \to \mathbf{V}$ be a class function. Then there exists a function $\mathbf{F} : \mathbf{On} \to \mathbf{V}$ such that:*

$$\begin{aligned} \mathbf{F}(0) &= 0 \\ \mathbf{F}(\alpha+1) &= \mathbf{G}(\mathbf{F}(\alpha)) \\ \mathbf{F}(\beta) &= \bigcup\{\mathbf{F}(\alpha) \mid \alpha < \beta\} \text{ for limit } \beta. \end{aligned}$$

PROOF. Our strategy is to *construct* $\mathbf{F}$ using a sequence of set-sized approximations to $\mathbf{F}$ which we can then collect together to form $\mathbf{F}$. First we define the notion of approximation; then we show that this notion of approximation works; finally, we show how to use the approximations to construct $\mathbf{F}$.

Let us say that f *is an α-approximation* (of $\mathbf{F}$) if $f : \alpha \to \mathbf{V}$ such that for all $\beta < \alpha$,

- if $\beta = 0$, then $f(\beta) = 0$;
- if $\beta = \alpha + 1$, then $f(\beta) = \mathbf{G}(f(\alpha))$; and
- if β is a limit ordinal, then $f(\beta) = \bigcup\{f(\alpha) \mid \alpha \in \beta\}$.

REMARK. If f is an α-approximation, f is intended to behave exactly the same as $\mathbf{F}$ for every $\beta \in \alpha$.

CLAIM. For every α, there is a unique f which is an α-approximation.

PROOF. We prove this by transfinite induction, letting $\varphi(\beta)$ say that there is a unique f such that f is a β-approximation. We then work through the three cases from Theorem 79. Once we have done this our conclusion then follows.

(1.) Let $f = \{\langle 0, 0\rangle\}$.

(2.) Suppose that $\beta = \alpha + 1$ and that $\varphi(\alpha)$ is the case; i.e., there is a unique $f^\dagger$ such that $f^\dagger$ is an α-approximation. Let $f = f^\dagger \cup \{\langle \alpha + 1, \mathbf{G}(f(\alpha))\rangle\}$.

(3.) Suppose that β is a limit ordinal and that $\varphi(\alpha)$ holds for all $\alpha < \beta$; i.e., for each $\alpha < \beta$, there is a unique $f^\dagger$ which is an α-approximation. Let $h : \beta \to \mathbf{V}$ which returns the α-approximation for all $\alpha < \beta$. Then by Replacement there is a set

$$y = \{h(\alpha) \mid \alpha \in \beta\}.$$

We then let

$$f = \bigcup y.$$

It should be relatively clear that $(h(\alpha))(\beta) = (h(\gamma))(\beta)$ for all $\alpha, \gamma > \beta$; i.e., that the approximations behave exactly the same on inputs within their domain. This can also be proven using transfinite induction, but we leave this as an exercise. Thus $\bigcup y$ is also a function, as required. □

Finally, let

$$\begin{aligned}\mathbf{F} = \{\langle \alpha, x\rangle \ \mid \exists f \exists \beta (\quad & \alpha \in \beta \ \wedge \\ & f \text{ is an } \alpha\text{-approximation } \wedge \\ & f(\alpha) = x)\}\end{aligned}$$

This is the class function we were looking for. □

EXERCISE 81. Show that $h(\alpha) \subseteq h(\beta)$ for $\alpha \leq \beta$ (as define in the proof of Theorem 80). (Hint: *let* β *be arbitrary and use transfinite induction on* $\alpha < \beta$.)

EXAMPLE 82. Let us define the function $\mathbf{r} : \mathbf{On} \to \mathbf{V}$ as follows:

$$\begin{aligned} \mathbf{r}(0) &= \emptyset \\ \mathbf{r}(\alpha+1) &= \mathcal{P}(\mathbf{r}(\alpha)) \\ \mathbf{r}(\beta) &= \bigcup_{\alpha<\beta} \mathbf{r}(\alpha) \text{ for limit } \beta. \end{aligned}$$

Clearly $\mathcal{P} : \mathbf{V} \to \mathbf{V}$ so $\mathbf{r}$ is a class function by transfinite recursion. This function is particularly important. Indeed it can be shown that.

THEOREM 83. $\mathbf{V} = \{x \mid x = x\} = \{\mathbf{r}(\alpha) \mid \alpha \in \mathbf{On}\}$*; i.e., the range of the* $\mathbf{R}$ *function is the entirety of the universe. Note that we may define* $\{\mathbf{r}(\alpha) \mid \alpha \in \mathbf{On}\}$ *as* $\{x \mid \exists\alpha\ x \in \mathbf{r}(\alpha)\}$ *so it is a class.*

PROOF. Suppose not. Then let $\mathbf{C} = \{x \mid \neg\exists\alpha\ \mathbf{r}(\alpha) \mid \alpha \in \mathbf{On}\}$. By Foundation, $\mathbf{C}$ must have an $\in$-least member. Let x be such a set. Then for all $y \in x$, it should be clear that there is some least α such that $y \in \mathbf{r}(\alpha)$. For $y \in x$, let

$$\alpha(y) = \text{the least } \alpha \text{ such that } y \in \mathbf{r}(\alpha).$$

Then using Replacement let

$$\beta = \{\alpha(y) \mid \alpha \in y\}.$$

It should then be clear that $x \subseteq \mathbf{r}(\beta)$ and so $x \in \mathcal{P}(\mathbf{r}(\beta)) = \mathbf{r}(\beta+1)$ as required. □

This is why the universe of sets is often known as the *cumulative hierarcy*. Moreover we usually write $\mathbf{V}_\alpha$ instead of $\mathbf{r}(\alpha)$ to indicate the functions relation to the universe. This also allows us to define what is known as the rank of a set:

DEFINITION 84. The *rank* of a set x, $rank(x)$, is the least α such that $x \in \mathbf{V}_{\alpha+1}$.

Intuitively, the rank of a set if a measure of how deep the $\{\}$'s it has are. For example,

- the empty set, $\emptyset$, has rank 0; and
- the set $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ has rank 3.

EXERCISE 85. It should be clear that:

- $\mathbf{V}_0 = \emptyset$;
- $\mathbf{V}_1 = \{\emptyset\}$;

- $\mathbf{V}_2 = \{\emptyset, \{\emptyset\}\}$;
- $\mathbf{V}_3 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

Write the contents of $\mathbf{V}_4$? How big is $\mathbf{V}_5$?

2.3. Cardinals

Intuitively speaking, the cardinality of a set is a measure of its size. With finite collections, we can just count the number of objects in a set and represent its cardinality by a natural numbers. We want to do something similar in the case of infinite sets, but counting will no longer work.

We introduce here the machinery required to address this problem. First we define some special properties of functions that will allow us to say when one set has the same *cardinality* as another. Then we'll show how to find a canonical representative for each cardinality using ordinals. This will introduce us to the infamous Axiom of Choice.

DEFINITION 86. (i) We say that a function $\sigma : D \to C$ is ***surjective***, often abbreviated $\sigma : D \twoheadrightarrow C$ if for every element $c \in C$, there is some $d \in D$ such that $\sigma(d) = c$. In other words, the functionσ *exhausts* its codomain. (Sometimes such maps are called ***onto***.)

(ii) We say that a map $\sigma : D \to C$ is ***injective***, often abbreviated $\sigma : D \rightarrowtail C$, if for any two elements $d_1 \neq d_2 \in D$, $\sigma(d_1) \neq \sigma(d_2)$. Thus different object in the domain are always output to different object in the codomain. (Sometimes such maps are called ***one-to-one***.)

(iii) We say that a function $\sigma : D \to C$ is ***bijective***, often abbreviated $\sigma : D \cong C$, if it is both surjective and injective.

First we need to be more precise about what it means to have a certain cardinality.

DEFINITION 87. Let us say that two collections A and B have the *same cardinality*, which we abbreviate $A \approx B$ if there is a function $f : A \to B$ which is a bijection.

Thus if A is a collection of five sheep and B is a collection of five ducks, then there is a map f which takes each sheep to a duck such that:

(1) every duck d is such that there is a sheep s for which $f(s) = d$; and

(2) if $s_1 \neq s_2$ are different sheep then they are mapped to different ducks ($f(s_1) \neq f(s_2)$);

i.e., f is a bijection.

DEFINITION 88. We shall say that collection A has less than or the same cardinality as B, abbreviated $A \preceq B$, if there is an injection between A and B.

Thus if in the previous example there had been six ducks, we still could have got a map with (2.) satisfied, but one duck would have to have been left out.

So we should be able to see that these definitions are in accord with our intuitions about finite sets, but what about infinite ones. The following fact should be obvious, but it is instructive to prove it from our definitions.

FACT 89. *Let E be the set of even numbers and O be the set of odd numbers. Then $E \approx O$.*

PROOF. To show this, it suffices to define a bijection between E and O. Intuitively, we let f take the first even number to the fist odd number, the second even number to the second odd number and so on. More formally,

$$f(2n) = 2n + 1.$$

It should be clear that this is a bijection. □

The following fact may be more surprising.

FACT 90. *Let E be the set of even numbers and ω be the set of natural numbers. Then $N \approx E$.*

PROOF. Again we define the bijection. Let $f : N \to E$ be such that

$$f(n) = 2n.$$

This is clearly a surjection: every even number is output. Moreover it is an injection. Suppose not. Then there would be some even e such that for some $m \neq n$,

$$f(m) = e = f(n).$$

But then $2m = 2n$ and $m = n$: contradiction. □

This might be surprising given that finite collections A and B is A is a proper subset of B, $A \subsetneq B$ (i.e., $A \subseteq B$ but $A \neq B$), then $A \preceq B$. This is not always the case with infinite collections.

2.3.0.2. *Enumeration.* We shall say that a collection A is *countable* if $\omega \approx A$; i.e., if there is a bijection f between ω (the natural numbers and A). We shall say that f *enumerates* A since that's exactly what it does.

2.3.1. Choice. We now introduce the infamous Axiom of Choice.

9. Choice $\forall x(\forall y(y \in x \to \exists z(z \in y)) \wedge \forall y \forall z(y \in x \wedge z \in x \to \neg\exists w(w \in y \wedge w \in z)) \to \exists u \forall y(y \in u \to \exists! z(z \in y)))$

Intuitively, this says that if x is a family of non-empty open sets, each pair of members of which has an empty intersection, then there is a set u which intersects each member of x by exactly one object.

It's pretty ugly.

This is a much nicer version, but it was a little difficult to state the extra set-theoretic vocabulary we've developed.

9. Choice $\forall x(\forall y \in x \; y \neq \emptyset \to \exists f(f \text{ is a function } \wedge \; dom(f) = x \; \wedge \; \forall y \in x \; f(y) \in y))$.

Intuitively, this says that if x is family of non-empty sets, there is a function f which chooses an element from each member of the family x. We call such an f, a *choice function* for f.

THEOREM 91. $(9) \leftrightarrow (9')$.

PROOF. Exercise. (Harder) □

This next proof is important. It will give us our canonical representatives of cardinalities and from there we get our first proof that uses transfinite recursion and Choice.

THEOREM 92. *For every set x there is some α such that $x \approx \alpha$ and for all $\beta < \alpha$, $x \not\approx \beta$.*

PROOF. We make our first substantive use of transfinite recursion.

Let us take $\mathcal{P}(x)$ (the powerset of x) and apply Choice to get a choice function f for $\mathcal{P}(x)$. Thus, if $y \subseteq x$ and $y \neq \emptyset$, then $f(y) \in y$.

Let $\mathbf{G} : \mathbf{V} \to \mathbf{V}$ be the function such that

$$\mathbf{G}(y) = \begin{cases} f(x \backslash y) \cup y & \text{if } x \backslash y \neq \emptyset \\ x & \text{otherwise.} \end{cases}$$

Now using transfinite recursion, let $\mathbf{F} : \mathbf{On} \to \mathbf{V}$ be such that:

$$\begin{aligned} \mathbf{F}(0) &= 0 \\ \mathbf{F}(\alpha + 1) &= \mathbf{G}(\mathbf{F}(\alpha)) \\ \mathbf{F}(\beta) &= \bigcup \{\mathbf{F}(\alpha) \mid \alpha \in \beta\} \text{ for limit } \beta. \end{aligned}$$

REMARK. So the idea of this function is that at every stage we keep picking a new element of x that we haven't already picked: this is what G does (while it can); and a limit stages we just collect everything up that's come before. When we eventually collect up all of x, we just make the function return x forever after.

