Maths 336 at Canterbury, Winter 2016

Thomas Forster

September 26, 2016

Office hours by negotiation. My office is Erskine 616 but i tend to work from home: 21a Creyke rd. My mob is 071-771-1729.

1 Some Set Theory

We give a historically motivated introduction.

Points at infinity are concretised as pencils of lines; imaginary divisors concretised as ideals (= sets). Here is the standard example: $\mathbb{Z}[\sqrt{-5}]$ is sold to us as the substructure of $\mathbb{C}$ generated by $\mathbb{Z}$ and $\sqrt{-5}$.

In $\mathbb{Z}[\sqrt{-5}]$ we can factorise 6 as $2 \cdot 3$ and also as $(1+\sqrt{-5}) \cdot (1-\sqrt{-5})$ (we can compute these products in $\mathbb{C}$) and all these four factors are irreducible in $\mathbb{Z}[\sqrt{-5}]$.

So we invent "lower" factors—four of them in fact. One to be a common factor of 2 and $1+\sqrt{-5}$, a second to be a common factor of 2 and $1-\sqrt{-5}$, the third to be a common factor of 3 and $1+\sqrt{-5}$, and finally the fourth to be a common factor of 3 and $1-\sqrt{-5}$. How are we to concretise these ficticious factors? The key observation is that, although we (think) we do not know what these new roots are, we know exactly what their nontrivial multiples are, and that gives us a way in.

Different ideal divisors will correspond to different sets, so we concretise the ideal divisor of 3 and $1+\sqrt{-5}$ *as that set*: $\{a \cdot 3 + b \cdot (1+\sqrt{-5}) : a, b \in \mathbb{Z}\}$.

At a later stage set theory was used to provide new concretisations of things—such as natural numbers—that either already had satisfactory concretisations, or were never felt in need of concretisations in the first place. (There had been a problem with complex numbers but set theory wasn't part of the solution)

Russell-Whitehead and Frege concretise some things as equivalence classes: e.g. cardinals and ordinals.

In the same operationalist spirit we concretise functions as functions-in-extension.

Sets are pluralities. Lots of different kinds of pluralities. A flock of sheep; a job lot of sheep in a stock auction. Sets are the metaphysically minimal plurality. Algebras are pluralities with inner structure. A job lot has structure only from an external point of view.

OK, so the set of multiples of the ideal divisor exists as a comprehended object, some suitably concrete object-in-extension. Ditto the pencil-of-lines. So there is an unproblematic object-in-extension corresponding to the two intension (ideal divisor, point at infinity) Does this always work? Does every set-in-intension have a corresponding set-in-extension? No! Russell was able to show this, using very old ideas going back at least to the Greeks. Russell's paradox. It's an interesting object proof-theoretically but for us it's just a pain. We are going to have to come up with some subset of the set of axioms of naïve set theory plus a good story about why we use that subset rather than any other.

There are various subsets one can use, but—altho' I am an expert on one particular one, due to Quine—i am not going to tell you about that subset, but talk only about the mainstream version which everyone uses. It's known as **Zermelo-Fraenkel Set theory** or 'ZF' for short. A guiding principle in trying to suss out the most suitable subset to use is the recurring thought that set theory started off (as outlined above) as a way of concretising abstract mathematical objects. But first we deal with the most fundamental axiom: extensionality. $(\forall x, y)(x = y \longleftrightarrow (\forall z)(z \in x \longleftrightarrow z \in y))$.

We concretise functions as sets of ordered pairs so let's concretise ordered pairs. We want a total [binary] function `pair` and two [unary] partial functions `fst` and `snd` (or π_1 and π_2 if you prefer) s.t.
$(\forall xy)(\texttt{fst}(\texttt{pair}(x, y)) = x)$ and
$(\forall xy)(\texttt{snd}(\texttt{pair}(x, y)) = y)$

One that works is $\texttt{pair}(x, y) = \{\{x\}, \{x, y\}\}$. This is the **Wiener-Kuratowski** pair. Then
$\texttt{fst}(p) = \bigcap\bigcap p$ and
$\texttt{snd}(p)$ = the unique member of $\bigcup p$ that belongs to exactly one member of p. (Thanks to Peter Johnstone for supplying this constructive formulation of `snd`.)

$$x = \texttt{snd}(p) \longleftrightarrow (\exists! z)(z \in p \wedge x \in z).$$

If ordered pairs are concretised as above, what axioms do we need if we are to construct and deconstruct them?

Pairing, sumset and a certain amount of separation.

State these axioms here

What other axioms. . . ? Well, it shouldn't matter how we concretise ordered pairs. Let's try to prove the existence of $X \times Y$ (which is a set, after all, even if ordered pairs aren't!) without knowing what an ordered pair is.

For any $x \in X$ we consider the function $f_x : y \mapsto \langle x, y\rangle$. The range of $f_x \restriction Y$ is just $\{x\} \times Y$. Consider now the function $F_x : x \mapsto \{x\} \times Y$. The range of $F_x \restriction X$ is $\{\{x\} \times Y : x \in X\}$ and $\bigcup$ of this is just $X \times Y$.

Notice that we have not made any assumptions about what particular set $\langle x, y\rangle$ might be for $x \in X$ and $y \in Y$. However we have assumed (twice) that *the image of a set in a function is a set*. This assumption is the **axiom scheme of replacement**. f"x notation: $f\text{"}x = \{f(y) : y \in x\}$ is the image of x in f.

[There is actually a converse to this: if $X \times Y$ always exists however you implement pairing and unpairing then the axiom scheme of replacement follows.]

If you have decided that ordered pairs are W-K pairs then you can prove the existence of $x \times y$ without using replacement. It's a useful basic exercise. You need pairing, power set and a bit of separation.

Power set, sumset, infinity, foundation, separation. Talk over them.

Foundation says that $\in$ is a wellfounded relation. That means you can prove facts about sets by $\in$-induction. For example, you can show that no set is a member of itself. Try it.

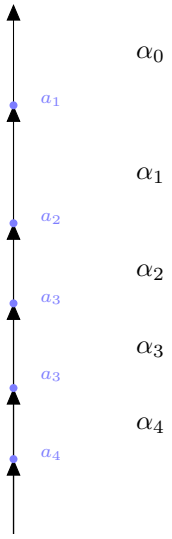
Prove that given any two worders there is a unique isomorphism between one and an initial segment of the other. Then prove carefully that $<_{On}$ wellfounded.

THEOREM 1

1. *Every wellordering is rigid (no nonidentity automorphisms)*
2. *If there is an isomorphism between two wellordering $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ then it is unique;*
3. *Given two wellorderings $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ one is isomorphic to a unique initial segment of the other.*

Proof:

1. Any automorphism of a total order is torsion-free—every cycle looks like $\mathbb{Z}$ and can have no least element.
2. Suppose σ and τ were two distinct isomorphisms $\langle A, <_A\rangle \to \langle B, <_B\rangle$; Then $\sigma \cdot \tau^{-1}$ would be a nontrivial automorphism of $\langle A, <_A\rangle$.
3. We define an isomorphism by recursion in the obvious way. It must exhaust either $\langle A, <_A\rangle$ or $\langle B, <_B\rangle$ and, by the earlier parts, it will be unique.



THEOREM 2 *$<_{On}$ is wellfounded.*

Proof: Let α be an ordinal. We will show that the ordinals below α are wellfounded. The long arrow represents a wellordering $\langle A, <_A\rangle$ of length $\alpha = \alpha_0$. If (*per impossibile*) there is a family $\{\alpha_i : i \in I\}$ of ordinals with no least member (and all of them $< \alpha$) then, for each $i \in I$, $\langle A, <_A\rangle$ has a (unique) proper initial segment of length α_i. For $i \in I$ let a_i be the supremum of that (unique) initial segment of $\langle A, <_A\rangle$ of length α_i. Then $\{a_i : i \in I\}$ is a subset of A with no $<_A$-least member. ■

Define the **Cumulative Hierarchy** by recursion on the ordinals.

$$V_\alpha =: \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta).$$

(Needs replacement and power set). Explain carefully why this needs replacement. The restriction of $\in$ to the cumulative hierarchy is wellfounded beco's ρ is a homomorphism $\twoheadrightarrow$ the ordinals.

If we have foundation we can use $\in$-recursion to define a **rank** function, written 'ρ'. $\rho(x) = \sup\{\rho(y) + 1 : y \in x\}$. The rank of a set is precisely its *birthday* of which i spoke earlier. Thus $\rho(x)$ is the least ordinal α such that $x \subseteq V_\alpha$.

Talk briefly about V_ω. Explain why it is countable. Observe that V_ω is a model for all the other axioms of set theory, in the sense that if you restrict the variable in the other axioms to range only over [members of] V_ω then they come out true. However everything in V_ω is finite. This is unsatisfactory, since if we are to think of natural numbers as sets then we will need infinitely many of them, and the set that is to do duty as $\mathbb{N}$ will have to be infinite. Accordingly

we need an **axiom of infinity** that says that there is an infinite set. Precisely what form this axiom will take we will determine later.

