LECTURE 12: AXIOMATIC SET THEORY

We start with the proof of theorem 2.8 that was left out last time:

Proof of theorem 2.8. First we prove the following:

Claim 2.1. For any $\beta \in \mathrm{ON}$ there is a *unique* function $g\colon \beta \to V$ (i.e. g is a set) satisfying

$$(\star_\beta) \qquad \forall \alpha < \beta \quad g(\alpha) = F(g \restriction \alpha).$$

Proof of the claim. Assume otherwise, and let β be the least counterexample. First, assume β is a successor ordinal, say $\beta = \tilde{\beta} + 1$. By minimality of β, there is a unique $\tilde{g}\colon \tilde{\beta} \to V$ such that $(\star_{\tilde{\beta}})$ holds. Letting $g = \tilde{g} \cup \{\langle \tilde{\beta}, F(\tilde{g})\rangle\}$, check that $g\colon \beta \to V$ satisfies $(\star_\beta)$ and is in fact the unique such function, contradicting the definition of β.

So we may assume β is a limit ordinal. By assumption, for each $\tilde{\beta} < \beta$ there is a unique $g_{\tilde{\beta}}\colon \tilde{\beta} \to V$ satisfying $(\star_{\tilde{\beta}})$. Now we glue them all together: using replacement, we obtain a (set) function $h\colon \beta \to V$ such that for $\tilde{\beta} < \beta$, $h(\tilde{\beta}) = g_{\tilde{\beta}}$. By the uniqueness, any two different $g_{\tilde{\beta}_0}$ and $g_{\tilde{\beta}_1}$ agree on their common domain, so $g = \bigcup h[\beta]$ (i.e. $\bigcup_{\tilde{\beta}<\beta} g_{\tilde{\beta}}$) is a function. Check that $g\colon \beta \to V$ satisfies $(\star_\beta)$, again contradicting the initial assumption and proving the claim. $\square$

Intuitively, G is obtained by taking the union of all g_β, for $\beta \in \mathrm{ON}$. Of course to be formally correct, we must display a formula ϕ_G defining G, in the sense that $y = G(\alpha) \iff \phi_G(y, \alpha)$. We can let $\phi_G(y, \alpha)$ be the formula expressing

$$\exists g \quad \alpha \in \mathrm{dom}(g) \in \mathrm{ON} \text{ and } \forall \beta \in \mathrm{dom}(g)\ (\star_\beta) \text{ holds, and } g(\alpha) = y$$

where of course you must replace the implicit mention of $y = F(x)$ by an adequate formula $\phi_F(y, x)$. In fact, this represents a simple procedure to obtain ϕ_G from ϕ_F (just insert a given ϕ_F in the above).[1] That ϕ_G actually defines a function follows from the uniqueness of each g, as in the proof of the claim. $\square$

The following lemma is often useful. It says that every set of ordinals is bounded in ON:

Lemma 2.2. *If X is a set of ordinals, $\sigma = \bigcup\{\xi + 1 \mid \xi \in X\}$ is an ordinal; and in fact, it is the least ordinal with the property that $X \subset \sigma$.*

Proof. It is clear that σ is well-ordered by $<$. To see $\sigma \in \mathrm{ON}$, it remains to show it is transitive: this is because the union of a set of transitive sets is itself transitive (let $x \in y \in \sigma$; there is $\xi \in X$ such that $y \in \xi + 1$, so $x \in \xi + 1$ and thus, $x \in \sigma$). This shows $\sigma \in \mathrm{ON}$.

If $\xi \in X$, $\xi \in \xi + 1$ so $\xi \in \sigma$; thus $X \subseteq \sigma$. It remains to show σ is least with this propery.

Assume $X \subseteq \beta \in \mathrm{ON}$, and show $\sigma \subseteq \beta$ (which is equivalent to $\sigma \leqslant \beta$): If $\alpha \in \sigma$, for some $\xi \in X$ we have $\alpha \leqslant \xi$ (see above). As $X \subseteq \beta$, $\xi \in \beta$. So $\alpha \leqslant \xi < \beta$ and we're done. $\square$

We shall write $\sup(X)$ for σ defined as above.