Now let γ be the least ordinal such that $\mathbf{F}(\gamma) = x$. Clearly $x \approx \gamma$, since $\mathbf{F} \restriction \gamma : \gamma \cong x$.

Thus we have shown that there is a γ such that there is some $\gamma \approx x$. By Transfinite Induction (in its Least Number Principle form) we then see that there must be some least α such that there is an $\alpha \approx x$, which is what we require. □

This licenses us to introduce a term picking out the least ordinal that bijects with a particular set x.

DEFINITION 93. The *cardinality* of x, denoted $|x|$, is the least α such that $\alpha \approx x$.

REMARK 94. For those with more logic background, the Axiom of Choice is equivalent over ZF to the downward Löwenheim-Skolem theorem.

2.4. Cantor's theorem - there are larger infinities!

2.4.1. Intuitive version. We are now ready to prove that there is more than one size of infinite collection. This is known as Cantor's theorem. The proof doesn't take that long, but we'll go through it somewhat informally to make it as clear as possible. Moreover, if we go through it too quickly, it can feel a little like a trick has been played on us.

Suppose we had countably many one pound coins and we lined them all up in a row. Each coin would either be heads-up or tails-up. Remember that since we have countably many coins, this means we can enumerate them with a function from N to the coins. Let us call this function c.

We may then represent this situation in a table as follows:

$c(0)$	$c(1)$	$c(2)$	$c(3)$	$c(4)$	$c(5)$	$c(6)$	
H	H	T	H	T	H	H	...

Now of course there are different ways that the coins could have been laid out. For example we might switch the third coin $c(2)$ from tails to heads.

Let us then consider the table which would result by placing each different arrangement of the coins in new rows of the table. Thus we get something like:

$c(0)$	$c(1)$	$c(2)$	$c(3)$	$c(4)$	$c(5)$	$c(6)$	...
H	H	T	H	T	H	H	...
T	T	T	T	T	T	T	...
H	T	T	H	T	T	H	...
T	H	T	T	T	H	H	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

We won't worry about the order in which the rows are filled in. We just want to ensure that every (infinite) arrangement of heads and tails is represented on exactly one row of the table.

Now our claim is that there are more rows than there are columns. Since there are infinitely many columns, this will suffice to show that there is more than one size of infinite collection.

So what does it mean for there to be more rows than columns? Since we know that the columns are countable, it will suffice to show that the rows are not countable. In order to do this, we must show that there is no bijective function r from the naturals to the rows, which provides an enumeration of them.

We shall demonstrate this by *reductio*. Thus, suppose that there was such an enumeration. Let us call it r. We might then represent this situation as follows:

	$c(0)$	$c(1)$	$c(2)$	$c(3)$	$c(4)$	$c(5)$	$c(6)$	...
$r(0)$	H	H	T	H	T	H	H	...
$r(1)$	T	T	T	T	T	T	T	...
$r(2)$	H	T	T	H	T	T	H	...
$r(3)$	T	H	T	T	T	H	H	...
$r(4)$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

To show that r cannot enumerate all of the rows, we are going to *construct* a row $r^\dagger$ that the enumeration r must miss. This is an arrangement of the coins.

We let $r^\dagger$ be the row constructed by flipping every coin down the diagonal of our table. So looking at the table above we get:

	$c(0)$	$c(1)$	$c(2)$	$c(3)$	$c(4)$	$c(5)$	$c(6)$	...
$r(0)$	**T**	H	T	H	T	H	H	...
$r(1)$	T	**H**	T	T	T	T	T	...
$r(2)$	H	T	**T**	H	T	T	H	...
$r(3)$	T	H	T	**H**	T	H	H	...
$r(4)$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Or more formally, we let $r^\dagger$ be such that the n^{th} column of $r^\dagger$ is:

- heads if the n^{th} column of the n^{th} row is tails; and
- tails if the n^{th} column of the n^{th} row is heads.

Now if $r^\dagger$ was in the enumeration, then there would have to be some n such that $r(n) = r^\dagger$. But this is not possible. We have defined $r^\dagger$ so that it is different from every $r(n)$ at exactly one place. Thus there can be no enumeration of the rows and there is a infinite cardinality which is *not countable*. We call such sizes *uncountable*.

REMARK 95. More informally, we observe that if there had been the same number of rows as columns, then the table above would be an infinite square. We have shown that there must always be an extra row beyond this.

2.4.1.1. *This also tells us ...* Consider each of the rows of the table. There is a sense in which each of them represents a particular subset of the natural numbers.

For example,

	$c(0)$	$c(1)$	$c(2)$	$c(3)$	$c(4)$	$c(5)$	$c(6)$	
$r(0)$	**H**	**H**	T	**H**	T	**H**	**H**	...

the 0^{th} row could pick used to pick out the set

$$\{0, 1, 3, 5, 6, ...\}.$$

We simply take the set be numbers of those coins which are facing heads-up. Moreover it should be clear that every set of natural numbers will be represented by exactly one of these rows.

Thus we have shown:

COROLLARY 96. *There are more sets of natural numbers than there are natural numbers.*

Moreover, it is possible to represent any real number by an infinitely long decimal number. We might then replace the coins in the example by a 10-sided dice. Then each of the rows could represent a real number between 0 and 1. We can then perform much the same trick as before to show that:

FACT 97. *There are more real numbers than natural numbers.*

The sizes of sets does not stop here either. We can repeat a (slightly more general) version of our argument above with the coins to get a collection which is even larger than the sets of all sets of natural numbers. Indeed we can repeat this indefinitely.

2.4.2. Formal version.

This one is much shorter.

THEOREM 98. *(Cantor) There is no surjection $f : X \to \mathcal{P}(X)$ for any set X.*

PROOF. Suppose for a contradiction that there is such an f and fix one. Using Separation, let

$$D = \{y \in X \mid y \notin f(y)\}.$$

Since f is a surjection, there must be some $d \in X$ such that $f(d) = D$.

But then we have

$$d \in D \leftrightarrow d \notin f(d) \leftrightarrow d \notin D$$

which is a contradiction. □

2.5. Infinite Cardinals & the Continuum hypothesis (CH)

2.5.1. The ℵs.

DEFINITION 99. Let κ be a cardinal. Then κ^+ is the least cardinal λ such that $\lambda > \kappa$.

REMARK 100. There's a tacit convention to refer to cardinals using κ and λ if they're available.

DEFINITION 101. Given two orderings $\mathcal{A} = \langle A, <_A \rangle$ and $\mathcal{B} = \langle B, <_B \rangle$, we say that $\mathcal{A}$ and $\mathcal{B}$ are *isomorphic*, abbreviated $\mathcal{A} \cong \mathcal{B}$, if there is a function $\sigma : A \to B$ such that:

- σ is a bijection; and
- for all $a_1, a_2 \in A$

$$a_1 <_A a_2 \;\Leftrightarrow\; \sigma(a_1) <_B \sigma(a_2).$$

As orderings (or more generally structures) go, if $\mathcal{A} \cong \mathcal{B}$, then we can do anything with $\mathcal{A}$ that we could already do with $\mathcal{B}$. For all intensive (in in particular, mathematical) purposes they are the same. We often say that are *identical up to isomorphism*. Thus, given a collection of orderings, it's convenient to deal remove redundant (isomorphic copies), so that every ordering in our collection is distinct, in the sense that is important.

We do this to the set of well-orderings of a particular cardinal κ using Choice below.

DEFINITION 102. Let κ be a cardinal. Then let

$$A = \{[R] \mid R \text{ is a well-ordering of } \kappa\}$$

where $[R] = \{S \subseteq \kappa \times \kappa \mid S \cong R\}$.

For each R, $[R]$ is the equivalence class (or set of) well-orderings of κ which are isomorphic to each other. These are known as *isomorphism types*.

Now let $f : A \to \{R \mid R \text{ is a well-ordering of } \kappa\}$ be a choice function for A and let

$$W = \{f([R]) \mid R \in A\}.$$

Clearly, W contains a single member of each *isomorphism type* of well-ordering of κ and thus, no redundant copies.

LEMMA 103. *Let κ be a cardinal, then $\kappa^+ = W$.*

PROOF. It will suffice to show there is a bijection $\sigma : W \to \kappa^+$. The proof has three parts, first we define σ; then we show σ is an injection; and finally we show that σ is a surjection.

We define σ using a new function **coll** (which *collapses* well-orderings). For a particular $R \in W$, let

- $\mathbf{coll}(0) = \emptyset$;
- $\mathbf{coll}(\alpha{+}1) = \begin{cases} \mathbf{coll}(\alpha) \cup \text{the } R\text{-least element of } field(R) \backslash \mathbf{coll}(\alpha) & \text{if such exists} \\ \mathbf{coll}(\alpha) & \text{otherwise} \end{cases}$;
  and
- $\mathbf{coll}(\beta) = \bigcup_{\alpha<\beta} \mathbf{coll}(\alpha)$ for β a limit ordinal.

REMARK. Essentially, we are just counting our way up the well-ordering R and assigning an ordinal to each stage as we go.

There must be some $\alpha < \kappa^+$ such that $\mathbf{coll}(\alpha + 1) = \mathbf{coll}(\alpha)$; i.e., we know that there are only κ many elements in the field of R. Thus we must run our of elements to count before then. This is known as a *fixed point*. If the fixed point occurred at a stage $\geq \kappa^+$, then we'd have shown that the field of R has cardinality κ^+ which would be a contradiction. We'll use a similar argument when we come to Kripkean truth definitions.

Let $\sigma(R) = \mathbf{coll}(\alpha)$ for the least α such that $\mathbf{coll}(\alpha + 1) = \mathbf{coll}(\alpha)$.

(*Injection*) It should be clear that for $R_1, R_2 \in W$ if $\sigma(R_1) = \sigma(R_2)$ (i.e., they collapse to the same ordinal), then $R_1 \cong R_2$. But W has no redundant isomorphic copies, so $R_1 = R_2$ which is what we needed to show.

(*Surjection*) Let $\alpha < \kappa^+$ be arbitrary. We just show that there is some $R \in W$ such that $\sigma(W) = \alpha$. Since $\alpha < \kappa^+$, we know that $|\alpha| < \kappa^+$. For simplicity (we'll leave the more complicated case as an exercise) let us assume that $|\alpha| = \kappa$. Then there is some $g : |\alpha| \cong \kappa$. Now we define an ordering (κ, S) such that for $\beta, \gamma \in \kappa$

$$\beta S \gamma \;\Leftrightarrow\; g(\beta) \in g(\gamma).$$

It should be clear that $(\kappa, S) \cong (\alpha, \in)$. Also since S is a well-ordering of κ, there must be some $R \in W$ such that $(\kappa, R) \cong (\kappa, S)$. Thus, it can be seen that $\sigma(R) = \alpha$.

Thus, we have a bijection and the proof is complete. □

EXERCISE 104. In the surjection case of the proof of Lemma 105, what happens if $|\alpha| < \kappa$. Complete the proof for those cases where this occurs.

DEFINITION 105. Let $\aleph : \mathbf{On} \to \mathbf{On}$ be such that:

- $\aleph_0 = \omega$;
- $\aleph_{\alpha+1} = (\aleph_\alpha)^+$; and
- $\aleph_\beta = \bigcup_{\alpha<\beta} \aleph_\alpha$.

EXERCISE 106. How is Definition 105 justified?

2.5.2. The Continuum Hypothesis. So we know a few interesting things at this point:

(1) $\omega = \aleph_0$ is the least infinite cardinal;
(2) $\aleph_1$ is, by definition, first infinite cardinal larger than ω; and
(3) $2^{\aleph_0}$ is greater than ω.

This leads to the following very natural question.

PROBLEM 107. (The Continuum Hypothesis) $\aleph_1 = 2^{\aleph_0}$.

After over one hundred years of set theoretic research, this question still lacks an answer. We shall see in the next two sessions that we have learned a great deal about this question and that its recalcitrance to solution has changed the way we think about set theory.

Perhaps the most important thing to note about it at this point is how, so to speak, embarrassing it is. It's surely the first question one wants to ask about infinite cardinalities. If you've just learnt that there are larger infinities, you want to know which one is the next largest.