We cannot straightforwardly concretise cardinals as equivalence classes if we have separation beco's $\bigcup \alpha = V$ whenever α is an equipollence class, and separation will give us Russell's paradox.

Remark 1 *If α is an equipollence class (other than $\{\emptyset\}$) then $\bigcup \alpha = V$.*

Proof:

Suppose not. Then there is b s.t $(\forall A \in \alpha)(b \notin A)$. Let A be a member of α (any will do). For any $a \in A$, the set $(A \cup \{b\}) \setminus \{a\}$ is in bijection with A and is therefore in α. But then $b \in \bigcup \alpha$ after all.

■

Corollary 1 *The only equipollence class that is a set is $\{\emptyset\}$.*

To prove an analogous result for ordinals we need to know how to think of structures (decorated sets) in set theory. (An ordinal—unlike a cardinal—is *prima facie* an equivalence class of structures rather than of sets so we have to be careful.) As tuples. Then you can prove that isomorphism classes of wellorderings cannot be sets. It's the same proof (morally) but the details are fiddly.

Next we need

Definition 1 Scott's trick

When you are trying to concretise/implement a mathematical entity that arises naturally from equivalence classes for an equivalence relation, then instead of the $\sim$-equivalence class of a set x, use the collection of things $\sim$ to x that are of minimal rank with that property.

Thus, if $\sim$ is an equivalence relation we instantiate the (as it might be, cardinal) not as the true equivalence class—which may be a proper class—but instead as $[x]_\sim \cap V_\alpha$ where α is the least ordinal α s.t. $[x]_\sim \cap V_\alpha$ is nonempty. Observe that x might not be a member of its (as-it-might-be) cardinal thus construed!

How are we to concretise/implement ordinals? If we have foundation we can use Scott's trick, but there are also other tricks up our sleeves. Every equivalence class (= abstract ordinal) contains a wellordering whose order relation is set membership. We prove this using ...

1.1 Mostowski Collapse

Suppose R is a wellfounded relation on a set X.

We declare the recursion

$$\pi(x) =: \{\pi(y) : R(y, x)\}$$

which has a unique solution if R is wellfounded.

Mostowski collapse shows that every wellfounded structure $\langle X, R\rangle$ has a homomorphism π onto a structure $\langle \pi\text{“}X, \in\rangle$ where $\pi\text{“}X$ is a transitive set. A class M is *transitive* if $(\forall x, y)(((x \in y) \wedge (y \in M)) \rightarrow (x \in M))$ holds. [Explain *homomorphism*]

In general there is no reason to expect that the homomorphism π is injective. It's simple to give illustrations where it is and also illustrations where it isn't. If $\{y : R(y, x_1)\} = \{y : R(y, x_2)\}$ then clearly $\pi(x_1) = \pi(x_2)$. Clearly if there is no such pair x_1 and x_2 then π will be injective. In these circumstances we say that R is **extensional**. Reflect that the axiom of extensionality says that $\in$ is extensional.

What happens in the cases where $\langle X, R\rangle$ is a wellordering? Wellorders are total orders so distinct things have distinct predecessors so the homomorphism is an isomorphism.

Thus every wellordering is isomorphic to a wellordering whose order relation is $\in$! And this wellordering is of course unique. We then take this canonical representative to be our ordinal. This trick is due to von Neumann, and the resulting concretisations are called *von Neumann ordinals* or (often!) just plain 'ordinals' [which is naughty].

[Notice that although we have shown that every wellordering is isomorphic to a special one (which we can use as its ordinal) namely the wellordering whose order relation is $\in$, there doesn't seem to be a similar move available for cardinals. Given a set x is there an obvious special set in bijection with x, something that we can use as its cardinal? Not clear. We will return to this later.]

Once we've implemented ordinals we can implement integers, rationals, reals and complexes.

Can do these in lots of ways:

> Naturals can be von Neumann naturals or Zermelo naturals or Scott's trick naturals.
>
> Integers can be signed naturals or equivalence classes of ordered pairs of naturals
>
> Rationals can be signed ordered pairs of naturals or equivalence classes of ordered pairs of integers.
>
> Reals can be Dedekind cuts in rationals or equivalence classes of Cauchy sequences of rationals.
>
> Complex numbers typically are thought of as ordered pairs of reals.

It would be a very helpful exercise to crunch out the ranks of the sets that implement these various mathematical objects under the assorted possible implementations

The answers themselves do not matter in the slightest—the ordinals obtained are properties of the implementing sets, not of the mathematical entities

themselves—but the exercise will give you experience in manipulating some purely set theoretic quantities, and prepare you for doing some more idiomatic set theory in the days to come—something you will not have done before.

The justification I gave of R-induction on the assumption that R is well-founded was an informal one. Now that we are doing set theory formally the time has come to formally deduce $\in$-induction from the assumption that $\in$ is wellfounded.

Suppose $(\forall x)(((\forall y \in x)(F(y)) \to F(x)) \to F(x))$. Suppose (with a view to obtaining a contradiction) that $\neg F(a)$ for some a. Naturally we want a to give rise to a set with no $\in$-minimal element, thereby contradicting wellfoundedness of $\in$. The obvious candidate is the set of things in $a \cup \bigcup a \cup \bigcup^2 a \ldots$ that are not F. So the challenge is to show that this collection $a \cup \bigcup a \cup \bigcup^2 a \ldots$ is a set. We have a special notation for $\bigcup_{n\in\mathbb{N}} \bigcup^n a$; we call it $TC(a)$ for **transitive closure** of a. How are we going to prove that this is a set? If we are to do it with the axioms we have seen so far we are clearly going to have to use replacement. (Use the function $n \mapsto \bigcup^n a$ and take the image of $\mathbb{N}$ in it; then do $\bigcup$ to the result.) If we wish to be thrifty we don't use full replacement but just an axiom to say that $TC(x)$ exists for all x.

Set theory without replacement is roughly the theory of $V_{\omega+\omega}$.

Cantor's theorem. Alephs. GCH: Set theory does actually have a life of its own, beyond being a device for concretising novel mathematical entities. A good example of a genuinely set-theoretical question is the **Generalised Continuum Hypothesis**. The axiom of choice has an elementarily set-theoretic formulation, but its meaning lies outside set theory so it doesn't really count.

An aleph is the cardinal of a wellordered set. Clearly anything in bijection with a set that can be wellordered can also be wellordered, so we are right to think of this as a property of cardinals not merely as sets. [equipollence is a congruence relation for the predicate "can be wellordered"]

Theorem 3 *Vital, central fact! (Cantor)*

Every ordinal is the order type of the set of ordinals below it in their natural order.

Equivalently, the order type of an initial segment of the ordinals is the least ordinal not in it.

You prove this by induction.

Corollary 2 *(The Burali-Forti Paradox)*

The collection On of all ordinals cannot be a set.

Proof:

By thm 2 $\langle On, <_{on}\rangle$ is a wellordering. Since it is downward-closed, thm 3 tells us that its order type must be the least ordinal not in it. But there is no such ordinal. ■

ω is a *countable ordinal*. This means two things (i) it has countably many predecessors; (ii) it is the order type of a wellordering of a countable set. By theorem 3 (i) and (ii) are equivalent. Observe that $\omega + 1$, ω^2 and lots of other ordinals are also countable. Are *all* ordinals perhaps countable ... ? No!

THEOREM 4 *Hartogs' theorem.*
For every set X there is a wellordered set Y s.t. $Y \not\hookrightarrow X$.

Proof:

We exhibit a uniform construction of such a Y.

Consider $\mathcal{P}(X \times X)$. This is the set of all binary relations on X. (We have seen how $X \times X$ is a set, and by the power set axiom its power set wil be a set too.) We define a map $f : \mathcal{P}(X \times X) \to On$. If $R \in \mathcal{P}(X \times X)$ is a wellordering we send it to its order type, its length; if it is not a wellordering we send it to 0. The range f"$(\mathcal{P}(X \times X))$ of f is the set Y that we want.

What is the cardinality of Y? Y is naturally wellordered, so what is its order-type in this ordering? Y is *downward-closed* so, by theorem 3 its order-type is the least ordinal not in Y. The ordinals in Y are precisely the ordinals of wellorderings of subsets of X. So the order type of Y is the least ordinal not the length of a wellordering of a subset of X. So Y is not the same size as any subset of X. *It's too big.*

■

It's natural to ask what happens if we do the construction of theorem 4 to $\mathbb{N}$. The answer is that we get the set of countable ordinals, a set that Cantor called the *second number class*. We need a name for the cardinal of this set: $\aleph_1$. The supremum of the second number class is the ordinal ω_1, the least uncountable ordinal.

Generally if X (in theorem 4) is of size $\aleph_\alpha$ then the Y we obtain will be of size $\aleph_{\alpha+1}$.