[1] Thus, the schema of transfinite recursion is actually a recursive set of formulas, which is nice to know.

4. Cardinals

We may consider ‘existence of a bijection’ as defining a (possibly partial) ordering between sets:

Definition 4.1. For any two sets x, y, write $x \preceq y$ *iff* there is an injective function $h\colon x \to y$. Write $x \approx y$ if there is a bijection between x and y.

Using the axiom of choice, one can show that $x \preceq y$ is also equivalent to: there is a surjective function $h\colon y \to x$ (use the the axiom to pick an element from the preimage under h of each $z \in y$ and define $\bar{h}\colon y \to x$ using these and some dummy value outside the range of h).

Interestingly, in the presence of the other axioms of ZF, AC is equivalent to the statement that $\preceq$ is a total order, that is, $\forall x, y\ x \preceq y$ or $y \preceq x$. It may be a little surprising that the following (which you have proved as an exercise) was shown *without* using AC:

Theorem 4.2 (Schröder-Bernstein). *If $x \preceq y$ and $y \preceq x$ we have $x \approx y$.*

If we look at the behavior of $\approx$ on ON we arrive at the notion of *cardinal number*:

Definition 4.3. An ordinal α is called a *cardinal number* (or just: a *cardinal*) if and only if there is no $\beta < \alpha$ such that $\alpha \approx \beta$.

It follows from the previous theorem that we can replace $\approx$ by $\preceq$ in the above definition. The proof of the following is left to you:

Proposition 4.4. *For any $\alpha \in \mathrm{ON}\backslash\omega$, $\alpha + 1$ is* not *a cardinal. Neither is $\alpha + n$ nor $\alpha + \omega$.*

Proposition 4.5. *Each $n \in \omega$ is a cardinal, as is ω.*

Proof. We show by induction that for each $n \in \omega$, the following holds:

$$(\ast_n) \qquad \text{for all } f\colon n \to n, \text{if } f \text{ is 1-to-1 then } f \text{ is onto.}$$

The case $n = 0$ is trivial. Say $(\ast_n)$ holds and let $f\colon n+1 \to n+1$ be arbitrary. First assume that $n+1$ is not in the range of f. Then $f \upharpoonright n$ must be onto n, contradicting that f is injective. So $n+1$ is in the range; let $f(k) = n+1$ and $f(n+1) = l$. The function $g\colon n \to n$ defined by

$$g(m) = \begin{cases} l & \text{if m=k,} \\ f(m) & \text{otherwise} \end{cases}$$

is injective, so by $(\ast_n)$ it is onto. But then f is onto as well.

Clearly, $(\ast_n)$ implies that there can be now bijection $f\colon n \to m$ for $m < n$, proving the first statement of the theorem.

If $f : \omega \to n$ is a bijection, $f \upharpoonright n\colon n \to n$ would be injective. But then it must be onto, contradicting that f was injective, so ω is a cardinal. $\square$

We say a cardinal α is *finite* if and only if $\alpha \in \omega$.

Definition 4.6. For any set, let $\mathrm{card}(x)$ denote the $<$-least cardinal α such that $\alpha \approx x$, if such exists.

Note that you can replace the word ‘cardinal’ by the word ‘ordinal’ in the above (by the definition of cardinal).

The first thing to note is:

Proposition 4.7. *If a set x can be well-ordered, i.e. there is $<_x \subseteq x^2$ which is a well-ordering of x, then* $\mathrm{card}(x)$ *is defined (this can be shown without using the axiom of choice).*

Proof. Given such x, and fixing a well-odering $<_x$ of it, we construct $h\colon \alpha \to x$ which is an isomorphism of the structures $\langle \alpha, < \rangle$ and $\langle x, <_x \rangle$. Obviously, h also witnesses $\mathrm{card}(x) = \alpha$.