CH is however, true in some restricted contexts. For example, CH is true for *closed* sets of reals as was noticed by Cantor. Moreover, in recent times it has been demonstrated that CH holds of all the sets that can be defined in second order arithmetic, so long as some particularly large cardinals exist.

Part 2

Overview of advanced set theory

CHAPTER 3

Advanced topics in set theory

At the end of the previous section, we encountered the problem of the continuum hypothesis (CH).

PROBLEM 108. (CH) $\aleph_1 = 2^{\aleph_0}$.

As we saw, once we have developed just a little of the theory of infinite cardinality, it is a very natural question to ask. We're wondering if the next largest infinite cardinal after ω is the cardinality of it powerset, which we know is larger by Cantor's theorem. It is so natural that it doesn't seem like a *merely* mathematical question. It asks a very simple question about the nature of the world of infinite cardinalities that Cantorian set theory opened up. It is a question with deep philosophical implications.

It has, however, proved extremely resistant to solution.

This week we are going to look at some of the progress that has been made toward addressing it. We'll break this into the following sections:

(1) Inner models, constructibility & CH; and
(2) Outer models, forcing and CH.

Each of these sections also describes a major contemporary set theoretic technique and programme. However, I should note that this list is not exhaustive and that these programmes are not exclusively concerned with addressing CH.

This week, we are going to leave a large number of black boxes unopened. Rather that giving full proofs of the relevant theorems, my goal is to convey the urgency of the problems and the changes in philosophical perspective provided by some of the solutions.

3.1. Inner models, constructibility & CH

3.1.1. The problem. A solution to the CH problem was the first of Hilbert's famous problems. In this section, we are going to give an overview of Gödel's work on what is known as the constructible hierarchy and what it tells us about CH.

But first let us take stock as to what we already know about the problem. We can articulate it, but very little more.

Ideally, if CH were false, then we'd be able to refute it from the axioms of ZFC; i.e.,

$$ZFC \vdash \neg CH$$

$\neg CH$ would then be a theorem like, say, transfinite induction. Given that we accept ZFC is correct, we'd be compelled to accept $\neg CH$ too.

In this section we are going to use Gödel's construction to show that

$$ZFC \nvdash \neg CH$$

under the assumption that ZFC is consistent. Clearly this is not an answer to the problem, but it is a kind of progress: we've learnt that ZFC does not rule out CH. So at this point, we might still wonder if we could prove CH from ZFC.

Spoiling the surprise, we'll learn in the next section that assuming ZFC is consistent, we also have

$$ZFC \nvdash CH.$$

Assuming ZFC is consistent, it's obvious that we can't prove both CH and $\neg CH$ from ZFC, but it might be surprising that we can prove neither of them. However, after Gödel's incompleteness theorems we know that this kind of thing can occur. Gödel showed us that there must be sentences γ from $\mathcal{L}_\in$ such that

$$ZFC \nvdash \gamma \text{ and } ZFC \nvdash \neg\gamma;$$

i.e., we can neither prove nor refute γ - ZFC is incomplete.

Even so, CH seems like a very different sort of sentence to the self-referential trickery involved in a Gödel sentence like γ. It's a very natural question to ask and it seems like it should have an answer. So it is disappointing that ZFC cannot do this. In the final section of this week, we'll ask whether this incompleteness can be addressed.

3.1.1.1. *The Strategy.* We shall now run through an overview of Gödel's proof that

$$ZFC \nvdash \neg CH.$$

Let's think a little about what this means: it says that there's not derivation from the axioms of ZFC to $\neg CH$. In other words, $ZFC \cup \{CH\}$ is consistent. Abusing notation slightly, we'll follow convention and write this as follows:

$$Con(ZFC + CH).$$

Recall the soundness theorem for first order logic:

THEOREM 109. *Let Γ be a set of sentences in some language $\mathcal{L}$. Then*

$$\exists \mathcal{M}\ \mathcal{M} \models \Gamma \ \Rightarrow\ Con(\Gamma).$$

REMARK. Of course we usually state this in its contrapositive form. Also note, that I am writing $\mathcal{M} \models \Gamma$ to mean that for all $\gamma \in \Gamma$, $\mathcal{M} \models \gamma$.

This tells us that it will suffice for our purposes to find a model $\mathcal{M}$ such that

$$\mathcal{M} \models ZFC + CH.$$

In essence, this is what we are going to do. However, there is a slight technical hitch that we need to take care of due to Gödelian issues around proving consistency that we need to deal with first. The reader less familiar issues may wish to skip the next subsection on first reading.

3.1.1.2. *A Gödelian hitch.* Suppose we were able to prove that there is a model $\mathcal{M}$ such that

$$\mathcal{M} \models ZFC + CH.$$

Then using Theorem 109, we have established $Con(ZFC+CH)$ and $Con(ZFC)$ as well.

Now let us consider the perspective from which we proved this theorem. One of the main attractions of using ZFC is that we are supposed to be able to any form of mathematics (including set theory) using it. This kind of theorem pushes the boundaries of this claim.

Recall Gödel's second incompleteness theorem:

THEOREM 110. *Let Γ be a theory extending ZFC. Then if Γ is consistent, then*

$$\Gamma \nvdash Con(\Gamma).$$

Since ZFC is a trivial extension of itself, we see that

$$ZFC \nvdash Con(ZFC).$$

But this, then, means - using soundness again - that

$$ZFC \nvdash \text{“}\exists \mathcal{M}\ \mathcal{M} \models ZFC + CH.\text{”}$$

So at this point, our hopes for this strategy may appear to be dashed. Fortunately, it turns out that there is an easy way around this.

Rather than showing there is a model $\mathcal{M}$ such that $\mathcal{M} \models ZFC$, we will show that for any finite subset $\Delta \subseteq ZFC + CH$,

$$ZFC \vdash \text{“}\exists \mathcal{M}\ \mathcal{M} \models \Delta\text{”}.$$

This is something we can show and it is not exposed to the Gödelian trouble.

REMARK. One might think there is trouble there. I appear to have shown that ZFC can show that every finite subset of $ZFC + CH$ is true in some $\mathcal{M}$. Why can't I think apply the compactness theorem to show in ZFC that there is a model $\mathcal{M}$ such that $\mathcal{M} \models ZFC + CH$.

The problem here is my informal reasoning makes a different assumption. It assumes we have shown that

$$ZFC \vdash \text{“}\forall_{\text{finite}} \Delta \subseteq ZFC + CH\ \exists \mathcal{M}\ \mathcal{M} \models \Delta\text{”}$$

which we will not and cannot show. The key point is that on pain of inconsistency we cannot bring that quantification over finite subsets of $ZFC + CH$ into the context of ZFC; or more formally, this is an example of the ω-incompleteness of ZFC.

Once we have shown this we can get to our desired result using the following lemma.

LEMMA 111. *Let S and T be theories extending ZFC. suppose that for all finite $\Delta \subseteq S$*

$$T \vdash \text{“}\exists \mathcal{M}\ \mathcal{M} \models \Delta.\text{”}$$

Then $Con(T) \to Con(S)$.

PROOF. Since T extends ZFC, the completeness theorem for first order logic is a theorem of T. Thus our first assumption tells us that for any finite subset $\Delta \subseteq S$,

$$T \vdash Con(\Delta).$$

Now suppose that $\neg Con(S)$. Then there is some finite $\Delta \subseteq S$ such that

$$\Delta \vdash \varphi \wedge \neg\varphi.$$

Since T extends ZFC, it can be shown since T is Σ^0_1-complete, that

$$T \vdash \text{“}\Delta \vdash \varphi \wedge \neg\varphi\text{”}$$

and so by the soundness theorem in T, we have

$$T \vdash \text{“}\forall \mathcal{M}\ (\mathcal{M} \models \Delta \ \rightarrow \ \mathcal{M} \models \varphi).\text{”}$$

But by assumption we have

$$T \vdash \text{“}\exists \mathcal{M}\ \mathcal{M} \models \Delta.\text{”}$$

Thus, $T \vdash$ “$\exists \mathcal{M}\ \mathcal{M} \models \varphi \wedge \neg\varphi$”, which is a contradiction since ZFC tell us that no model can make both a sentences and its negation true. □

So letting T be ZFC and S be $ZFC + CH$, we get the proposition that we shall demontrate in sketch form:

$$Con(ZFC) \rightarrow Con(ZFC + CH).$$

REMARK. This is why I made the constant hedge “assuming ZFC is consistent” above. So this formulation above is another way of saying that

$$ZFC \nvdash \neg CH$$

assuming ZFC is consistent.

3.1.2. Definability and L. Our goal now is to outline the model that we'll use for the proof. To describe it very briefly, we are going to construct the *thinnest* model of set theory that we can. This is known as L. Recall the fact that all of the sets could be deliverd by using the following function defined by transfinite recursion.

$$\begin{aligned}\mathbf{V}_0 &= \emptyset\\ \mathbf{V}_{\alpha+1} &= \mathcal{P}(\mathbf{V}_\alpha)\\ \mathbf{V}_\beta &= \bigcup_{\alpha<\beta}\mathbf{V}_\alpha \text{ for limit } \beta.\end{aligned}$$

At the all important successor stages we ensured that we added all of the subsets of the previous level to the construction. We might, for the moment, think of this as providing the *fattest* model of set theory.

Gödel's insight was to think of a different, seemingly more conservative, method of generating sets at the successor stages. The idea that rather than taking *all* of the subsets of the previous level, we only take those subsets which are *definable* using the previous level.

We shall show that this way of doing things gives us a model of ZFC and is such that the continuum hypothesis is true in it.

In order to get a good feel for it we'll spend a little time introducing some useful notions.

3.1.2.1. *Definability over a model.*

DEFINITION 112. Let $\mathcal{L}$ be a language and $\mathcal{M}$ be a model for $\mathcal{L}$. We say that $X \subseteq M$, the domain of $\mathcal{M}$, is *definable over* $\mathcal{M}$ if there is some formula $\varphi(v_0, v_1, ..., v_n)$ of $\mathcal{L}$ and $m_1, ..., m_n \in M$ such that

$$X = \{n \in M \mid \mathcal{M} \models \varphi(n, m_1, ..., m_n)\}.$$

So the idea here is that a set X is definable in a model, $\mathcal{M}$, if it's the extension of some formula (with parameters from the model) in $\mathcal{M}$.

Now we want to use this notion of definability to generate new sets. To do this we use the following function:

DEFINITION 113. Let $\mathcal{M}$ be a model of some language $\mathcal{L}$. Let

$$Def(\mathcal{M}) = \{X \subseteq M \mid X \text{ is definable over } \mathcal{M}\}$$

3.1.2.2. *Nice, big models.* Now the next thing we need to note is that the target model **L** that we are building toward will not, strictly speaking, be a model. It is not a model in the sense that it is too big or long. Like **V** it will have the length of the ordinals and as such it will be a proper class.

In fact it will just be a (possibly) thinner version of **V**. So the domain will be (possibly) restricted but we can use the ordinary $\in$ relation of **V**. We are just (possibly) restricting the domain of quantification.

This gives us a way of getting around things. Since **L** won't be a model in the strict sense, we won't be able to define the satisfaction relation for it. Nonetheless, we can still talk about what's true in **L** since we can talk about the its domain of quantification.

The basic idea here is that we are going to figure out if a sentence φ is true in **L**, we just restrict all its quantifiers to **L**.

More formally we write $\varphi^{\mathbf{L}}$ be be the result of this translation which is defined recursively as follows:

- if $\varphi := x \in y$, them $\varphi^{\mathbf{L}} = \varphi$;
- if $\varphi := \neg\psi$, then $\varphi^{\mathbf{L}} = \neg(\psi^{\mathbf{L}})$;
- if $\varphi := \psi \wedge \chi$, then $\varphi^{\mathbf{L}} = \psi^{\mathbf{L}} \wedge \chi^{\mathbf{L}}$; and
- if $\varphi := \forall x\psi(x)$, then $\varphi^{\mathbf{L}} = \forall x(\mathbf{L}(x) \to \psi(x))$.