This notation is legitimate because, if X is wellorderable, the Y that we obtain from the construction in the proof of theorem 4 is of minimal size $\not\leq |X|$. So, if $|X|$ is the αth $\aleph$, $|Y|$ is the $\alpha + 1$th aleph. Is this OK? Yes: the alephs are wellordered by $<_{card}$ so we can enumerate them using ordinals. I'm leaving this as an exercise. You will need theorem 2

Cofinality; initial ordinals

Initial ordinals;

For the moment write 'card(α)' for $|\{\beta : \beta <_{On} \alpha\}|$. (This 'card' notation is in the literature, but it is not in common use, and you do not need to know it). Then an ordinal α is **initial** if $(\forall \beta <_{On} \alpha)(\text{card}(\beta) <_{card} \text{card}(\alpha))$.

We enumerate the initial ordinals as ω_0, ω_1, ... ω_α ..., and we define $\aleph_\alpha$ to be card(ω_α) which of course was $|\{\beta : \beta <_{On} \omega_\alpha\}|$.

$\aleph_\alpha$ is of course also the αth aleph. (I am going to leave this to you to verify)

1.2 Independence of the axioms

We have seen how V_ω is a model for all the axioms except infinity. Let's be a bit more rigorous about this, and see what other results we have.

Relativisation.

Restricting to M:

> Replace every quantifier '$(\forall x)\phi(x))$' by '$(\forall x)(x \in M \to \phi(x))$' (usually written '$(\forall x \in M)\phi(x))$'
>
> Replace every quantifier '$(\exists x)\phi(x))$' by '$(\forall x)(x \in M \wedge \phi(x))$' (usually written '$(\exists x \in M)\phi(x))$'.

The result of doing this [to all quantifiers in] an expression Φ is notated 'Φ^M.

Notice that, for example, if Φ is an axiom other than the axiom of infinity then Φ^{V_ω} holds—is a theorem of ZF.

We need to talk about the difference between being closed under an operation and being a model for the axiom that says you are closed under that operation.

An old example sheet question

"Show that if M is a transitive class, then the structure $\langle M, \in\rangle$ satisfies the axiom of extensionality, and that it satisfies each of the empty-set, pair-set and union-set axioms if and only if M is closed under the corresponding finitary operation on V. What more do you need to know about M to get a similar result for the power-set axiom?"

Discussion

"finitary operation"??

Most of the axioms of set theory assert that the world of sets is closed under some operation or other. You might think that for a set to be a model of the axiom that says the world of sets is closed under operation blah it is necessary and sufficient for that set to be closed under operation blah. But you'd be wrong! In what follows $\mathfrak{M}$ is the structure consisting of the set M equipped with the membership relation $\in$.

$\mathfrak{M} \models$ the axiom of pairing iff

$$(\forall x \in \mathfrak{M})(\forall y \in \mathfrak{M})(\exists z \in \mathfrak{M})(\forall w \in \mathfrak{M})(w \in z \longleftrightarrow w \in x \vee w \in y)$$

$\mathfrak{M}$ is closed under the pair set operation iff $(\forall x, y \in \mathfrak{M})(\{x, y\} \in \mathfrak{M})$.

Are these two equivalent? Clearly yes. So far so good.

In contrast $\mathfrak{M} \models$ the axiom of power set iff

$$(\forall x \in \mathfrak{M})(\exists y \in \mathfrak{M})(\forall z \in \mathfrak{M})(z \in y \longleftrightarrow (\forall w \in \mathfrak{M})(w \in z \to w \in x))$$

Now, since $\mathfrak{M}$ is transitive, the last bit—$(\forall w \in \mathfrak{M})(w \in z \to w \in x)$—is equivalent to $z \subseteq x$, so the displayed formula simplifies slightly to

$$(\forall x \in \mathfrak{M})(\exists y \in \mathfrak{M})(\forall z \in \mathfrak{M})(z \in y \longleftrightarrow z \subseteq x)$$

$\mathfrak{M}$ is closed under the power set operation iff

$$(\forall x \in \mathfrak{M})(\mathcal{P}(x) \in \mathfrak{M})$$

Are these two equivalent? Clearly not. [Reflect that, by Skolemheim, ZF has a countable transitive model $\mathfrak{M}$. In a countable transitive model every set must be countable. So the thing in $\mathfrak{M}$ that $\mathfrak{M}$ believes to be power set of $\mathbb{N}$ will be a countable set and cannot possibly be the true power set of the naturals.]

The way to understand this stuff is to grasp the concept of **restricted quantifier** and Δ_0-**formula**, and to understand why we restricted our attention to transitive models in the first place. The idea is that if i give you a set x i must also give you all its members—since a set, after all, is nothing more than the set of all its members. So any sensible model with an element x must contain everything in the transitive closure of x as well. Hence our restriction to transitive models only.

We now want to check what conditions we have to put on $\phi(x, y)$ if we are to be confident that the truth value of $\phi(x, y)$ does not depend on the model in which we evaluate it. We say of such ϕ that they are **absolute**. The illustrations in the example sheet question are entirely fit for our purpose:

It is very important in the later development of set theory that practically everything to do with von Neumann ordinals is absolute

(i) $x = \{y, z\}$ is just

$$y \in x \ \wedge \ z \in x \ \wedge \ (\forall w \in x)(w = y \vee w = z)$$

and we observe that this can be checked just by looking inside x.

(ii) $x = \bigcup y$ is

$$(\forall w \in x)(\exists z \in y)(w \in y) \wedge (\forall w \in y)(\forall z \in w)(z \in x)$$

(iii) In contrast $x = \mathcal{P}(y)$ is

$$(\forall w \in x)(\forall z \in w)(z \in y) \wedge (\forall w)((\forall u)(u \in w \to u \in y) \to y \in x)$$

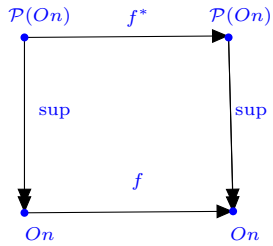
Observe the difference in the quantifiers. As we have seen, a *restricted* quantifier is one in the style "$(\forall x \in y)\ldots$" or "$(\exists x \in y)\ldots$". It turns out that $\phi(x, y)$ is absolute [between transitive models] iff all quantifiers within it are restricted. Observe that "$x = \mathcal{P}(y)$" has an unrestricted quantifier in it. We say that a formula is Δ_0 iff all the quantifiers within are restricted. "$x = \mathcal{P}(y)$" is not Δ_0 and is not absolute. If we want power sets to be preserved we have to ensure that the models we deal with not only have all *members* of all of their inhabitants, but also all *subsets* of their inhabitants.

2 Some Infinite Cardinal Arithmetic

Definition 2 *Normal functions*

A total function $f : On \to On$ is **normal** *if it is total, strictly increasing and continuous.*

"Continuous"? It means that the following diagram commutes.



"f^*" is a nonce notation for the function $X \mapsto f``X$.

Addition, multiplication and exponentiation on the Right are normal. Not on the Left! $\alpha \mapsto \alpha^2$ not normal.

Lemma 1 *Division Algorithm for Normal Functions.*

If $f : On \to On$ is normal, and α is any ordinal, then there is β such that $f(\beta) \leq \alpha < f(\beta + 1)$.

Proof:

The β we want is $\sup\{\beta : f(\beta) \leq \alpha\}$. What is $f(\beta)$? By normality it must be $\sup\{f(\beta) : f(\beta) \leq \alpha\}$, which is clearly $\leq \alpha$. So β is not merely the *supremum* of $\{\beta : f(\beta) \leq \alpha\}$, it is actually the *largest element* of $\{\beta : f(\beta) \leq \alpha\}$. But then $f(\beta + 1)$ must be strictly greater than α. ■

We need ordinal subtraction for Cantor Normal Forms.

Uniqueness of ordinal subtraction. What might we mean by '$\alpha - \beta$'? If $\beta < \alpha$ then whenever $\langle B, <_B\rangle$ belongs to β and $\langle A, <_A\rangle$ belongs to α then there is an isomorphism $\pi : \langle B, <_B\rangle$ to a unique initial segment of $\langle A, <_A\rangle$. The truncation $\langle A \setminus \pi``B,\ <_A{\restriction}(A \setminus \pi``B)\rangle$ is our wellordering of length $\alpha - \beta$. This definition ensures that $\beta + (\alpha - \beta) = \alpha$.

Cantor Normal Forms [lectured but not written up]

Theorem 5 *Every normal function $On \to On$ has a fixed point.*

Proof:

Let $f : On \to On$ be normal, and let α be any ordinal. Then $\{f^n(\alpha) : n < \omega\}$ exists, by replacement. Then, by normality, we have

$$\sup\{f^n(\alpha) : n < \omega\} = \sup\{f^{n+1}(\alpha) : n < \omega\}) = f(\sup\{f^n(\alpha) : n < \omega\})$$

is a fixed point, by normality. ■

In fact this shows that every normal function has *arbitrarily large* fixed points.

Then we can contemplate the ordinal ϵ_0. Veblen hierarchy!

[lectured but not yet written up]

Remark 2 *Every regular ordinal is initial*

Proof:

It's not a particularly deep or important fact but it's basic and will help you orient yourself. And the proof is idiomatic. Actually we prove the contrapositive.