Apply the schema of transfinite recursion to obtain a class function $F\colon V \to V$ satisfying:

$$F(\alpha) = \begin{cases} \text{the } <_x\text{-least } u \in x \setminus \mathrm{ran}(F \restriction \alpha, \text{ if it exists} \\ w_0 \text{ otherwise,} \end{cases} \tag{1}$$

for $\alpha \in \mathrm{ON}$–where w_0 is some arbitrary set which is not an element of x. We show that for some ordinal α, $F \restriction \alpha$ is the isomorphism we are looking for.

Letting $D = \{\xi \mid F(\xi) \neq w_0\}$, $F \restriction D$ is injective, so $G = (F \restriction D)^{-1}$ is a a definable class function whose domain is the set x. Thus, D must be a set by replacement. By the a previous lemma, $\alpha = \bigcup D$ is an ordinal. By construction, $h = F \restriction \alpha$ is the isomorphism we were looking for. □

Corollary 4.8. *For any well-ordering $\langle x, <_x \rangle$, there is precisely one alpha* $\in \mathrm{ON}$ *and precisely one map h such that h witnesses the structures $\langle \alpha, < \rangle$ and $\langle x, <_x \rangle$ are isomorphic.*

Proof. The previous proof shows that such α and h exist. To prove uniqueness, say $\bar{h}$ is an isomorphism of $\langle x, <_x \rangle$ with $\langle \bar{\alpha}, < \rangle$. The map $g = h \circ \bar{h}^{-1}\colon \bar{\alpha} \to \alpha$ is a bijection and clearly preserves $<$. Towards a contradiction, assume there is $\xi \in \bar{\alpha}$ such that $g(\xi) \neq \xi$; as g is order preserving, the least ξ with this property must satisfy $\xi \in \alpha \setminus \mathrm{ran}(g)$, contradicting surjectivity of g. So g is the identity on $\bar{\alpha}$ and $\bar{\alpha} = \alpha$. □

Definition 4.9. The α descibed in the previous corollary is called the *length of* $<_x$. We write $\mathrm{length}(<_x) = \alpha$.

This makes it easy to show:

Proposition 4.10. *There is no largest cardinal.*

Proof. Let $\alpha \in \mathrm{ON}$. Note first that we can form the set of all well-orderings of α, by comprehension, as this is a definanble subset of the powerset of $\alpha \times \alpha$:

$$\{r \subseteq \alpha \times \alpha \mid r \text{ is a well-ordering of } \alpha\}.$$

Any $\beta \in \mathrm{ON}$ such that $\beta \approx \alpha$ induces a well-ordering $\prec$ of α: letting $h\colon \alpha \to \beta$ be some bijection, define $\prec$ by $\xi \prec \nu \iff h(\xi) < h(\nu)$. Thus, the map $r \mapsto \mathrm{length}(r)$ is a surjection from the set of well orderings of α onto the class of β such that $\beta \approx \alpha$. Thus, by replacement, the latter is in fact a set. By Schröder-Bernstein, it is moreover an interval of ordinals. Its supremum is the least cardinal above α. □

Definition 4.11. For any ordinal α, write α^+ for the least cardinal above α. Define the '*aleph-function*' by transfinite recursion:

$$\aleph_0 = \omega, \text{ the least infinite cardinal}$$

$$\aleph_{\xi+1} = (\aleph_\xi)^+$$

$$\aleph_\lambda = \bigcup \{\aleph_\xi \mid \xi < \lambda\} \text{ when } \lambda \text{ is a limit ordinal.}$$

Thus, for each $\alpha \in \mathrm{ON}$, $\aleph_\alpha$ is the α-th cardinal. The least uncountable cardinal is $\aleph_1$.

So far, we only know that well-ordered sets have a cardinality, and in fact, if we don't assume AC, there may be sets whose cardinality in the above sense is undefined.[2]

Theorem 4.12. *Every set can be well-ordered; moreover, the previous statement is equivalent to the axiom of choice (in the presence of the other ZF-axioms).*

Proof. Let x be arbitrary. We have seen that it suffices to find a bijection of x with some ordinal.