All the action is at the quantifier stage. It should then be clear that

$$\varphi \text{ is true in } \mathbf{L} \;\Leftrightarrow\; \varphi^{\mathbf{L}}.$$

Our goal will then be to show that for every $\delta_1, ..., \delta_n \in ZFC + CH$, we have

$$ZFC \vdash \delta_1^{\mathbf{L}} \wedge ... \wedge \delta_n^{\mathbf{L}}.$$

A simple adaptation of Lemma 111 will then get us to our target

$$Con(ZFC) \to Con(ZFC + CH).$$

3.1.2.3. *Defining* **L**. We are now ready to define **L** using transfinite recursion.

DEFINITION 114. Let $\mathbf{L}_{\cdot} : \mathbf{On} \to \mathbf{V}$ be defined by transfinite recursion as follows:

$$\begin{aligned}
\mathbf{L}_0 &:= \emptyset \\
\mathbf{L}_{\alpha+1} &:= Def(\mathbf{L}_\alpha) \\
\mathbf{L}_\beta &:= \bigcup_{\alpha<\beta} \mathbf{L}_\alpha.
\end{aligned}$$

This is known as the *constructible hierarcy* and the sets in it are called *constructible.*

The key change here is in the successor clause. We are no longer adding all the subsets of the previous level, but merely those sets which are definable from the previous level.

With a little work, it's relatively easy to see that all the finite levels will be the same; i.e., $\forall n \in \omega$

$$\mathbf{L}_n = \mathbf{V}_n$$

and thus

$$\mathbf{L}_\omega = \mathbf{V}_\omega.$$

However, things come apart at $\omega + 1$. In $\mathbf{V}_{\omega+1}$ we add all the subsets of $\mathbf{V}_\omega$. Since $|\mathbf{V}_\omega| = \omega$, this means we are adding $|\mathcal{P}(\omega)| = 2^{\aleph_0}$ sets to the hierarchy.

On the other hand, $\mathbf{L}_{\omega+1}$ merely consists of those sets which are definable over $\mathbf{L}_\omega$. So since there are countably many formulae and countable many parameters to use, we can see that $|\mathbf{L}_{\omega+1}| = \aleph_0$.

So this clearly tells us that $\mathbf{L}_{\omega+1} \neq \mathbf{V}_{\omega+1}$. However, it does not also tell us that $\mathbf{V} \neq \mathbf{L}$. The essential reason for this is that the $\mathbf{L}$ function might, so to speak, catch up with $\mathbf{V}$. For example, although only countably many subsets of $\mathbf{L}_\omega$ are added to $\mathbf{L}_{\omega+1}$, more subsets of $\mathbf{L}_\omega$ will be added at stage $\mathbf{L}_{\omega+2}$. In the next section, we'll see how far we have to go into $\mathbf{L}$ before no new subsets of $\mathbf{L}_\omega$ are added.

3.1.2.4. *Some properties of* $\mathbf{L}$.

FACT 115. *(i) For all* α, $|\mathbf{L}_\alpha| = |\alpha|$.

REMARK 116. (ii) I won't provide a formal proof, but it's not too hard to see in the countable case and then generalise from there. If you consider what we added when we formed $\mathbf{L}_{\omega+1}$, we used formulae of the language of set theory to define subsets of $\mathbf{L}_\omega$. We are also permitted to use parameters from $\mathbf{L}_\omega$ (i.e., names for things in $\mathbf{L}_\omega$).

However, note that every member in $\mathbf{L}_\omega$ is finite. Indeed every member of ever membe of $\mathbf{L}_\omega$ is finite and so on. We call such objects *heridarily finite*: they contain finite elements, so to speak, all the way down. This means we don't actually need the names: we can form all the finite sets using pairing and union terms - the language of set theory is sufficiently rich to be able to describe all of its heritarily finite elements. Now this means that we are adding at most one subset of $\mathbf{L}_\omega$ for every formula of $\mathcal{L}_\in$; thus, it can be seen that $|\mathbf{L}_{\omega+1}| = \aleph_0 = |\omega + 1|$.

A similar argument then shows that $|\mathbf{L}_{\omega+2}| = \aleph_0 = |\omega + 2|$. If we keep going through the countable ordinals $\alpha < \omega_1$, there will come a point where we do need parameters, however, there will only be countable many of them. So enriching our language with names for elements of $\mathbf{L}_\alpha$ will result in a language which also has countably many formulae.

3.1.3. Condensation & $Con(ZFC) \to Con(ZFC + CH)$.

3.1.3.1. *The condensation lemma.* The next lemma is the key to showing that CH hold in $\mathbf{L}$. It will tell us how far we need to go into $\mathbf{L}$ before subsets of a certain set will be added.

Before we prove this, we need a few technical pieces of information.

DEFINITION 117. Let $\mathcal{M} = \langle M, \sigma \rangle$ and $\mathcal{N} = \langle N, \rho \rangle$ be arbitrary models of some language $\mathcal{L}$ with $M \subseteq N$. We say that $\mathcal{M}$ is an *elementary submodel* of $\mathcal{N}$, abbreviated $\mathcal{M} \prec \mathcal{N}$ if for all $m_1, ..., m_k \in M$ and $\varphi(v_1, ..., v_k)$ from $\mathcal{L}$, we have:

$$\mathcal{M} \models \varphi(m_1, ..., m_k) \;\Leftrightarrow\; \mathcal{N} \models \varphi(m_1, ..., m_k).$$

REMARK 118. Let's descirbe a little more informally what this says. Suppose we have $\mathcal{M}$ and $\mathcal{N}$ where the domain of $\mathcal{M}$ is a subset of the domain of $\mathcal{N}$. Then if $\mathcal{M}$ is an elementary submodel of $\mathcal{N}$ we can take any tuple of objects from $\mathcal{M}$ we like and say whatever we like about that tuple using our language $\mathcal{L}$. Then if this was true in $\mathcal{N}$ then it's true in $\mathcal{M}$ too; and vice versa. So from the point of view of what we can express in $\mathcal{L}$ and the objects of $\mathcal{M}$ there is no difference between $\mathcal{M}$ and $\mathcal{N}$.

We might then wonder if $\mathcal{M}$ and $\mathcal{N}$ would need to be identical (in the sense of isomorphic) for this to hold. The downward Löwenheim-Skolem tells us otherwise.

FACT 119. *Let $\mathcal{N}$ be a model of some $\mathcal{L}$. Let $X \subseteq N$. Then there is a model $\mathcal{M}$ of $\mathcal{L}$ such that:*

(1) $\mathcal{M} \prec \mathcal{N}$*;*
(2) $X \subseteq M$*;*
(3) $|M| = \max\{\aleph_0, |X|, |\mathcal{L}|\}$.

We shall make use of this in a moment, but we'll also need the following technical fact.

FACT 120. *Let M be set such that $(\text{Extensionality})^M$ holds. Then there is a unique transitive set N, known as the transitive collapse of M, such that:*

- *$\pi : M \cong N$ for some π; and*
- *for any transitive set $x \in M$, $\pi(x) = x$.*

REMARK. Recall that a transitive set M is such that every member of every member of every member etc of M is already in M. Transitive models are particularly respectable in that the truth of relatively simply (but natural) sentences are preserved between a transitive model M and the real world of set theory, $\mathbf{V}$.

In order to prove this, we need to, so to speak, we need to remove all the gaps that might be in M. For example, suppose we had $x, y \in M$ where $y = \{x, z\}$ and $z \notin M$. Then M is not transitive. However, we could replace y by $\{x\}$ since for all M knows, this is what y is anyway. Thus our collapse function π would be such that

$$\pi(y) = \{x\}.$$

To fully define $\pi : M \cong N$ where N is transitive, we use a function constructed using transfinite recursion and which is similar to the function we used in showing that every set has a cardinality.

FACT 121. *If M is a transitive model (i.e. set), then for $x \in M$*

$$(x \in \mathbf{L})^M \;\Leftrightarrow\; x \in \mathbf{L}.$$

REMARK 122. The basic idea here is that our definition of being in $\mathbf{L}$ is sufficiently simply and natural that any transitive model will agree about it.

LEMMA 123. *(Condensation) Suppose $M \prec \mathbf{L}_\alpha$ for some α and let N be the transitive collapse of M. Then*

$$N = \mathbf{L}_\beta$$

for some $\beta \leq \alpha$.

REMARK. This proof requires some more technical background to give a proper sketch, but we might make a few remarks about *why* it is true.

- The basic idea is that the definition of $\mathbf{L}$ is so simple that any two respectable (i.e., transitive) models will agree on what they think is $\mathbf{L}$, even if they disagree about other matters.

- So if we take an elementary submodel M of some $\mathbf{L}_\alpha$, then all the properties we can ascribe to elements of M (in the language $\mathcal{L}_\in$ with parameters) will be the same according to $\mathbf{L}_\alpha$ and M.
- However, M is not sufficiently respectable so we then collapse M into the isomorphic and transitive N using some $\pi :: M \to N$.
- Now since N is respectable, the N and $\mathbf{V}$ (the most respectable model) agree on what they think $\mathbf{L}$ is.
- But N thinks (like $\mathbf{L}_\alpha$ and M) that it is some initial segment of $\mathbf{L}$, so that is indeed what it is.

With these tools in hand, we can prove our target theorem.

THEOREM 124. *CH is true in $\mathbf{L}$.*

PROOF. It suffices to show that $\mathcal{P}(\omega) = \aleph_1$ is true in $\mathbf{L}$. Let's work inside $\mathbf{L}$; i.e., from $\mathbf{L}'s$ perspective.

Let $x \subseteq \omega$. It will suffice to show the following claim:

CLAIM. $x \in \mathbf{L}_\alpha$ for some $\alpha < \aleph_1$.

This suffices since it shows that $\mathcal{P}(\omega) \subseteq \mathbf{L}_{\aleph_1}$ where we know by Fact 115 that $|\mathbf{L}_{\aleph_1}| = \aleph_1$.

PROOF. (of claim) Clearly there is some β such that $x \in \mathbf{L}_\beta$. By the downward Löwenheim-Skolem theorem let $M \prec \mathbf{L}_\beta$ be such that:

(1) $x \in M$; and
(2) $|M| = \aleph_0$.

Let N be the transitive collapse of M. Then by Lemma 123 and Fact 115, we can see that $N = \mathbf{L}_\alpha$ for some $\alpha < \aleph_1$. Moreover, since x is transitive it will collapse to itself; thus, $x \in \mathbf{L}_\alpha$ as required.

□

□

THEOREM 125. *$Con(ZFC) \to Con(ZFC + CH)$.*

Nowwe can also articulate the fact that $\mathbf{V} = \mathbf{L}$ in the language of set theory. Thus we also have,

THEOREM 126. *$Con(ZFC) \to Con(ZVC + \mathbf{V} = \mathbf{L})$.*

3.1.4. Inner models.

FACT 127. *(i) If* M *is an inner model, then*

$$(x \in \mathbf{L})^{\mathbf{M}} \Leftrightarrow x \in \mathbf{L}.$$

3.2. Outer models, forcing and CH

3.2.1. The problem. This time we are going to establish that:

$$ZFC \nvdash CH$$

assuming that ZFC is consistent.

But to get things started, we'll work with start with something a little less complicated. We'll show that

$$ZFC \nvdash \mathbf{V} = \mathbf{L}$$

assuming ZFC is consistent. Theorem 126 tell us that we don't have $ZFC \vdash \mathbf{V} \neq \mathbf{L}$, but we might still have wondered if the constructible sets are all the sets and whether ZFC can tell us this. We shall see that this is not the case.

3.2.1.1. *A hitch.* The obvious thing to try is repeating the strategy of the previous proof. Thus, we might try to define a class model, say $\mathbf{R}$, such that for all finite $\delta_1, ..., \delta_n \in ZFC + \mathbf{V} \neq \mathbf{L}$, we have

$$ZFC \vdash \delta_1^{\mathbf{R}} \wedge ... \wedge \delta_n^{\mathbf{R}}.$$

But if we do this then we'll show in ZFC

$$(\mathbf{V} \neq \mathbf{L})^{\mathbf{R}}.$$

Thus there is some $x \in \mathbf{R}$ such that

$$(x \notin \mathbf{L})^{\mathbf{R}}$$

and by Lemma 127, we see that

$$x \notin \mathbf{L};$$

i.e., this is true in the *real* universe $\mathbf{V}$. But since x is then a non-constructible set, we would have established that

$$ZFC \vdash \mathbf{V} \neq \mathbf{L}$$

which is contrary to 126, unless ZFC is inconsistent.