We need a factoid. Suppose $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ are (strict) total orders, with $<_A$ a wellorder, and there is a bijection $f : A \twoheadrightarrow B$. We are *not* assuming that f is order-preserving! Nevertheless f does have a maximal order-preserving restriction, a rather special one: there is $A' \subseteq A$ s.t $f \restriction A'$ is order-preserving, and f"A' is cofinal (unbounded) in $\langle B, <_B\rangle$.

We obtain A' by recursion on $\langle A, <_A\rangle$. The first member of A' is the bottom element of $\langle A, <_A\rangle$. Thereafter the next member is always the $<_A$-least element a of A. s.t. $f(a) >_B f(a')$ for all $a' <_A a$ that we have already put into A'. Suppose f"A' were bounded in $\langle B, <_B\rangle$. Consider the subset $B' \subseteq B$ consisting of things not dominated by any $f(a)$ for $a \in A'$, and consider the $b \in B'$ s.t. $f^{-1}(b)$ is $<_A$-minimal. $f^{-1}(b)$ should have been put into A'.

Now suppose β is not an initial ordinal. (As I said, we are proving the contrapositive). Then there is $\alpha < \beta$ s.t. α has as many predecessors as β. Let $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ (as in the factoid) be the ordinals below α and the ordinals below β respectively. The factoid gives us a set of ordinals cofinal in β whose order type $\leq \alpha < \beta$. So β is not regular. ■

Here is an illustration of a particular case.

The picture shows why every countable limit ordinal has cofinality ω. The long right-pointing arrow represents a countable ordinal manifested as a wellordering of naturals ($\mathbb{N}$ in a funny order). The (unbounded!) increasing sequence of natural numbers reading from the left are the numbers chosen as in the recursion ... 1001 is the least natural number > 257 that is above 257 in both orders. The semicircle represesents where this increasing sequence of naturals comes to a halt, closes off. Are there any natural numbers in the region flagged by the question marks? Suppose there were—347, say. OK, so what were doing declaring 1001 to be the 6th member of the sequence? We should have used 347!

$\aleph^2 = \aleph$

We start by noting that $\aleph = \aleph + \aleph$. (Well, what we will *actually* need is $\aleph + \aleph + \aleph = \aleph$, but never mind). Beginners might like to have this spelled out, and it holds because $2 \cdot \omega_\alpha = \omega_\alpha$. How so? Any order of limit order-type consists of lots of concatenated copies of $\mathbb{N}$, each of length ω. You can interleave two worders of length ω to get a worder of length ω so you can do this for all the copies simultaneously.

We start by defining a function $\mathfrak{S} : On \to On$. Given an ordinal α, take a wellordering $\langle A, <_A \rangle$ of order type α, make disjoint copies of all its proper initial segments, and then concatenate the copies ... with longer things appended after shorter things. The result is a wellordering and its order type is defined to be $\mathfrak{S}(\alpha)$. [This notation is not standard].

Lemma 2

(i) $\mathfrak{S} : On \to On$ *is a normal function;*
(ii) Every initial ordinal is a value of $\mathfrak{S}$.

Proof:
(i)
$\mathfrak{S} : On \to On$ evidently also has a recursive definition:

$\mathfrak{S}(\alpha + 1) = \mathfrak{S}(\alpha) + \alpha$ and
$\mathfrak{S}(\lambda) = \mathrm{Sup}\{\mathfrak{S}(\alpha) : \alpha < \lambda\}$ for λ limit.

... from which it is clear that $\mathfrak{S}$ is a normal function.

(ii)
Use the division algorithm for normal functions to show that there is a β s.t $\mathfrak{S}(\beta) \leq \omega_\alpha < \mathfrak{S}(\beta + 1)$. If $\mathfrak{S}(\beta) < \omega_\alpha$ then we have $\omega_\alpha \leq \mathfrak{S}(\beta + 1) = \mathfrak{S}(\beta) + \beta$ which is impossible, since $\mathfrak{S}(\beta)$ and β both have cardinality below $\aleph_\alpha$. ■

We want to show that $(\aleph_\alpha)^2 = \aleph_\alpha$. $\aleph_\alpha$ is defined as the cardinal $\{\beta : \beta < \omega_\alpha\}$, which means that the canonical set of size $(\aleph_\alpha)^2$ is the cartesian product $\{\beta : \beta < \omega_\alpha\} \times \{\beta : \beta < \omega_\alpha\}$. We partition this last set into three pieces:

(i) the [graph of] the identity relation restricted to $\{\beta : \beta < \alpha\}$, and

(ii), (iii)

the two triangles above-and-to-the-left, and below-and-to-the-right of the diagonal.

To be slightly more formal about it, we partition the cartesian product $\{\beta : \beta < \alpha\} \times \{\beta : \beta < \alpha\}$ into the three pieces $\{\langle \beta, \gamma \rangle : \beta < \gamma < \alpha\}$, $\{\langle \beta, \gamma \rangle : \beta = \gamma < \alpha\}$ and $\{\langle \beta, \gamma \rangle : \gamma < \beta < \alpha\}$.

It is clear that the third piece is of order type $\mathfrak{S}(\alpha)$ in the lexicographic order.

The idea is to show that these three pieces all have cardinality $\aleph_\alpha$. That's obvious for the second piece, the identity relation. Also there is an obvious bijection between the first and third piece ("flip your ordered pairs") so it will suffice to prove that the third piece ("the bottom-right triangle") has cardinality $\aleph_\alpha$.

Now we can prove

THEOREM 6 $(\forall\alpha)(\aleph_\alpha = (\aleph_\alpha)^2)$

Proof:

By induction on α.

Assume true for all alephs $< \aleph_\alpha$. By lemma 2, ω_α is a value of $\mathfrak{S}$; we want to show that it is actually a fixed point. Now ω_α is an initial ordinal, which is to say that for any $\beta < \omega_\alpha$, the cardinal $|\{\gamma : \gamma < \beta\}|$ is less than $\aleph_\alpha$, and (by induction hypthesis) is equal to its own square. Suppose ω_α were $\mathfrak{S}(\beta)$ for some $\beta < \omega_\alpha$. This would entail that the size of the cartesian product $\{\gamma : \gamma < \beta\} \times \{\gamma : \gamma < \beta\}$ is at least $\aleph_\alpha$, contradicting the induction. So ω_α is a fixed point of $\mathfrak{S}$. This means that the lower-right triangle of the cartesian product $\{\gamma : \gamma < \omega_\alpha\} \times \{\gamma : \gamma < \omega_\alpha\}$—which can be wellordered to length $\mathfrak{S}(\omega_\alpha) = \omega_\alpha$—is of cardinality $\aleph_\alpha$. It's clearly naturally isomorphic to the upper-left triangle (as remarked earlier) so the cartesian product is now a union of three sets each of size $\aleph_\alpha$, giving $(\aleph_\alpha)^2 = \aleph_\alpha + \aleph_\alpha + \aleph_\alpha = \aleph_\alpha$ as desired. ■

So every aleph is equal to its square. Suppose every (infinite) cardinal is equal to its square ... so that all infinite cardinals behave like alephs, in this respect at least ... might that mean that all infinite cardinals are alephs?

The answer is *yes*, and for the moment i am leaving it as an exercise.

Ah, a bit of notation, an *overloading* of the letter '$\aleph$'. $\aleph(\alpha)$ (where α is a cardinal not an ordinal) is the least aleph $\not\leq \alpha$. Thus, for example, $\aleph(\aleph_\beta) = \aleph_{\beta+1}$.

Talk very briefly about the Veblen hierarchy

CH and GCH and Jordan-König.

We can't put off talking about AC any longer!

AC = axiom of Choice; ZL is Zorn's Lemma[1] WO is the Wellordering Principle every set can be wellordered)

REMARK 3 *WO implies AC.*

Proof: Suppose you can wellorder anything that is shown to you, and you want a choice function on a family X of nonempty sets. You wellorder $\bigcup X$ by $<$ and then, for each $x \in X$, declare $f(x)$ to be the $<$-least element of x. ■

[1] What is yellow and equivalent to the axiom of choice?

Remark 4 *ZL implies WO.*

Proof: You are given a set X and you want to wellorder it. Your weapon is ZL, which means that whenever you have a chain-complete poset, it will have a maximal element. How do you use ZL? Well, you seek a chain-complete poset such that a maximal element of it is a wellordering of X. How about taking your chain-complete poset to be the poset of wellorderings of subsets of X (thought of as subsets of $X \times X$) ordered by $\subseteq$? Not quite. The problem is that a union of a chain of wellorderings under $\subseteq$ might not be a wellordering. You need to partially order the wellorderings by **end-extension**. ("Any new stuff must be put after all the old stuff") ■

Remark 5 *AC implies WO*

Again we have a matching challenge. We want to wellorder a set X and we are told we can have a choice function on any family of nonempty sets that we like. The obvious suspect is $\mathcal{P}(X) \setminus \{\emptyset\}$. We now define, by recursion on the ordinals, a sequence s of elements of X indexed by ordinals By AC, the family $\mathcal{P}(X) \setminus \{\emptyset\}$ of nonempty sets has a choice function f. Then we declare $s(\alpha)$, the αth member of our sequence to be $f(X \setminus \{s(\beta) : \beta < \alpha\})$.