First, using the axiom of choice, we find a function $t\colon \mathcal{P}(x)\backslash\{\varnothing\} \to x$ such that for all $y \subseteq x$, $f(y) \in y$: look at the disjoint union of $\mathcal{P}(x)\backslash\{\varnothing\}$, i.e. $\bigcup\{\{y\} \times y \mid y \in \mathcal{P}(x)\}$. This is easily seen to be a set; construct it as $\bigcup \mathrm{ran}(h)$ (using union and replacement), where $h\colon \mathcal{P}(x) \to V$ is given by $h(x) = \{x\} \times x$ (using pairing, cartesion product, powerset). Now if t is such that for each non-empty $y \in \mathcal{P}(x)$, $\mathrm{card}(t \cap \{y\} \times y) = 1$, t is a function with the above property.

Now we can argue as in the proof of theorem 4.7, to obtain $\alpha \in \mathrm{ON}$ and a function $f\colon \alpha \to x$ satisfying

$$f(\alpha) = t(x\backslash \mathrm{ran}(F \restriction \alpha))) \text{ if } \mathrm{ran}(F \restriction \alpha)) \neq x; \tag{2}$$

as before, we can assume that α is the least ordinal such that the $f(\alpha)$ as above is not well-defined. Clearly, f is a bijection.

For the other direction, we prove AC assuming every set can be well-ordered. Let X be some set consisting of pairwise disjoint, non-empty sets. Let $\prec$ be a well-ordering of $\bigcup X$. Let t consist of those $z \in \bigcup X$ such that for the unique $y \in X$ with $z \in y$, z is the $\prec$-least element of y. That this is a set is just an application of the comprehension schema. Clearly, for any $y \in X$, $t \cap y$ contains exactly one element. □

5. Cardinal arithmetic

We now define cardinal addition, multiplication and exponentiation and establish their most basic properties. Unfortunately, it is customary to use the same notation for both cardinal and ordinal operations, although they usually *don't* agree–if the necessary, writers typically issue a parenthetical remark, e.g. α^β (ordinal exponentiation). In this section, we avoid confusion by momentarily denoting *ordinal* operations by $+_{\mathrm{ON}}$, $\cdot_{\mathrm{ON}}$ and $\exp_{\mathrm{ON}}$.

Definition 5.1. Let α and β be ordinals. Define

$$\alpha + \beta = \mathrm{card}(\alpha \dot\cup \beta),$$

where $\alpha \dot\cup \beta$ denotes the *disjoint union* $\{0\} \times \alpha \cup \{1\} \times \beta$;

$$\alpha \cdot \beta = \mathrm{card}(\alpha \times \beta)$$

and

$$\alpha^\beta = \mathrm{card}({}^\beta\alpha),$$

where we write ${}^\beta\alpha$ for the set of functions $f\colon \beta \to \alpha$.

[2]Today we have to work with lots of models of ZF where AC fails, as these come up naturally in some arguments even when working in ZFC. Note that we could, independently of AC, regard the 'cardinality' of a set to be the class of all sets that can be brought in bijection with it, but that's not nearly as nice as the present definition.

Note that some people write α^β for both the set ${}^\beta\alpha$ and its cardinality, which causes surprisingly little confusion.

Proposition 5.2. *(1) For all ordinals α, β we have*

$$\alpha + \beta = \operatorname{card}(\alpha +_{\mathrm{ON}} \beta)$$
$$\alpha \cdot \beta = \operatorname{card}(\alpha \cdot_{\mathrm{ON}} \beta)$$

(2) For $m, n \in \omega$, $m + n$, $m \cdot n$ and m^n agrees with the usual meaning.
(3) For cardinals α, β which are not both finite, we have $\alpha + \beta = \alpha \cdot \beta = \max\{\alpha, \beta\}$.

Proof. The proof of (1) is only sketched: $\alpha +_{\mathrm{ON}} \beta$ is can equivalently be defined as the length of a particular well-ordering of $\alpha \dot\cup \beta$ (put one ordinal 'above' the other), $\alpha \cdot_{\mathrm{ON}} \beta$ as the length of a particular well-ordering of $\alpha \times \beta$ (lexicographic). That this definition is equivalent to ours is a somewhat tedious, but straightdorward induction (which I skip). So the length of these well-orderings is of course in bijection to their respective underlying set ($\alpha \dot\cup \beta$ or $\alpha \times \beta$ respectively).