So this tells us that this strategy cannot work. Indeed, until Cohen came along it was thought that this was evidence that no proof of $Con(ZFC) \rightarrow Con(ZFC + \mathbf{V} \neq \mathbf{L})$ would ever be provided [Shepherdson, 1952].

3.2.1.2. *A solution.* Fortunately, there is a way around this (actually there are a number of them). We going to exploit the approach used by Kunen in [Kunen, 2006]. Rather than defining a class model, we are going to start with a small model and use that to create a new model with the properties we want. This technique is known as *forcing* [Cohen, 1966].

To make a model in which $\mathbf{V} \neq \mathbf{L}$, we are going to add a new elemement that is not constructible. This is known as a *generic element*. But in contrast to the previous section, we are going to use a countable model, M. Indeed, we'll take it that the model is also transitive: thus, a countable transitive model.

REMARK. Note where the argument for class models breaks down for countable transitive models. Suppose we have added a generic element G to a countable transitive model M and call the result $M[G]$. Since M is countable, there must be some $\alpha < \omega_1$ such that $M \cap \mathbf{On} = \alpha$. We're also going to ensure that $M[G] \cap \mathbf{On} = \alpha$ too. But in adding a generic element to M, we'll have shown that

$$(G \notin \mathbf{L})^{M[G]}.$$

So we might then think that

$$G \notin \mathbf{L}.$$

But this is too strong. All we are allowed to conclude is that

$$G \notin \mathbf{L}_\alpha.$$

Moreover, since α is countable, there may be some β with $\alpha < \beta < \omega_1$ such that $G \in \mathbf{L}_\beta$; i.e., G might get into the real $\mathbf{L}$ a little later on.

3.2.1.3. *The metalogical technicalities.* This section may be best skipped on first reading.

The main proof actually establishes something of the following form:

THEOREM 128. *For all* $\varphi_1, ..., \varphi_n \in ZFC + V \neq L$ *there exists* $\psi_1, ..., \psi_m \in ZFC$ *such that*

$$ZFC \vdash \text{“}\exists M(M \models \bigwedge_{1 \leq i \leq m} \psi_i \wedge M[G] \models \bigwedge_{1 \leq i \leq n} \varphi_i)\text{”}.$$

From this we immediately get:

COROLLARY 129. *For all* $\varphi_1, ..., \varphi_n \in ZFC + \mathbf{V} \neq \mathbf{L}$

$$ZFC \vdash \text{“}\exists N\ N \models \bigwedge_{1\leq i\leq n} \varphi_i.\text{”}$$

From which we can generalise Lemma 111 to establish that

$$Con(ZFC) \to Con(ZFC + \mathbf{V} \neq \mathbf{L}).$$

REMARK 130. Note that unlike the $\mathbf{L}$ construction, we are not necessarily working with a particular model through the proof: it could be different for different finite subsets of $ZFC + \mathbf{V} \neq \mathbf{L}$.

Despite these metalogical technicalities, the actual construction has a pleasing combinatorial feel.

3.2.2. A generic object from a partial order.

3.2.2.1. *Intuitive idea.* So our first goal will be to make a model $M[G]$ in which $\mathbf{V} \neq \mathbf{L}$. For this, all we need to do is show that there is some subset G of the natural numbers in $M[G]$ that wasn't already in M. This is known as a *Cohen real.* Then since M and $M[G]$ agree about what's in their versions of $\mathbf{L}$, $G \notin \mathbf{L}$. There are two things we need to do:

(1) *adjoin* the generic element, G; and
(2) retain *control* of the model, $M[G]$.

The basic idea of the approach can be outlined as follows:

- We can't just assume that there is such a G, we need to establish that it exists.
- So we're going to consider a partial order $\mathbb{P} \in M$ consisting of all the different ways one might go about constructing such a G. Every $p \in \mathbb{P}$ will be a finite (since M is countable) part of a potential generic G; and if $q \leq p$, then q may be a larger part of such a G. So when we add a Cohen real, p will (essentially) consist in a finite subset of the natural numbers.
- This will allow us to *control* $M[G]$ in the sense that these finite pieces $p \in \mathbb{P}$ of information about potential G's tells us about all of the potential G's that could be constructed from p. p thus *forces* every

suitable $M[G]$ to be a certain way. Moreover, M will be in a position to *know* these facts.

- This *control* allows us to verify the axioms of ZFC are upheld in $M[G]$.
- The *adjunction* allows us to verify that $\mathbf{V} \neq \mathbf{L}$.

This technique can then be generalised to give other kinds of models of ZFC that contain new elements.

3.2.2.2. *Making this goal more precise.* Rather than take a long lead up to how we might do this, let's just have a look at the target theorems. These won't make proper sense yet as we haven't given proper definitions of all the terms, but we can still make a few informal remarks at this point, which should make things a little clearer:

(1) We are going to use the symbol $\Vdash$ to represent the forcing relation. We shall write $p \Vdash \varphi$ which we want to mean that p forces φ to the be the case in our model $M[G]$ if $p \in G$.
(2) Our generic element G will in fact be a subset of the partial order $\mathbb{P}$. It will have some special properties that we'll further clarify later on.

Our main theorem can be stated as follows:

THEOREM 131. *Let M be a countable transitive model, $\mathbb{P} \in M$ be a partial order and $G \subseteq \mathbb{P}$ be generic for M. Then*

$$M[G] \models \varphi(c_1^G, ..., c_n^G) \Leftrightarrow \exists p \in G\ M \models \text{“}p \Vdash \varphi(c_1, ..., c_n).\text{”}$$

This theorem is all about control. It tells us both that:

(1) If we can state some proposition φ about how things are in $M[G]$, then
  (a) there is some partial information $p \in G$ (a preliminary stage in $G's$ construction) such that p forces that φ to be the case (p while only having partial information about G has sufficient information to ascertain φ); and
  (b) the fact that p forces φ can be ascertained in M.
(2) If there is some $p \in G$ such that M thinks that p forces some proposition φ, then φ is true in $M[G]$.

The surprising thing about this result is that G is not even in M and yet we can tell a great deal about what any M would be like using the forcing relation inside M.

This will allow us to ensure that some facts about M are preserved in $M[G]$. In particular, this is what allows us to show that:

THEOREM 132. *If M is a countable transitive model of ZFC, then $M[G] \models ZFC$.*

REMARK 133. We haven't yet defined the forcing relation or what a generic set is. This is where Cohen's amazing insight came. It involves a delicate balancing act between the control of the shape of the generic object and the ability of M to work out facts about the forcing relation.

3.2.2.3. *Ensuring that M can know facts about forcing.* I'm not going to work through the full details of these definitions and proofs, but I'd like to provide a feel for how the construction works from a strategic point of view.

We are first going to look at the forcing relation $\Vdash$. The guiding idea is figure out what we could say about the model $M[G]$ if we know that $p \in G$.

We shall work in an enriched language with special names whose reference will change depending on the content of G. We shall denote these names by $c_1, ..., c_n$. A great deal of the work required is involved in the atomic forcing clauses. However, we don't have time to go into this detail here. See my slides "A Philosopher's Guide to Forcing" for a detailed examination of how to do this.

Let's assume that we've done this work and introduce the symbol $\Vdash_{Atom}$ to deal with the forcing of atomic formulae. We can give an example of the kind of thing this will need to capture. We take the specific example of adding a Cohen real.

EXAMPLE 134. Suppose we want to add a generic real in the form of a function $g : \omega \to 2$ (we can extract a subset of the naturals such a function). To do this, we shall use the partial order $\mathbb{P}$ consisting of all the partial functions $p : \omega \rightharpoonup 2$ with finite domains. These are just functions from some finite subsets of ω into 2. For example

$$p = \{\langle 5, 0\rangle, \langle 7, 1\rangle, \langle 257, 0\rangle\}$$

is an example of such a partial function and thus $p \in \mathbb{P}$. Now let $\dot{g}$ be a special name for our generic real, g. Then we might ask whether the following hold:

(1) $p \Vdash_{Atom} \dot{g}(5) = 0$;
(2) $p \Vdash_{Atom} \dot{g}(7) = 0$;

(3) $p \Vdash_{Atom} \dot{g}(4) = 1$;
(4) $p \Vdash_{Atom} \dot{g}(4) = 0$;

For (1), we see that $\langle 5, 0 \rangle \in p$, so p does force that $g(5) = 0$. For (2), since $\langle 7, 1 \rangle \in p$, there is no way that p could force that $g(7) = 0$. So we'd actually want p to force that $g(7) \neq 0$. For (3) & (4), p does contain any information that about what g does to 4, so p doesn't force either of these statements.

Now assuming that we have a working $\Vdash_{Atom}$ in hand, we can define the forcing relation for the rest of the formulae in our language.

DEFINITION 135. We say that $p \Vdash \varphi(c_1, ..., c_n)$ iff:

- if $\varphi := c_i \in c_j$, then $p \Vdash_{Atom} c_i \in c_j$;
- if $\varphi := c_i = c_j$, then $p \Vdash_{Atom} c_i = c_j$;
- if $\varphi := \neg\psi$, then $\forall q \leq p \ q \nVdash \psi$;
- if $\varphi := \psi \vee \chi$, then $p \Vdash \psi$ or $p \Vdash \chi$;
- if $\varphi := \exists x \psi(x)$, then there is some c such that $p \Vdash \psi(c)$.

REMARK 136. If you're familiar with modal logic, this should remind you of Kripke's semantics for intuitionist logic [Priest, 2008]. We can think of each possible world as being some $p \in \mathbb{P}$; i.e., a partial state of information about a potential generic G.

It also turns out that this definition is very simple and natural. Even when we add the atomic forcing clause using transfinite recursion, we end up with something that every respectable model agrees on (like **L**). This means that if $p \Vdash \varphi$, then M can work that out too.

It should be pretty clear that some sentences will not be decided by $\Vdash$. We already saw this with the atomic case, but we can use this to form more pressing example

EXAMPLE 137. Suppose we used the partial order $\mathbb{P}$ and $p \in \mathbb{P}$ from Example 134 then we'd have

$$p \nVdash \dot{g}(4) = 0 \vee \dot{g}(4) \neq 0$$

so excluded middle fails here.

This is both a problem an opportunity. We might reason about this situation as follows:

(1) For the forcing relation $\Vdash$ to do the work we want, it needs to be able to describe a model.
(2) But the lack of bivalence at any world $p \in \mathbb{P}$, means that no world p described a model.
(3) So instead of considering worlds $p \in \mathbb{P}$ we might try a set of worlds that fills in all the details.

So if we stay with the $\mathbb{P}$ of Example 134, we could use a total function, $f : \omega \to 2$ rather than a partial function. However, it turns out that our definition of forcing won't necessarily give us us back bivalence even if we use a total function.

This gives us the *final clue* we need for our definition of a generic set: we simply build it into the definition of generic that it must do this.

DEFINITION 138. Let $\mathbb{P} \in M$ be a partial order. We say that $G \subseteq \mathbb{P}$ is generic for M if for all $\varphi(c_1, ..., c_n)$ there is some $p \in G$ such that either:

$$M \models \text{``}p \Vdash \varphi(c_1, ..., c_n) \text{ or } p \Vdash \neg\varphi(c_1, ..., c_n)\text{''}$$

With this in hand, our generic sets G will now give bivalence which puts us in a position to prove Theorem 131.

We may also prove the existence of generic sets using a construction like a completeness or compactness proof.

THEOREM 139. *Let M be a countable transtive model of ZFC and let $\mathbb{P} \in M$ be a partial order. Then there exists a $G \subseteq \mathbb{P}$ which is generic for M.*

REMARK 140. First we note that this proof is done from the perspective of $\mathbf{V}$: the *real* universe. The basic construction works as follows:

(1) Make an enumeration $\langle \varphi_n | n \in \omega \rangle$ of all the sentences in special name lanaguage.
(2) We run a construction by recursion as follows:
    (a) At stage $n+1$ suppose we have already picked a sequence $p_0, p_1, ..., p_n$ of partial information about G. Now consider φ_{n+1} and let p_{n+1} be such that either:

$$p_{n+1} \Vdash \varphi_{n+1} \text{ or } p_{n+1} \Vdash \neg\varphi_{n+1}.$$

    (The $\Vdash$ definition's negation clause ensure that we'll be able to find such a p_{n+1}).