How can we be sure that we do not run out of ordinals? Hartogs' lemma (theorem 4) tells us that there is a wellording too big to be embedded in X. ■

A function $f : \langle X, \leq_X \rangle \to \langle X, \leq_X \rangle$ is **inflationary** if $(\forall x \in X)(x \leq_X f(x))$.
For AC $\to$ ZL we need

Theorem 7 *The Bourbaki-Witt theorem*
Every inflationary function from a chain-complete poset into itself has a fixed point.

Proof: Let $\langle X, \leq_X \rangle$ be a chain complete poset, and let $f : X \to X$ be inflationary. The idea is to build a chain, starting at some (any) $x \in X$, extend it at successor stages by doing f to the latest element obtained, and at limit stages by taking sups—$\langle X, \leq_X \rangle$ is chain complete. If we reach a fixed point at any stage we have our hearts' desire. If we do not reach a fixed point then we have injected On into X. But then, by replacement and separation, On must be a set, since it is a 1-1 copy of a subset of X. But we know from corollary 2 that On is not a set. ■

Corollary 3 *AC implies ZL*

Proof: Now let $\langle X, \leq_X \rangle$ be a chain-complete poset. By AC we have a choice function f on $\mathcal{P}(X) \setminus \{\emptyset\}$. The function

$$x \mapsto \texttt{if } x \texttt{ is } \leq_X\text{-maximal } \texttt{then } x \texttt{ else } f(\{x' \in X : x <_X x'\})$$

is inflationary and must have a fixed point by theorem 7. That fixed point will be maximal by construction. ■

3 Large Cardinals

3.1 Filters

A filter in a boolean algebra is a subset F of the carrier set that is closed under $\geq$ and $\wedge$, that is, it satisfies the two conditions:

$$x \in F \wedge x \leq y \rightarrow y \in F$$

and

$$x, y \in F \rightarrow x \wedge y \in F.$$

A filter in the power set algebra $\langle \mathcal{P}(X), \subseteq \rangle$ is said to be a filter **on** X. We should think of a filter on X as a concept of largeness (of subsets of X). This seems reasonable if we reflect on the easiest examples: the cofinite subsets of $\mathbb{N}$ (these are the sets X such that $\mathbb{N} \setminus X$ is finite) are clearly large in some sense. This motivates two other clauses that we almost always assume and which one can easily overlook.

1. Proper filters. According to the definition of filter, the whole algebra is a filter. However, it is not a **proper** filter. All other filters are proper. If the filter generated by a set of elements is proper, we say the set is a **filter base**.

2. Nonprincipal filters. There are pathological filters that do not accommodate the "largeness" intuition. If b is any element of a boolean algebra $\mathcal{B}$, then $\{b' \in \mathcal{B} : b' \geq b\}$ is a filter in $\mathcal{B}$. It is the **principal filter generated by** b. We will think of principal filters as pathological and will not be interested in them. The remaining filters are **nonprincipal**, like the filter of cofinite subsets of $\mathbb{N}$ we saw earlier.

Exercise 1 *Check that the proper filters in a fixed boolean algebra form a chain-complete poset.*

Let F be a filter in a boolean algebra $\mathcal{B}$. $\{\neg y : y \in F\}$ is an **ideal**. Indeed it is the **dual ideal** to F. Ideals in boolean algebras are so-called because they correspond to ideals in boolean rings. But we haven't got time for boolean rings!

Definition 3 *A filter F satisfying any of the conditions below is said to be an* **ultrafilter**.

1. *F is $\subseteq$-maximal among proper filters.*

2. *$(\forall x \in \mathcal{B})(x \in F \vee \neg x \in F)$.*

3. *For all $a, b \in \mathcal{B}$, if $(a \vee b) \in F$, then either $a \in F$ or $b \in F$. (F is* **prime**.*)*

The ideal dual to an ultrafilter is **prime**.

We tend to use $\mathcal{CALLIGRAPHIC}$ font capitals for variables ranging over ultrafilters.

Exercise 2 *Prove that the definitions of definition 3 are equivalent.*

There are natural examples of filters on sets: we saw earlier the filter of cofinite subsets of $\mathbb{N}$, and indeed for *any* infinite set X the collection of cofinite subsets of X is a filter on X. Unfortunately, the only natural examples of ultrafilters are trivial. If x is any element of a set X, then $\{X' \subseteq X : x \in X'\}$ is a principal ultrafilter on X, and, unless we assume something like the axiom of choice, this is the only kind of ultrafilter whose existence can be demonstrated.

If we do assume the axiom of choice we can prove that there are ultrafilters aplenty:

Theorem 8 *(The prime ideal theorem)*
Every boolean algebra has an ultrafilter.

Proof: Consider the set of filters in a boolean algebra $\mathcal{B}$. They are partially ordered (by $\subseteq$, as we have remarked earlier in exercise 1); also, any $\subseteq$-chain of filters has an upper bound (which is simply the union of them all) so the assumptions of Zorn's lemma are satisfied. Therefore there are maximal filters. These are ultra, by exercise 2. ∎

Since proving that there are ultrafilters is the same as proving that there are maximal ("prime") ideals, the name should not cause puzzlement: it is simply a question of which terms you propose to think in.

Since, as we noted on page **??**, upper sets in (complete) posets are (complete) posets, we can even prove the apparently stronger assertion that every filter in a boolean algebra $\mathcal{B}$ can be extended to an ultrafilter. It is not in fact any stronger because, if we seek an ultrafilter extending a given filter F, we form the quotient algebra $\mathcal{B}/F$, use theorem 8 to find an ultrafilter, and then form the set of all elements of $\mathcal{B}$ that got sent to the ultrafilter in the quotient. This set is an ultrafilter extending F.

In fact, by being careful in the choice of a chain-complete poset we can even prove:

Exercise 3 *If $\mathcal{B}$ is a boolean algebra with nonprincipal filters, then it has a nonprincipal ultrafilter.*

Algebras have products and quotients. A homomorphism from $\mathcal{A}$ to $\mathcal{B}$ is a map h such that, if a tuple $\vec{a}$ of elements of $\mathcal{A}$ stands in some (atomic) relation R in $\mathcal{A}$, then the tuple $h(\vec{a})$ stands in the same relation R in $\mathcal{B}$. In fact, there is usually more one can say about homomorphisms than this. In the case of boolean algebras (which are the only algebras we are going to be interested in here), any filter gives rise to a homomorphism. As noted earlier, a filter corresponds to a notion of largeness. Thus, if we have a filter F in a boolean algebra $\mathcal{B}$, it is natural to think of b and b' in $\mathcal{B}$ being similar if their symmetric difference $b\Delta b'$ is *small*, which is to say, its complement is in the filter. Thus we have $b \sim_F b'$ iff the complement of $(b\Delta b') \in F$.

Exercise 4

1. *Check that $\sim_F$ is the same as $(\exists c \in F)(c \wedge b = c \wedge b')$.*
2. *Check that $\sim_F$ is a congruence relation for the boolean operations.*
3. *Prove that the function sending elements of $\mathcal{B}$ to their equivalence classes is a boolean algebra homomorphism.*

Definition 4 *The algebra whose elements are equivalence classes under $\sim_F$ is the* **quotient algebra modulo** *F. The* **kernel** *of a homomorphism of boolean algebras is the set of elements sent to 0.*

This enables us to prove the *Stone representation theorem*. A representation theorem you already know is the representation theorem for groups: every group is (isomorphic to) a group of permutations of a set. The most obvious examples of boolean algebras all have sets as their elements and set inclusion ($x \subseteq y$) as their partial order, but not all do: quotient algebras typically do not. The Stone representation theorem is the assertion that nevertheless

Theorem 9 *(Stone's representation theorem) Every boolean algebra is isomorphic to a boolean algebra whose elements are sets, whose partial order is $\subseteq$, and whose $\vee$ and $\wedge$ are $\cup$ and $\cap$.*

Not to be lctured

Proof: The hard part is to find the isomorphic algebra; the rest is easy. Given $\mathcal{B}$, construct $\mathcal{B}'$ as follows. Send each $b \in \mathcal{B}$ to $\{\mathcal{U} : b \in \mathcal{U}\}$ (the set of all ultrafilters in $\mathcal{B}$ containing b). $\mathcal{B}'$ will be the image of $\mathcal{B}$ in this map. Obviously, if $b \leq c$, then any ultrafilter containing b will contain c but not vice versa, unless $c \leq b$. If b is strictly below c, then consider the principal filter generated by $c \wedge \neg b$. Extend this to an ultrafilter by theorem 8. This ultrafilter will contain c but not b. Thus $b \leq c \longleftrightarrow \{\mathcal{U} : b \in \mathcal{U}\} \subseteq \{\mathcal{U} : c \in \mathcal{U}\}$.

■

Theorems 8 and 9 are in fact equivalent. Although we used Zorn to prove them, there is no converse. Nevertheless, there is a list of natural assertions equivalent to them, though it is not as long as the list of equivalents of AC. The most interesting item is probably: a product of compact Hausdorff spaces is compact Hausdorff, but that is *hard*!

3.2 Ultraproducts and Łoś's theorem

Let T be a first-order theory. Clearly, if every finitely axiomatised subsystem of T has a model, then every finitely axiomatised subsystem of T is consistent. This tells us that T itself is consistent (by compactness) and thus that T itself has a model, by the completeness theorem.

We have inferred that T has a model from the news that all its finite subsets have models, but our proof has involved something very like a *détour* through syntax. In the spirit of the interpolation lemma one might expect that there

should be a function that will accept a set of models and output a model, so that we can feed it models of finite subsets of T and obtain models of T.

There certainly are constructions that accept sets of models and output (single) models: recall that $\prod_{i\in I}\mathcal{A}_i$ is the direct (sometimes called *Cartesian*) product of the A_i.

If $\{\mathcal{A}_i : i \in I\}$ is a family of structures, we define the product

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

to be the structure whose carrier set is the set of all functions f defined on the index set I such that $(\forall i \in I)(f(i) \in A_i)$ and the relations of the language are interpreted by $R(f,g)$ iff $(\forall i \in I)(R(f(i), g(i)))$. The $\{\mathcal{A}_i : i \in I\}$ are said to be the **factors** of the product $\prod_{i\in I}\mathcal{A}_i$. For this operation to make sense it is of course necessary that all the A_i should have the same signature!

Products are nice in various ways. They preserve horn sentences. What do we mean by "preserve"?

Definition 5 *Let Γ be a class of formulæ. Products* preserve Γ *if, whenever $\prod_{i\in I}\mathcal{A}_i$ is a product of a family $\{\mathcal{A}_i : i \in I\}$ and $\phi \in \Gamma$, then $\prod_{i\in I}\mathcal{A}_i \models \phi$ iff $(\forall i \in I)(\mathcal{A}_i \models \phi)$. In these circumstances we also say that ϕ is preserved, when $\phi \in \Gamma$.*

By the definition of product, products preserve atomic formulæ. Clearly they also preserve conjunctions of anything they preserve, and similarly universal quantifications over things they preserve.

Exercise 5 *Verify that products preserve Horn formulæ.*

(This was proved by a man named 'Horn'!) However, they do not always preserve formulæ containing $\vee$ or $\neg$. How so? If ϕ is preserved, then the product will fail to satisfy it if even *one* of the factors does not satisfy it but all the rest do. In these circumstances the product $\models \neg\phi$, but it is not the case that all the factors $\models \neg\phi$. As for $\vee$, if ϕ and ψ are preserved, it can happen that $\phi \vee \psi$ is not, as follows. If half the factors satisfy ϕ and half satisfy ψ, then they all satisfy $\psi \vee \phi$. Now the product will satisfy $\psi \vee \psi$ iff it satisfies one of them. But in order to satisfy one of them, that one must be true at *all* the factors, and by hypothesis it is not. Something similar happens with the existential quantifier.

Given a filter F over the index set, we can define $f \sim_F g$ on elements of the product if $\{i \in I : f(i) = g(i)\} \in F$. This equivalence relation is a congruence relation for all the operations and relations that the product acquires from the factors. At this point it is customary to take a quotient by this congruence relation and call this structure a **reduced product**. This new structure has a different carrier set from the product, but the interpretation of '=' in it is indeed equality. It is possible to keep the same carrier set and obtain much of the effect of "reducing" by $\sim_F$ by taking the interpretation of '=' in the new structure to be $\sim_F$, but this of course gives a nonstandard model.

Then we *either* take this $\sim_F$ to be the interpretation of '=' in the new product we are defining, keeping the elements of the carrier set of the new product the same as the elements of the old, *or* we take the elements of the new structure to be equivalence classes of functions under $\sim$. These we will write $[g]_{\sim_F}$ or $[g]$ if there is no ambiguity. Whichever way you look at it $\sim_F$ is a congruence relation on $\prod_{i\in I}\mathcal{A}_i$.

This new object is denoted by the following expression:

$$(\prod_{i\in I}\mathcal{A}_i)/F.$$

Similarly, we have to revise our interpretation of atomic formulæ so that

$$(\prod_{i\in I}\mathcal{A}_i)/F \models R(f,g) \text{ iff } \{i : R(f(i),g(i))\} \in F.$$

The reason for proceeding from products to reduced products was to complicate the structure and hope to get more things preserved. In fact, nothing exciting happens (we still have the same trouble with $\vee$ and $\neg$) unless the filter we use is ultra. Then everything comes right.

Theorem 10 *Łoś's theorem*
Let $\mathcal{U}$ be an ultrafilter $\subseteq \mathcal{P}(I)$. For all first-order expressions ϕ,

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi \text{ iff } \{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}.$$

Proof: We do this by structural induction on the rectype of formulæ. For atomic formulæ it is immediate from the definitions.

As we would expect, the only hard work comes with $\neg$ and $\vee$, though $\exists$ merits comment as well.

Disjunction

Suppose we know that $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi$ iff $\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}$ and
$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \psi$ iff $\{i : \mathcal{A}_i \models \psi\} \in \mathcal{U}$.
We want to show $(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models (\phi \vee \psi)$ iff $\{i : \mathcal{A}_i \models \phi \vee \psi\} \in \mathcal{U}$.
The steps in the following manipulation will be reversible. Suppose

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi \vee \psi.$$

Then

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \phi \text{ or } (\prod_{i\in I}\mathcal{A}_i)/\mathcal{U} \models \psi.$$

By induction hypothesis, this is equivalent to

$$\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U} \text{ or } \{i : \mathcal{A}_i \models \psi\} \in \mathcal{U},$$

both of which imply

$$\{i : \mathcal{A}_i \models \phi \vee \psi\} \in \mathcal{U}.$$

which implies

$$\{i : \mathcal{A}_i \models \phi\} \cup \{i : \mathcal{A}_i \models \psi\}.$$

Now we exploit the fact that $\mathcal{U}$ is ultra: for all A and B it contains $A \cup B$ iff it contains at least one of A and B, which enables us to reverse the last implication.

Negation

We assume $(\prod_{i\in I} \mathcal{A}_i)/\mathcal{U} \models \phi$ iff $\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}$ and wish to infer $(\prod_{i\in I} \mathcal{A}_i)/\mathcal{U} \models \neg\phi$ iff $\{i : \mathcal{A}_i \models \neg\phi\} \in \mathcal{U}$.

Suppose $(\prod_{i\in I} \mathcal{A}_i)/\mathcal{U} \models \neg\phi$. That is to say,

$$(\prod_{i\in I} \mathcal{A}_i)/\mathcal{U} \not\models \phi.$$

By induction hypothesis this is equivalent to

$$\{i : \mathcal{A}_i \models \phi\} \notin \mathcal{U}.$$

But, since $\mathcal{U}$ is ultra, it must contain I' or $I \setminus I'$ for any $I' \subseteq I$, so this last line is equivalent to

$$\{i : \mathcal{A}_i \models \neg\phi\} \in \mathcal{U},$$

as desired.

Existential quantifier

The step for $\exists$ is also nontrivial:

$$(\prod_{i\in I} \mathcal{A}_i)/\mathcal{U} \models \exists x\phi$$

which, by definition of satisfaction for '$\exists$' is equivalent to

$$(\exists f)((\prod_{i\in I} \mathcal{A}_i)/\mathcal{U} \models \phi(f))$$

which, by induction hypothesis, is equivalent to

$$(\exists f)(\{i \in I : \mathcal{A}_i \models \phi(f(i))\} \in \mathcal{U}). \tag{A}$$

Now: if there is an f that, for almost all i, can pick a ϕ-flavoured thing from A_i then almost all A_i must contain a ϕ-flavoured thing:

$$\{i \in I : \mathcal{A}_i \models \exists x\phi(x)\} \in \mathcal{U}. \tag{B}$$

To step backwards from (B) to (A) we use AC to pick a ϕ-flavoured thing from those A_i that have one (which is almost all of them). We don't care what we pick from those A_i that do not contain a ϕ-flavoured thing. ■

If all the factors are the same, the ultraproduct is called an **ultrapower**, and we write '$A^K/\mathcal{U}$' where K is a set and $\mathcal{U}$ an ultrafilter on K.

Define elementary embedding. Give examples

THEOREM 11 *The ("constant function") embedding $i_{\mathcal{U}} : \mathfrak{M} \to \mathfrak{M}^{(\kappa/\mathcal{U})}$ defined by $m \mapsto [\lambda i.m]_{\mathcal{U}}$ is elementary.*

Proof: Of course we have to do this by structural induction on formulæ, but the only hard case is the existential quantifier, and even that is hard in only one direction. After all, if $\mathfrak{M} \models (\exists x)(\phi(x))$, then i of any witness will satisfy ϕ in the ultraproduct. So it will be sufficient to show that, for any $m \in \mathfrak{M}$, if there is $x \in \mathfrak{M}^\kappa/\mathcal{U}$ such that $\mathfrak{M}^\kappa/\mathcal{U} \models \phi(x, i(m))$, then there is $x \in \mathfrak{M}$ s.t. $\mathfrak{M} \models \phi(x, m)$. Consider such an $x \in \mathfrak{M}^\kappa/\mathcal{U}$. It is the equivalence class $[f]_{\mathcal{U}}$ of a family of functions f such that $\{\alpha < \kappa : \phi(f(\alpha), m)\} \in \mathcal{U}$. But then this thing in $\mathfrak{M}$ that is $f(\alpha)$ will serve as the witness in $\mathfrak{M}$. ■

Ultraproducts enable us to give a particularly slick proof of the compactness theorem for predicate calculus. [probably not to be lectured]

THEOREM 12 *(Compactness theorem for predicate logic)*
If Δ is a set of sentences of predicate calculus such that every finite $\Delta' \subseteq \Delta$ has a model, (we say Δ is "finitely satisfiable") then Δ has model.