(3) Derive G from the sequence $\langle p_n | n \in \omega \rangle$.
(4) Derive a full description of the model $M[G]$ from the set of sentence forced by some $p \in G$.

3.2.3. Using generic sets. We are now ready to put our generic sets to use. However, our definition of a generic sets is a little clunky when it comes to finding interesting ways to add generic objects. Fortunately, we can also give a more *combinatorial* characterisation of genericity. We need a quick definition first.

DEFINITION 141. Let $\mathbb{P}$ be a partial order and let $D \subseteq \mathbb{P}$. We say that D is dense in $\mathbb{P}$ if for every $p \in \mathbb{P}$ there is some $d \in D$ such that $d \leq p$.

REMARK. So informally speaking, if $\mathbb{P}$ is a partial order we are using to construct a generic object, then D is dense in $\mathbb{P}$ if no matter which way you construct, there will always be a stage later on where you could incorporate a stage from D. Now matter what we always have an opportunity to incorporate some of D. We might say that D is *perenial.*

LEMMA 142. *Let $\mathbb{P} \in M$ be a partial order. Then the following are equivalent:*

(1) $G \subseteq \mathbb{P}$ *is generic for* M.
(2) $G \subseteq \mathbb{P}$ *is such that:*
   (a) *for all dense* $D \subseteq \mathbb{P}$, $G \cap D \neq \emptyset$;
   (b) *if* $p \in G$ *and* $p \leq q$, *then* $q \in G$ (G *is closed under weaker information);*
   (c) *if* $p \in G$ *and* $q \in G$, *then there is some* $r \in G$ *such that* $r \leq p, q$ *(we can assemble a unique* G *from the information contained in it).*

REMARK. So this new characterisation, tells us that a generic G will incorportate a stage from every perenially availble D.

We'd like to show that $G \notin M$. To do this we note that all the partial orders we shall consider satisfy the following condition.

DEFINITION 143. Let $\mathbb{P}$ be a partial order. We say that $\mathbb{P}$ is *splitting*, if for all $p \in \mathbb{P}$ there are $q_1, q_2 \in \mathbb{P}$ such that:

(1) $q_1, q_2 \leq p$; and
(2) there is no $r \in \mathbb{P}$ such that $r \leq q_1, q_2$.

REMARK. Intuitively speaking $\mathbb{P}$ is splitting if at any stage there are further stages which necessarily lead to different generic G's. So we never get stuck on a path to constructing a particular generic G, there are always options.

THEOREM 144. *Let M be a countable transitive model of ZFC, $\mathbb{P} \in M$ be a partial order and $G \subseteq \mathbb{P}$ be generic for M. Then $G \notin M$.*

PROOF. Suppose $G \in M$. Let $D = \mathbb{P} \backslash G$. We first prove a claim.

CLAIM 145. D is dense in $\mathbb{P}$. Let $p \in \mathbb{P}$.

PROOF. If $p \notin G$, then we are done, so suppose $p \in G$. Now since $\mathbb{P}$ is splitting, there must be $q_1, q_2 \in \mathbb{P}$ such that $q_1, q_2 \leq p$ and no $r \in \mathbb{P}$ is such that $r \leq q_1, q_2$.

Suppose $q_1 \in G$ and $q_2 \in G$, then by Lemma 142, there must be some $r \in \mathbb{P}$ such that $r \leq q_1, q_2$: contradiction. Thus, at most one of q_1 and q_2 is in G. So pick one that isn't in G. Thus, we have shown that D is dense in $\mathbb{P}$ □

Since D is dense in $\mathbb{P}$, $G \cap D \neq \emptyset$. Fix some $d \in G \cap D$. But since $d \in D$, we have $d \in \mathbb{P} \backslash D$; i.e., $d \notin G$: contradiction. □

3.2.3.1. *Cohen reals.* Let us take an aribtrary countable transitive model M of ZFC and consider the partial order $\mathbb{P}$ we used in Example 134. It was the set of all finite partial functions with domain some finite subset of ω and range $2 = \{0, 1\}$.

Now let's consider a generic object $G \subseteq \mathbb{P}$ for M. We would like to be able to derive a total function $g : \omega \to 2$ from G. Each of the finite partial functions is a finite set of ordered pairs and a total function is an infinite set of ordered pairs.

So let's make $g = \bigcup G$; i.e, we'll take all of the pairs contained in any element of G. We then need to verify that

LEMMA 146. *G is a function.*

PROOF. Condition 2(c) from Lemma 142 ensures that the resulting relation is a function. It stops us from having $p, q \in G$ with $\langle n, 0 \rangle \in p$ and $\langle n, 1 \rangle \in q$ for some $n \in \omega$.

We also need to ensure that g is total; i.e., for every $n \in \omega$, there is some $i \in \{0, 1\}$ such that $g(n) = i$. To do this we exploit condition 2(a) from Lemma

142. Working from the perspective of M, for $n \in \omega$, let

$$D_n = \{p \in \mathbb{P} \mid \exists i \in \{0,1\}\ p(n) = i\}.$$

CLAIM. D_n is dense in $\mathbb{P}$.

PROOF. Let $p \in \mathbb{P}$. Suppose $\langle n, i\rangle \in p$ for some $i \in \{0,1\}$. Then $p \in D_n$ and we are done.

Suppose $\langle n, i\rangle \notin p$ for any $i \in \{0,1\}$. Then let $q = p \cup \{\langle n, 0\rangle\}$. Clearl $q \in D_n$ and we are done. □

Then we see that $G \cap D_n \neq \emptyset$ for all $n \in \omega$. For all n, let $p_n \in G \cap D_n$. Then we see that

$$\bigcup_n p_n \subseteq \bigcup G = g$$

and so g must decide every n; i.e., for all n, $g(n) = i$ for some $i \in \{0,1\}$. □

So we can see that we have a function $g : \omega \to 2$ and since $G \notin M$ neither is g.

Now since M and $M[G]$ are respectable models they agree about what is in $\mathbf{L}$. Thus, we see that

$$M[G] \models G \notin \mathbf{L}.$$

Thus

$$M[G] \models \mathbf{V} \neq \mathbf{L}.$$

This fact in conjunction with Theorem 132 and the considerations of Section 3.2.1.3 then gives us that:

COROLLARY 147. $Con(ZFC) \to Con(ZFC + \mathbf{V} \neq \mathbf{L})$.

So we have learned that neither $\mathbf{V} = \mathbf{L}$ nor its negation is entailed by ZFC.

3.2.3.2. *Collapses.* We next consider a different kind of forcing that produces a somewhat counterintuitive effect.

As before, we start our work in some countable transitive model M.

Let κ be some cardinal; i.e., M thinks κ is a cardinal, although we in $\mathbf{V}$ know better.

Let $\mathbb{P} \in M$ consist of functions p with domain some finite subsets of ω and range a subsets of κ. We then let $p \leq q$ iff $p \supseteq q$.

Now consider a generic $G \subseteq \mathbb{P}$ and the function $g = \bigcup G$ dervied from it. This function has a surprising effect.

THEOREM 148. $M[G] \models$ "$\kappa < \omega_1$"*; i.e.,* $M[G]$ *thinks that* κ *is a countable ordinal and not a cardinal at all.*

PROOF. Again we are going to make use of condition 2(a) of 142. This time, for all $\alpha < \kappa$, let

$$D_\alpha = \{p \in \mathbb{P} \mid \exists n \; p(n) = \alpha\}.$$

CLAIM. D_α is dense in $\mathbb{P}$ for all $\alpha < \kappa$.

PROOF. Let $p \in \mathbb{P}$. Suppose that $\langle n, \alpha \rangle \in p$ for some n. Then we'd be done. Suppose $\langle n, \alpha \rangle \notin p$ for any $n \in \omega$. Then let $q = p \cup \{\langle n, \alpha \rangle\}$ for some $n \notin dom(p)$. This suffices. □

Now from this claim we see that $G \cap D_\alpha \neq \emptyset$ for all $\alpha < \kappa$. From this we can see that $g = \bigcup G$ will be such that for all $\alpha < \omega_1$ there is some $n \in \omega$ such that $g(n) = \alpha$; i.e., there is a surjection from ω onto κ. But this then means that

$$M[G] \models \text{"}\kappa < \omega_1.\text{"}$$

□

So what M thought was an uncountable cardinal is countable from the perspective of $M[G]$. This strange effect is know as *collapsing a cardinal.*

3.2.3.3. *Adding lots of reals.* We are now ready to sketch the proof of the main theorem regarding CH.

Again, we're going to work in a countable transitive model M and define a partial order that suits our purpose.

Our goal this time is to find a generic G which is such that

$$M[G] \models 2^\omega \geq \aleph_2;$$

i.e., that there are at least $\aleph_2$ many functions $f : \omega \to 2$. Since there are $2^{\aleph_0}$ many such functions, this will show suffice to show that CH is false in $M[G]$.

To do this we want to be able to easily derive $\aleph_2$ many new functions $g : \omega \to 2$ from G.

We design our partial order $\mathbb{P}$ as follows. Let $\mathbb{P}$ be the set of functions p with domain a finite set of $\aleph_2 \times \omega$ and range 2.

REMARK. The idea here is that we are adding $\aleph_2$ many Cohen reals by using the first place of the function as an index.

Let $G \subseteq \mathbb{P}$ be generic for M and let $g = \bigcup G$. Using a similar argument to that in Lemma 146, we see that:

PROPOSITION 149. $g : \aleph_2 \times \omega \to 2$.

Now it should then be clear that from g we may derive a sequence $\langle h_\alpha | \alpha \in \aleph_1 \rangle$ of functions such that for all $\alpha < \aleph_2$:

(1) $h_\alpha : \omega \to 2$;
(2) $h_\alpha(n) = g(\alpha, n)$ for all $n \in \omega$; and
(3) $h_\alpha \notin M$.

From this we are almost in a position to show that:

THEOREM 150. $M[G] \models$ "$2^\omega \geq \aleph_2$.

From the argument above we know that there are many new reals in $M[G]$, but we only know that there are $\aleph_2$ of them from the perspective of M. We don't, however, know whether $\aleph_2$ has been collapsed in $M[G]$.

Fortunatley, it can be shown that:

FACT 151. *For $\mathbb{P}$ defined as above and $G \subseteq \mathbb{P}$ generic for M, all cardinals of M remain cardinals in $M[G]$.*

This fact is dependent on the type of partial order we use and requires a quite technical proof. With this in hand, we could then establish Theorem 150.

And from here using Theorem 132 noting the remarks of Section 3.2.1.3 we may establish that:

COROLLARY 152. $Con(ZFC) \to Con(ZFC + \neg CH)$

CHAPTER 4

Advanced topics in foundations

This week, we are going to look at:

(1) Large objects, palpable problems & determinacy; and
(2) Multiverse theories, is there more than one mathematical universe?

This will give us some idea of more recent progress regarding CH through what is known as Gödel's programme and we'll finish by investigating some recent and perhaps pessmistic thoughts regarding the prospects of its completion.

4.1. Large objects, palpable problems & determinacy

This week we are going to look at the effect of augmenting ZFC with axioms asserting the existence of large cardinals. A large cardinal is so big that ZFC cannot prove the existence of such an object.

One of the smallest large cardinals in an *inaccessible* cardinal. Given an inaccessible cardinal κ, we can show that V_κ is actually a models of ZFC, i.e.,

$$V_\kappa \models ZFC.$$

It's so large powerful axioms like Replacement and Powerset cannot be used to break free of it. But we then see that

$$ZFC + \text{“}\exists\kappa\ \kappa \text{ is inaccessible.”} \vdash \text{“}\exists\mathcal{M}\ \mathcal{M} \models ZF\text{”}$$

and so

$$ZFC + \text{“}\exists\kappa\ \kappa \text{ is inaccessible.”} \vdash \text{“}Con(ZFC)\text{”}.$$

Thus, $ZFC \nvdash \exists\kappa\ \kappa$ is inaccessible.”

We say that $ZFC +$ “$\exists\kappa\ \kappa$ is inaccessible.” has greater consistency strength that ZFC.