Proof: Let Δ be a set of wffs that is finitely satisfiable. Let $\mathcal{S}$ be the set of finite subsets of Δ (elsewhere in these notes notated $\mathcal{P}_{<\aleph_0}(\Delta)$), and let $X_s = \{t \in \mathcal{S} : s \subseteq t\}$. Pick $\mathfrak{M}_s \models s$ for each $s \in \mathcal{S}$. Notice that $\{X_s : s \in \mathcal{S}\}$ generates a proper filter on $\mathcal{S}$. Extend this to an ultrafilter $\mathcal{U}$ on $\mathcal{S}$. Then

$$(\prod_{s \in \mathcal{S}} \mathfrak{M}_s)/\mathcal{U} \models \Delta.$$

This is because, for any $\phi \in \Delta$, $X_{\{\phi\}}$ is one of the sets that generated the filter that was extended to $\mathcal{U}$. For any $s \in X_{\{\phi\}}$, $\mathfrak{M}_s \models \phi$, so $\{s : \mathfrak{M}_s \models \phi\} \in \mathcal{U}$. ■

COROLLARY 4 *A formula is equivalent to a first-order formula iff the class of its models is closed under taking ultraproducts.*

The effect of the ultraproduct construction is to add a lot of things whose presence cannot be detected by finitistic first-order methods. An important effect of this is a direct proof of the Upward Skolem-Löwenheim theorem which we saw earlier. But we have no time to go into that here.

3.2.1 Nonstandard Models of Arithmetic

In an ultrapower of the reals one finds elements like the equivalence class of the function $\lambda n.(1/n)$. This is clearly an infinitesimal: it is everywhere bigger than 0 and, for each n, eventually less than $1/n$. This enables us to do something the eighteenth century wanted to do but couldn't, namely, do differential and integral calculus using infinitesimals, and do it rigorously. Presenting analysis in this way has not caught on, but it well might yet. See Keisler (1976a, 1976b).

3.3 Measurable cardinals

Definition 6 *A filter F is κ-complete if $X \subseteq F \wedge |X| < \kappa \rightarrow \bigcap X \in F$. ($F$ is "closed under fewer than κ intersections".)*

...except that when people say a filter is countably complete they always mean that it is $\aleph_1$-complete. Notice that the definition of a filter *tout court* says that a filter is $\aleph_0$-complete!

We can use Zorn's lemma to prove that there is a nonprincipal $\aleph_0$-complete ultrafilter on any countable set. Can we find uncountable cardinals with the analogous property?

Definition 7 *An uncountable cardinal κ is* **measurable** *iff there is a nonprincipal ultrafilter $\mathcal{U}$ on $\{\alpha : \alpha < \kappa\}$ which is κ-complete.*

We insert the qualification 'uncountable' into definition 7 because it is only *uncountable* measurables that have the special properties we are interested in.

Remark 6 *Every measurable cardinal is weakly compact and therefore strongly inaccessible.*

We don't know yet what these mean

(Explain *strongly inaccessible*)

We appeal to this later, in part 3 of proposition 2.

3.4 Measurable Cardinals and Elementary Embeddings

Let us consider the ultrapower $V^\kappa/\mathcal{U}$. There is a problem here because the equivalence classes (of which the model is composed) are proper classes. We get round this by *Scott's trick* of which I told you earlier.

In general there is no reason to suppose that the ultrapower is wellfounded. Suppose we take an ultrapower $V^\omega/\mathcal{U}$ where $\mathcal{U}$ is nonprincipal over ω. Consider the (Von Neumann) integers of the ultrapower. We want to show that they are not wellfounded. Since $\leq_{\mathbb{N}}$ is the same as $\in$ where von Neumann integers is concerned, this will show that the ultrapower is not really well founded (though it will be a model of the (first-order) axiom of foundation—cf Peano arithmetic). To do this it will be sufficient to find a countable family $\langle u_i : i \in \mathbb{N}\rangle$ of elements of $\mathcal{U}$ such that each integer belongs to only finitely many u_i. That way we

can form an infinitely descending sequence of functions $\mathbb{N} \to \mathbb{N}$ (i.e., natural numbers in the sense of the ultrapower) where the values start off sufficiently big, and we create a smaller natural number by decreasing the values at all coordinates in u_1, then at all coordinates in u_2, then $u_3 \ldots u_n$ so that we get an infinitely descending sequence without any coordinate being decreased more than finitely many times. This is easy: just take u_i to be $\{n \in \mathbb{N} : n > i\}$.

However if the ultrafilter is countably complete the ultrapower will be well-founded, and then we get some action.

Proposition 1 *If V is wellfounded and $\mathcal{U}$ is countably complete over κ then $V^\kappa/\mathcal{U}$ is wellfounded.*

Proof. Suppose $\langle [f_i] : i \in \mathbb{N} \rangle$ satisfies $(V^\kappa/\mathcal{U}) \models [f_{i+1}] \in [f_i]$ for each $i \in \mathbb{N}$. Use AC_ω to pick f_i for each i. Then, for each $i \in \mathbb{N}$, let A_i be $\{\alpha < \kappa : f_{i+1}(\alpha) \in f_i(\alpha)\}$. All A_i are in $\mathcal{U}$ by hypothesis. But then, by countable completeness of $\mathcal{U}$, the intersection $\bigcap_{i\in\mathbb{N}} A_i$ is nonempty (it is actually in $\mathcal{U}$), and for any address β in it, it is the case that $(\forall i)(f_{i+1}(\beta) \in f_i(\beta))$, contradicting the assumption that there are no ω-descending $\in$-chains in V.

■

We could have stated something more general:

Exercise 6 *If $\mathcal{U}$ is a κ-complete ultrafilter over a measurable cardinal κ, then Łoś's theorem holds for sentences of $\mathcal{L}_{\kappa\kappa}$ and ultraproducts modulo $\mathcal{U}$*

We then obtain proposition 1 as a consequence of the fact that wellfoundedness can be captured by an expression of $\mathcal{L}_{\omega_1,\omega_1}$.

This does make it sound as if what is really important is the countable completeness of $\mathcal{U}$, since that it what enables us to show that the ultrapower is wellfounded. So why the κ-completeness in the definition of "measurable"? Suppose there is a countably complete ultrafilter $\mathcal{U}$ on κ, and $f : A \twoheadrightarrow \kappa$. Then $\{f^{-1}\text{“}X : X \in \mathcal{U}\}$ is a countably complete ultrafilter on A, so that if we had taken the posession of a countably complete nonprincipal ultrafilter to be our definition of measurable, then any cardinal (surjectively) larger than a measurable one would be measurable.

Next we take the Mostowski collapse of the ultrapower to obtain a new model, which it is customary to denote '$\mathfrak{M}$'. We will do this sufficiently often to have a notation. Let us use 'ψ' for this. We have a picture:

$$V \hookrightarrow_{i_\mathcal{U}} V^\kappa/\mathcal{U} \simeq_\psi \mathfrak{M}$$

giving an elementary embedding $V \hookrightarrow \mathfrak{M}$—namely $\psi \cdot i_\mathcal{U}$—which tradition in this area requires us to write 'j'. Tradition also requires that the measurable cardinal be 'κ', and the Mostowski collapse of the ultrapower be '$\mathfrak{M}$'.

Lemma 3 *j is not the identity.*

Proof:

Think about $\psi[\lambda\alpha.\kappa]$ and $\psi[id]$. They are both ordinals of $\mathfrak{M}$. Clearly we have $\psi[\lambda\alpha.\kappa] > \psi[id]$. Clearly j sends κ to $[\lambda\alpha.\kappa]$ and thence to $\psi[\lambda\alpha.\kappa]$. Now, for any $\beta < \kappa$, Łoś's theorem tells us that $\psi[\lambda\alpha.\beta]$ is an ordinal and, since $\lambda\beta.\alpha$ is almost everywhere less than the identity, it is a smaller ordinal than $\psi[id]$. (Again by Łoś's theorem).

So the situation is this. Ordinals below κ get sent to ordinals less than $\psi[id]$. On the other hand κ itself gets sent to $\psi[\lambda\alpha.\kappa]$ which is strictly bigger. This means that j is not continuous at κ and cannot be the identity. ■

This isn't the same as saying that $\mathfrak{M} \neq V$. For all we know j might be a nontrivial elementary embedding from the universe into itself. As it happens, it isn't—there are none!—but we won't have time to prove that.