We might be able to provide a non-mathematical argument for the existence of inaccessible cardinals. By adding them we get to prove that there is a natural model of ZFC. It would be very strange to take ZFC seriously and not believe that there was such a model. First, let's imagine that we thought there was no model at all for ZFC, then we'd think that ZFC was inconsistent. But if we thought that, we'd think that any sentence from $\mathcal{L}_\in$ was a theorem of ZFC. It would be trivial. So it would be pretty strange to also think ZFC was a good foundation for mathematics. So we might then think that we do have reason to think that ZFC is consistent and thus, that is also has a model. We might still doubt that is ZFC has a natural model in thee sense of being V_κ where κ is inaccessible, but a little further probing into the notion of inaccessiblity and the axioms of ZFC might assuage that doubt too.

This is known as a *reflection argument*. We can repeat it to get a variety of larger and larger cardinals. But all of the cardinals that developed, so to speak, from below are very small in comparison to those we'll be considering this week. Our goal is to illustrate how assuming the existence of staggeringly large objects can have a major effect on what we can learn about comparitively mundane objects like the real numbers.

4.1.1. Games and determinacy. Our tool for this illustration will be games of perfect information and determinacy theorems. Unlike some of the more abstract set theory we've seen over the last few weeks, this material is quite palpable.

4.1.1.1. *Infinite games of perfect information.* We first introduce a game of perfect information. Let's take the logicians' reals ω^ω; i.e., function $f : \omega \to \omega$. Now let's consider a particular set $A \subseteq \omega^\omega$. We can play a game G on this set as follows. I'll be player I and you can be player II.

- I'll start play by playing a number n_0.
- Then you as II can play a number n_1.
- We keep alternating back and forth infinitely.

We allow each player to see what the other player is doing: this is why it's known as a game of *perfect information.*

After the game has finished, we take the sequence $\langle n_0, n_1, n_2, ..., \rangle$ and observe that we can extract a function $f : \omega \to \omega$ form it such that for all $m \in \omega$,

$$f(m) = n_m.$$

With this in hand we may then determine the winner for the game as follows:

- I win (as player I) if $f \in A$; and
- You win (as player II) if $f \notin A$.

This provides a good means of modelling a large variety games. The model is, of course, idealised in that games go on for an *infinite* number of moves.

Now clearly if you or I want to win, we're going to need some kind of strategy. We can describe this mathematically, but the idea is very simple. A *strategy* is a kind of book that tells us what to do if we are at a particular position in the game and our opponent makes a certain move. In order to be comprehensive, the strategy will need to tell us what to do whatever our opponent does.

Now we might then wonder whether given a particular set $A \subseteq \omega^\omega$ whether there is a winning strategy that one of the players could take up. Clearly at most one of the players can have such a strategy, but perhaps neither of them do.

Let's consider a few simple cases.

EXAMPLE 153. (i) Suppose we let $A = \omega^\omega$; i.e., we simply let it be all the functions $f : \omega \to \omega$. Then it should be pretty clear that I has to win. So any strategy they employ will let them win.

(ii) On the other hand, suppose $A = \emptyset$. Then II will win and by any strategy.

(iii) Now suppose $A \subseteq \omega^\omega$ is countable. Let $\langle f_n \mid n \in \omega\rangle$ enumerate the elements of A. Then II can win this game by using Cantor's trick and diagonalising out of A. Thus suppose it is II's move and there have already been $2n + 1$ moves played in the game so far. Then II's strategy is to play

$$n_{2n+2} = f_n(2n + 2) + 1.$$

So the idea is that II ensures that the sequence $f = \langle n_0, n_1, ...\rangle$ resulting from a full play of the game cannot be equal to any f_n for $n \in \omega$. Thus $f \notin A$ and II wins.

4.1.1.2. *Determinacy of games on open and closed sets.* These examples are quite straightforward. The following is more interesting and provides the basis for what comes later. We shall now show that games on what are known as closed sets are determined.

Let $\omega^{<\omega} = \bigcup_n \omega^n$; i.e., the set of all finite sequences of natural numbers.

Let $T \subseteq \omega^{<\omega}$ be a tree if for all $p \in T$ if $q \subseteq p$, then $q \in T$; i.e., T is closed under initial segements.

Let $[T] = \{x \in \omega^\omega \mid \forall n \in \omega \exists p \in T\ p = x \upharpoonright n\}$. This is known as the *body of the tree* T. The basic idea is very simple. If a tree contains all the initial segments of some real $x \in \omega^\omega$, then $x \in [T]$.

We can now define a closed set as follows:

DEFINITION 154. $A \subseteq \omega^\omega$ is *closed* if there is some $T \subseteq \omega^{<\omega}$ such that $A = [T]$. Let us say that such a tree *represents* A.

LEMMA 155. *(Upward winning strategy preservation) Let G be a game played on $A \subseteq \omega^\omega$. Let $p \in \omega^{<\omega}$ represent the play of the game so far. Then:*

(1) *If $lh(p)$ is even (i.e., it's I's turn to move) and whatever I does leads to a situation where II has a winning strategy in the subsequent game, then II already has a winning strategy at stage p.*
(2) *If $lh(p)$ is odd (i.e., it's II's turn to move) and there is move available to II such that II has a winning strategy for the rest of the gaem, then II already has a winning strategy at stage p.*

THEOREM 156. *All games on closed sets are determined.*

PROOF. Let $A \subseteq \omega^{<\omega}$ be closed and let T represent A. Now suppose II does not have a winning strategy.

CLAIM. I has a winning strategy.

PROOF. We describe a strategy for I and show that it leads to a win as follows.

First up it's $I's$ turn to move. Lemma 155 (1) (contraposed) tells us that I has a move that will lead to a game in which II still lacks a winning strategy.

Now suppose I has made some move, so it's II's turn. Then Lemma 155 (2) tells us no matter what II does, the subsequent game is one in which II still lacks a winning strategy.

We then keep going in the same fashion to define a sequence of moves $\langle n_m \mid m \in \omega \rangle$ such that at every stage II lacks a winning strategy. We then derive a real, $f \in \omega^\omega$, from the sequence.

It suffices to show $f \in A$. Suppose not. Then there is some $m \in \omega$ such that $f \upharpoonright m \notin T$. But if that were the case, then the game ensuing from

$$f \upharpoonright m = \langle n_0, ..., n_{m-1} \rangle$$

would be one in which II had a winning strategy: we've broken out of the tree representing A, so there is no way for a play that continues from this point to be in the body of T. This, however, contradicts what we already know about f; thus, $f \in A$. □

Thus, the game is determined since either II has a winning strategy or I has a winning strategy. □

REMARK 157. Note that if II wins a game on a closed set, then there'll be some finite point in the game after which they've already won. This because all II needs to do is break out of the tree T that represents the closed set A.

4.1.1.3. *Why determinacy is interesting.* So now we've seen a few determinacy theorems, but why should we be interested in this? One important reason is that determinacy theorems show that sets of reals behave in particularly regular ways. I'll give a particularly salient example of this.

THEOREM 158. *Let $A \subseteq \omega^\omega$ be a set of reals of a particular complexity class. Then if all the games on sets of that determinacy class are determined, then either:*

(1) $|A| = 2^{\aleph_0}$ *; or*
(2) A *is countable.*

REMARK. The idea of the proof is to define a new game (of the same complexity) such that I wins if $|A| = 2^{\aleph_0}$ and II wins if A is countable.

So this tells us that if we restrict our attention to sets of reals from complexity classes satisfying determinacy theorems, then the continuum hypothesis is satisfied with respect to them; i.e., there are no infinite sets of that complexity class which have cardinality intermediate between $\aleph_0$ and $2^{\aleph_0}$.

Given that determinacy leads to such pleasing properties, we might countenancy adding an axiom that says that every game is determined. This is known as the *Axiom of Determinacy*, abbreviated AD. However, against this we have the following:

THEOREM 159. $ZFC \vdash \neg AD$.

REMARK. To prove this we need to employ the Choice to define a game that neither player has a winning strategy in. So the point here is perhaps that AC and AD are incompatible.

4.1.2. Π^1_1 **sets.** Our goal for the rest of this section will be to build up to a sketch of the proof of Π^1_1 determinacy. This proof requires more resources than ZFC and will exploit large, large cardinals.

4.1.2.1. *Well-foundedness of trees in* $\omega^{<\omega}$. Let $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$ be a tree. We say that T is *recursive* if there is a Turing machine φ_e such that for $p, q \in \omega^{<\omega}$ we have

$$\varphi_e(\langle p, q\rangle) \;\Leftrightarrow\; (p, q) \in T.$$

DEFINITION 160. Let $A \subseteq \omega^{<\omega}$. We say that A is Π^1_1 if there is some recursive tree $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$ such that for all $f \in \omega^\omega$

$$f \in A \;\Leftrightarrow\; \forall g \exists n\ (f \upharpoonright n, g \upharpoonright n) \notin T.$$

REMARK. The underlying idea here is quite intuitive. Given such an $A \subseteq \omega^\omega$ and $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$, let

$$T(f) = \{p \in \omega^{<\omega} \mid (f \upharpoonright lh(p), p) \in T\}.$$

We might think $T(f)$ as being the x-slice of T. Now we are then saying that

$$f \in A \;\Leftrightarrow\; T(f) \text{ is well-founded.}$$

Π^1_1 sets are not that exotic. For a salient philosophical example, Kripke's truth definition is a Π^1_1 set of natural numbers.
The complement of a Π^1_1-set is a Σ^1_1 set.

DEFINITION 161. Let $A \subseteq \omega^{<\omega}$. We say that A is Σ^1_1 if there is some recursive tree $T \subseteq \omega^{<\omega} \times \omega^{<\omega}$ such that

$$\begin{aligned} f \in A &\Leftrightarrow T(f) \text{ is ill-founded} \\ &\Leftrightarrow [T(f)] \neq \emptyset \\ &\Leftrightarrow \exists g \forall n\ (f \upharpoonright n, g \upharpoonright n) \in T \end{aligned}$$

We call sets that can be defined in this way via ill-foundedness *Suslin*.

4.1.2.2. Π^1_1 *sets are also Suslins.* In the previous section we say that Σ^1_1 sets can be defined using ill-foundedness and are thus Suslin. It turns out that we can show that Π^1_1 sets are Suslin.
This idea is crucial for the determinacy proof.
First we observe that:

PROPOSITION 162. *A tree* $T \subseteq \omega^{<\omega}$ *is well-founded iff there is an order preserving map*

$$\sigma : T \to \omega_1;$$

i.e., such that for all $p, q \in T$

$$\textit{if } p \subseteq q, \textit{ then } \sigma(p) \geq \sigma(q).$$

Now consider what σ is: it's a set of ordered pairs of elements $p \in T$ and ordinal $\alpha < \omega_1$. Any $p \in T$ is just a finite sequence of natural numbers. Using a bit of coding trickery, we can make a simple enumeration

$$\mathbf{s}_\cdot : \omega \cong \omega^{<\omega}$$

of all the finite sequences. Now let us consider what happens if we compose these functions:

$$\sigma \circ \mathbf{s} : \omega \cong \omega^{<\omega} \to \omega_1;$$

i.e., we have an ω-sequence of ordinals below ω_1. (This function is actually partial, unless $T = \omega^{<\omega}$, but this can be cleaned up quite easily.)

The basic idea for representing a Π^1_1 set in a Suslin (i.e., ill-founded way) is to use such an ω-sequence to witness its ill-foundedness.

We define a new tree $\tilde{T} \subseteq \omega^{<\omega} \times \omega_1^{<\omega_1}$ on the basis of T where the second component of the tree builds up an order-preserving mapping (via the coding) of the tree into ω_1.

More formally, for $p \in \omega^{<\omega}$ and $x \in \omega_1^{<\omega}$ we have

$$(p, x) \in \tilde{T} \quad \Leftrightarrow \quad \forall n < m < lh(p)$$
$$(\ (p \restriction lh(\mathbf{s}_n), \mathbf{s}_n) \subseteq (p \restriction lh(\mathbf{s}_m), \mathbf{s}_m) \in T \ \to \ x(n) \geq x(m)\ )$$

From here, we can then establish that:

THEOREM 163. *Let A, T and $\tilde{T}$ be as described above. Then for any $f \in \omega^\omega$ the following are equivalent:*

(1) $f \in A$;
(2) $T(f)$ *is well-founded.*
(3) *There is some* $x \in \omega_1^\omega$ *such that for all* $n \in \omega$, $(f \restriction n, x \restriction n) \in \tilde{T}$.
(4) $\tilde{T}(f)$ *is ill-founded.*

4.1.3. Strongly Compact Cardinals. We now introduce some large, large cardinals.

4.1.3.1. *The infinitary logic $\mathcal{L}_{\kappa,\omega}$ and its compactness theorem.* First let us recall the compactness theorem.

THEOREM 164. *Let Γ be a set of sentences which is unsatisfiable, then there is some finite $\Delta \subseteq \Gamma$ which is unsatisfiable.*

Given the completeness theorem it is a trivial corollary. It tells us that there is always a finite set of witnesses to unsatifiability of a set of sentences.

But this only works for first order logic. Suppose we move into $\mathcal{L}_{\omega_1\omega}$, the logic with countable conjunctions and disjunctions. $\mathcal{L}_{\omega_1\omega}$ is not compact.

EXAMPLE 165. Let $\{c_\alpha \mid \alpha \in \omega_1\}$ be an uncountable set of constant symbols and let $\{d_n \mid n \in \omega\}$ be a further countable set of constant symbols. Let Γ be the following set of sentences from $\mathcal{L}_{\omega_1\omega}$

$$\Gamma = \{c_\alpha \neq c_\beta \mid \alpha < \beta \in \omega_1\} \cup \{\forall x \bigvee_{n\in\omega} x = d_n\}.$$

Let $\Delta \subseteq \Gamma$ be finite. Then Δ is satisfiable. However, Γ itself is not satifiable. The first half says there are uncountably many things while the second half says there are only countably many.

This means we cannot find a finite witness to the unsatisfiability of Γ. But perhaps this is unfair. In moving to a logic that allows countable conjunctions rather than finite conjunction, perhaps we should be looking for a countable witness. But this also fails. Every countable subset of Γ above is also satisfiable.

So is compactness gone altogether? Perhaps we can move to an even stronger logic where compactness returns. We have to go a long way for this to occur. Let's voice our question as follows:

Question: Is there some κ such that for any set Γ of sentences from $\mathcal{L}_{\kappa\omega}$ if Γ is unsatisfiable, there is some $\Delta \subseteq \Gamma$ with $|\Delta| < \kappa$ and Δ inconsistent.

DEFINITION 166. A cardinal κ satifying the condition above is known as a *strongly compact cardinal.*

It turns out that ZFC cannot prove the existence of such a cardinal. However, if they do exist then this has some interesting consequences.

FACT 167. *There are unboundedly many inaccessible cardinals below a strongly compact cardinal.*

4.1.3.2. *Indiscernibles.*

DEFINITION 168. A set $C \subseteq \mathbf{On}$ is called a set of *(order) indiscernibles* if for $\varphi(v_1, ..., v_n)$ and all $\alpha_1 < ... < \alpha_n, \beta_1 < ... < \beta_n$ from C if

$$\varphi(\alpha_1, ..., \alpha_n) \;\leftrightarrow\; \varphi(\beta_1, ..., \beta_n).$$

The following is not optimal.

FACT 169. *If there is a strongly compact cardinal κ, then there a set C such that:*

(1) *C is a set of indisceribles for L;*
(2) *$|C| = \omega_1$; and*
(3) *$C \subseteq \omega_1$.*

4.1.4. Π^1_1 determinacy. With these tools in hand, we can now prove that every game on a Π^1_1 set is determined.

4.1.4.1. *Auxilliary games.* The general strategy for proving determinacy theorems for more complext sets is quite interesting. It can be described as follows:

> Translate the game into a *more complex* space in which the game becomes *more simple*.

In general, we describe an auxilliary game (related to the first game) which is closed. Then by Theorem 156, we see that this auxilliary game is determined. Using the relationship between the games, we then show that having a winning strategy for some player is preserved back into the original game.

In the case of Π^1_1 sets, we are going to exploit the Suslin property we described in Theorem 163. Let $A \subseteq \omega^\omega$ be Π^1_1 and let T be the recursive tree representing it. We are going to let player I play (in addition to their natural number) and ordinal $\alpha < \omega_1$ at each move.

I	(n_0, α_0)		(n_2, α_1)		$\dots$
II		n_1		n_2	

The idea here is that a play of the game will result in a function $x : \omega \to \omega_1$ which will verify the well foundedness of the tree $T(f)$ that is played out by the natural number moves of players I and II.

If player, I succeeds in doing this, they will have verified that $f \notin A$. On the other hand, if this is not possible, then this will have been because $T(f)$ is not well-founded and thus contains an infinite branch, g. Then any attempt to assign ordinals to the elements of this branch will come undone as there are no infinite descending chains of ordinals. Moreover, this will necessarily be revealed after a finite amount of time. Using Remark 157, this tells us that the game is actually based on a closed set, albeit in a much more complex space. Thus the auxilliary game is determined by Theorem 156.

4.1.4.2. *Sketch of the proof.*

THEOREM 170. *Let* $A \subseteq \omega^\omega$ *be* Π^1_1 *then a game on* A *is determined.*

PROOF. (Sketch) Since the auxilliary game is closed and thus determined. Moreover, since the game is very simply, we can actually play it inside L and since $L \models ZFC$ the game is also determined there. Let's play the game in L. It then suffices to show that if I or II has a winning strategy in the auxilliary game, then either I or II respectively has a winning strategy in the original game.

For player I this is simple. We just forget the extra ordinal moves. Then if f results from a play for I, it can be seen that $f \in A$.

We merely sketch the argument for player II. We suppose that II has a winning strategy in the auxilliary game and define a strategy for II in the original game.

(1) Let σ^* be a a winning strategy for II in the auxilliary game.
(2) We define σ using σ^* by considering all the ways I could have played, but resricting $I's$ ordinal moves to $\alpha \in C$.
(3) Since C is a set of indisernibles and this definition can be expressed in $\mathcal{L}_\in$, we'll get a unique move out of this.
(4) Call the resultant strategy σ.

We now claim σ is winning for II in the original game.

(1) Suppose not, and let $f \in A$ be a play where II followed σ.
(2) Then since $\tilde{T}(f)$ is ill-founded, we see that there is an order preserving $x : \omega \to \omega_1$.
(3) Since $|C| = \omega_1$, there is an order preserving $y : \omega \to C$.
(4) But this then corresponds to a run of the auxilliary game where II plays according to σ^* and yet loses, contradicting our initial hypothesis.

□

4.2. Multiverse theories, is there more than one mathematical universe?

- Large cardinals
  - Success with analysis
  - Failure with CH

4.3. Hamkins

- Argues that ideal approach to solving CH cannot be realised.
- Essentially, set theorists and mathematicians are to familiar working in worlds where CH is true and CH is false for any principle implying either CH or its negation to be taken seriously.
- A naturalistic approach to set theory
  - Set theorists are mostly concerned with different models of set theory.
  - These different models are part of the furniture of contemporary set theory, so they should also be part of the ontology - we should take them at face value.
  - We end up with the set theoretic multiverse.
- Multiverse set theory
  - No new - bifurcation has been taken seriously since the emergence of forcing arguments.
  - This is close to Cohen's ultimate view, but this is regarded as a species of formalism - we can write this stuff down and jumble the symbols about but it's doubtful that we're doing anything serious.
  - Hamkins view is not formalist. He is a Platonist about the multiverse. Each of the alternative worlds is out there and (presumably) within the range of our quantifiers.
- Hamkins principles
  - Forcing, well-foundedness, countability.
  - It has a (very strange) model.
    * Is this too radical?

4.4. Steel

- Re-invisaging Gödel's programme

"Gödel's Program: Decide mathematically interesting questions independent of ZFC in well-justified extensions of ZFC."

- maximising interpretative power
- The large cardinal programme has been very successful at this.
- But some questions appear to be untouched by it. Since

$$Con(ZFC) \leftrightarrow Con(ZFC + CH) \leftrightarrow Con(ZFC + \neg CH)$$

- No interpretability strength is gained by adding CH or its negation to set theory.
- This is revealed by forcing arguments.
  "None of our current large cardinal axioms decide CH, because they are preserved by small forcing, whilst CH can be made both true and false by small forcing. Because CH is provably not generically absolute, it cannot be decided by large cardinal hypotheses that are themselves generically absolute"

- The generic multiverse.
  - Example: closure under forcing.
  - Supervaluate for the meaningful sentences.
  - No higher order (multi-multi-verse) is required - every world in the generic agrees on super-true (this is a recent revelation).
- Questions
  - **Weak relativist thesis (Steel):** Every (*meaningful*) proposition that can be expressed in the standard language LST can be expressed in the multiverse language. (All mathematics done to date can be expressed in the multiverse language, and proved in extensions of the multiverse theory by large cardinals).
  - **Strong relativist thesis:** (No individual world is definable in the multiverse language. In particular, there is no inclusion minimal world.)
  - **Strong absolutist thesis:** "$\dot{V}$" makes sense, and that sense is not expressible in the multiverse language. ("V" makes sense, regardless of whether we can define it in the multiverse language.)

  - **Weak absolutist thesis (Steel hopes):** There are individual worlds that are definable in the multiverse language; that is, the multiverse has a core.
- Is this still set theory?
  - What is set theory for? Are we trying to give the definitive theory of collections or as close as we can? or are we trying to provide a foundation for mathematics? Are these goals compatible?

4.5. Woodin

- The generic multiverse could be flawed, so perhaps CH does have an answer.
- Ω-logic and the generic multiverse
  - $\models_\Omega$
  - $\vdash_\Omega$
  - The Ω-conjecture - a completeness theorem for Ω-logic.
- Assuming Ω-logic is complete, then the multiverse account may have a problem.
  - If the Ω conjecture is true, then there is an initial segment of every universe in which the proof theory for Ω-logic may be carried out.
  - The proof theory for ordinary first order logic is carried out in H_ω. The proof theory for a Kripke fixed point is carried out in $H_{\omega_1^{CK}}$. The proof theory for $\mathcal{L}_{\omega_1\omega}$ is carried out in H_{ω_1}.
  - Woodin claims that set theory has a distinctive feature. In other areas of science, like say physics, we would welcome a reduction of our ontology and theoretical resources.
  - Set theory is not like this. If we could figure out all of the meaningful truth of sets using an initial segement of the universe, then there is a sense in which we've betrayed the underlying motivation for set theory. The stuff beyond the initial segment in question isn't really adding anything meaningful.
  - There is a Hilbertian formalist shadow looming here. Recall that Hilbert thought that number theory was where substantive mathematics takes place. His idea was that we could talk of real numbers and exotic sets, but these were mere constructions which could ultimately be reducible back to the real work

of number theory. To vindicate our use of such exotic theories, he hoped that we might prove their consistency using number theory: a hope which Gödel ruined.

– Assuming the Ω-conjecture, the sets beyond the initial segment are, in some sense, merely convenient fictions which we can reduce back down to the "real" sets. So if Woodin's conjecture is true, then we could not do this at the same time as holding some kind of Gödelian Platonist committment to the existence of (large) sets.
– However, it is still just a conjecture. Moreover, it is not clear to me that this kind of formalism is necessarily a problem for Gödel's programme as Steel envisages it. While Platonism is gone, with respect to large sets, I'm not sure why that should worry us in the goal to increase the intepretative strength of our theories.

Bibliography

George Boolos. To be is to be the value of a variable (or to be some values of some variables). *The Journal of Philosophy*, 81:430–450, 1984.

Paul J. Cohen. *Set Theory and the Continuum Hypothesis*. W. A. Benjamin Inc., New York, 1966.

Keith J. Devlin. *Constructibility*. Springer-Verlag, Berlin, 1984.

Kenneth Kunen. *Set Theory: an introduction to independence proofs*. Elsevier, Sydney, 2006.

M.D. Potter. *Set Theory and Its Philosophy: A Critical Introduction*. Oxford University Press, 2004.

Graham Priest. *An Introduction to Non-Classical Logic: From If to Is*. Cambridge University Press, Melbourne, 2008.

J. C. Shepherdson. Inner models for set theory - part II. *Journal of Symbolic Logic*, 17(4):225–237, 1952.