Next we show

Lemma 4 *A non-trivial elementary embedding must send some ordinal κ to something $j(\kappa) > \kappa$, and the least such ordinal is initial.*

Proof: (*Idea: if all ordinals are fixed then by induction on rank all sets are fixed too. No ordinal can be sent to anything smaller so some ordinal is moved to something bigger*)

We observe that by elementarity of j we have

$$(\forall x)(\forall y)(y \in x \longleftrightarrow j(y) \in j(x))$$

Now we also have (for any injective j whatever!)

$$(\forall x)(\forall y)(y \in x \longleftrightarrow j(y) \in j\text{“}(x))$$

This implies $j\text{“}x \subseteq j(x)$. (Q: Why don't we get equality? A: $j(x)$ might have members not in the range of j!). This enables us to prove by induction that j can never move anything to a set of lower rank. If j is not the identity, consider an object x of minimal rank that is moved by j. Suppose x is moved but $\rho(x)$ fixed. Then

Suppose $y \in j(x)$. Then $\rho(y) < \rho(j(x)) = \rho(x)$ so $j(y) = y$. So $j(x) \subseteq$ range of j.

$$j(y) \in j(x)$$

iff

$$y \in x$$

by elementarity, so $j(x) = x$ by extensionality and x is not moved. So if x is a thing of minimal rank, $\rho(x)$ is also moved.

It's probably worth noting here that we are exploiting the fact that (the graph of) j is locally a set, in the sense that its intersection with any set is a set. If j had been merely some random external automorphism (in which case—admittedly—we would be working in a nonstandard model) we would not have been able to say "consider an object x of minimal rank that is moved by j ..."

So if anything is moved, an ordinal is moved. Now let κ be the first ordinal moved by j, we must show that κ is an initial ordinal. First we notice that it must be limit. If it is not initial, we have $\langle \kappa, \leq_{On} \rangle \simeq \langle \mu, R \rangle$ for some relation R on some ordinal $\mu < \kappa$. But then $j\langle \kappa, \leq_{On} \rangle \simeq j\langle \mu, R \rangle$ by elementarity, and $j\langle \mu, R \rangle = \langle \mu, R \rangle$ by minimality of κ (this is where we need κ to be limit, so that $\rho\langle \mu, R \rangle < \kappa$) so $j\langle \kappa, \leq_{On} \rangle \simeq \langle \kappa, \leq_{On} \rangle$ contradicting assumption that κ is moved.

■

This has a nice side-effect which we will need:

COROLLARY 5 *If $j : V \hookrightarrow \mathfrak{M}$ with κ the first ordinal moved, then things of rank less than κ are fixed by j, so $\mathfrak{M} \cap V_\kappa = V_\kappa$.*

Proof: One direction of the inclusion we have just proved, and for the other direction, remember $\mathfrak{M} \subseteq V$!

■

PROPOSITION 2 *Fix a κ-complete nonprincipal ultrafilter $\mathcal{U}$ on κ and consider $\mathfrak{M}$, the transitive collapse of the ultrapower. Then*

1. *The elementary embedding $j : V \hookrightarrow M$ is not the identity and κ is the first ordinal moved;*
2. *$\mathfrak{M}$ is closed under κ sequences but not κ^+-sequences;*
3. *$\kappa < 2^\kappa < j(\kappa) < (2^\kappa)^+$.*

We write "$\mathfrak{M}$ is closed under κ sequences" as "$\mathfrak{M}^\kappa \subseteq \mathfrak{M}$' (even tho' we could have written it as "$\mathcal{P}_{\kappa^+}(\mathfrak{M}) \subseteq \mathfrak{M}$' using a notation we already have) because this is the notation most commonly used in the literature.

Proof:

1. Let β be an ordinal below κ. The elementary embedding into the ultrapower will send β to $[\lambda x.\beta]_{\mathcal{U}}$, and this must get sent to β in the Mostowski collapse. This is because anything below it in the ultrapower is $[f]$ where f is almost everywhere less than β. Consider the preimages in f of the ordinals below β. There are only β of them and they add up to a set of measure 1. Therefore one of them is a set of measure 1, which is to say that $f \sim_{\mathcal{U}} \lambda x.\alpha$ for some $\alpha < \beta$. Thus there are precisely β things in $\mathfrak{M}$ below $j(\beta)$, so $j(\beta) = \beta$.

   For any $\alpha < \kappa$ we have $[\lambda x.\kappa] >_{\mathcal{U}} [id] >_{\mathcal{U}} [\lambda x.\alpha]$, so κ is certainly moved. Everything below κ is fixed, so κ is the first thing moved.

2. RTP: $\mathfrak{M}$ is closed under κ-sequences but not κ^+-sequences.

   We will show that if $j\text{“}x \in \mathfrak{M}$, $y \subseteq M$ and $|y| \leq |x|$ then $y \in \mathfrak{M}$. (This will be sufficient to show that $\mathfrak{M}$ is closed under κ-sequences because

$j``\kappa = \kappa \in \mathfrak{M}$.) Represent $y$ as $\{\psi([t_a]) : a \in x\}$ and define $T : j``x \to y$ by $T(j(a)) = \psi([t_a])$. Since T enumerates y, it suffices to show that $T \in \mathfrak{M}$. So, we need a g so that $\psi([\lambda\alpha.g]) = T$, that is, $dom\psi([\lambda\alpha.g]) = j``x$ and for all $a \in x$, $\psi([\lambda\alpha.g])(j(a)) = \psi([t_a])$. Let $[f] = \psi^{-1}(j``x)$. So there is a thing in the ultrapower with the properties we need. By Łoś's theorem if for each $i \in \kappa$ we set $dom(g(i)) = f(i)$, and $(g(i))(a) = t_a(i)$ for each $a \in dom(g(i))$, then clearly g is a thing in V with the required properties.[2]

To show that $\mathfrak{M}$ is not closed under κ^+-sequences it will be sufficient to show that the particular κ^+-sequence $j``\kappa^+$ is not in $\mathfrak{M}$. Suppose $j``\kappa^+ (= \psi([f]_{\mathcal{U}})) \in \mathfrak{M}$. If $A = \{i < \kappa : |f(i)| \leq \kappa\} \in \mathcal{U}$ since κ^+ is regular, there is $\alpha \in \kappa^+ \setminus \bigcup f``A$. But then $j(\alpha) \notin \psi([f])$. If, on the other hand, $B = \{i < \kappa : |f(i)| > \kappa\} \in \mathcal{U}$, define h by induction on ordinals so that $h(i) \in f(i) \setminus \{h(j) : j < i \wedge j \in B\}$. Then $[h] \in_{\mathcal{U}} [f]$, yet h is not constant on any set in $\mathcal{U}$, as $\mathcal{U}$ is nonprincipal. Hence in either case we get a contradiction from the assumption that $j``\kappa^+ = \psi[f]$.

3. Notice that the set of $<_{V^\kappa/\mathcal{U}}$-predecessors of $[\lambda x.\kappa]$ is a quotient of κ^κ (Here κ^κ is of course not an ordinal but is the set of maps from ordinals-below-κ to ordinals-below-κ! O the joys of overloading). κ^κ is the same size as $\mathcal{P}(\kappa)$ so the wellorder they form must be shorter than $(2^\kappa)^+$. Also, (in $\mathfrak{M}$) by elementarity, $j(\kappa)$ is measurable and hence strongly inaccessible (by remark 6) so that $(2^\kappa)^{\mathfrak{M}} < j(\kappa)$. But $\mathfrak{M}$ is closed under κ-sequences by (2) so $\mathcal{P}(\kappa) = (\mathcal{P}(\kappa))^{\mathfrak{M}}$, so $2^\kappa \leq (2^\kappa)^{\mathfrak{M}}$.

Notice that the fact that $\mathfrak{M}$ is not closed under κ^+-sequences entails that $V \neq \mathfrak{M}$. However we can give slightly more information than this.

■

THEOREM 13 $\mathcal{U} \notin \mathfrak{M}$.

Proof: Assume $\mathcal{U} \in \mathfrak{M}$. [3]. If $\mathcal{U} \in \mathfrak{M}$ then the whole of $\mathcal{P}(\kappa)$ is in $\mathfrak{M}$ too, as is κ^κ. So all the equivalence classes $[f]$ are also in $\mathfrak{M}$, in particular $[\lambda x.\kappa]$. Therefore we can reproduce in $\mathfrak{M}$ the proof that $j(\kappa) \leq (2^\kappa)^+$, but $\kappa < j(\kappa) \leq (2^\kappa)^+$ means that $j(\kappa)$ is not strong limit, contradicting the fact that $j(\kappa)$ is measurable in $\mathfrak{M}$.

■

Notice that we have here used that $\mathcal{U}$ is κ-complete, not just countably complete.

[2] Thanks to Nathan Bowler and Phil Ellison for tidying up some infelicities in this proof.

[3] It would be nice to be able to argue: "Then $\mathfrak{M} \models$ "κ is measurable". But $\mathfrak{M} \models$ "$j(\kappa)$ is the first measurable" and $j(\kappa) > \kappa$" but presumably this works only when κ is the first measurable!