(2) can be easily shown from (1) and using induction to prove the corresponding fact for $+_{\mathrm{ON}}$ and $\cdot_{\mathrm{ON}}$ and $\exp_{\mathrm{ON}}$ (I also leave this to you).

It remains to show (3). We start with $+_{\mathrm{ON}}$. It is enough to show that for all ordinals κ, $\kappa \approx \kappa + \kappa$, for then $\alpha + \beta \leqslant \max\{\alpha, \beta\}$ and the result follows from Schröder-Bernstein.

Towards a contradiction, let $\kappa \in \mathrm{ON}$ be least such that $\kappa \not\approx \kappa + \kappa$; κ is clearly a cardinal. Let $\prec$ be the lexicographic ordering on $K = \kappa \times \{0\} \cup \kappa \times \{1\}$, defined by

$$(\xi, i) \prec (\eta, j) \iff \xi < \eta \text{ or } (\xi = \eta \text{ and } i < j.$$

We have that

$$\kappa \leqslant \kappa + \kappa \leqslant \operatorname{length}(\prec)$$

. If $\kappa < \operatorname{length}(\prec)$, look at "the κ-th element of $\prec$": letting $h\colon \kappa \times \{0\} \cup \kappa \times \{1\} \to \operatorname{length}(\prec)$ be the unique order-isomorphism, fix $(\beta, j) \in \kappa \times 2$ such that $h(\beta, j) = \kappa$. Note that

$$\kappa \approx \{(\xi, i) \in K \mid (\xi, i) \prec (\beta, j)\} \subseteq \beta \times \{0\} \cup \beta \times \{1\} \cup \{(\beta, 0)\} \approx (\beta + \beta) +_{\mathrm{ON}} 1.$$

Since κ was the least counterexample, this means that $\kappa \preceq \beta$, contradicting that κ is a cardinal.

The proof of (3) is somewhat similar: I only give a sketch. Again, it suffices to show $\kappa \cdot \kappa = \kappa$, so again assume κ is the least counterexample. We think of $\kappa \times \kappa$ as well-ordered by the *maximo-lexicographic ordering*:

$$(\xi, \nu) \prec (\bar\xi, \bar\nu) \iff \begin{cases} \max\{\xi, \nu\} < \max\{\bar\xi, \bar\nu\} \text{ or} \\ (\max\{\xi, \nu\} = \max\{\bar\xi, \bar\nu\} \text{ and } \xi < \bar\xi) \text{ or} \\ (\max\{\xi, \nu\} = \max\{\bar\xi, \bar\nu\} \text{ and } \xi = \bar\xi \text{ and } \nu < \bar\nu). \end{cases}$$

You should draw a picture![3] Now one can show that by leastness of κ, the length of $\prec$ (on $\kappa \times \kappa$) is κ. □

[3]In fact, this ordering well-orders the class of pairs of ordinals and is frequently used to find a definable (class) bijection between this class and ON, called 'Gödel pairing'.

Note that cardinal addition and multiplication are well-defined even in the absence of the axiom of choice; the above proof shows that $\alpha\dot{\cup}\beta$ and $\alpha \times \beta$ can always be well-ordered.

Thus, cardinal exponentiation and multiplication are trivial. Very much unlike cardinal exponentiation, which admits an extremely complex theory. Some basic observations are:

Proposition 5.3. *Let α be any cardinal.*

(1) for $n \in \omega$, $\alpha^n = \overbrace{\alpha \cdot \ldots \cdot \alpha}^{n \text{ times}} = \alpha$.
(2) for any cardinal, $2^\alpha > \alpha$ and for infinite cardinals, $\alpha^\alpha = 2^\alpha$.

Proof. The first item follows from the previous theorem.

Towards a contradiction, assume $2^\alpha \leq \alpha$ (i.e. we prove a little more than what is asked for, but the present proof gives you information even in non-choice settings). Then by Schröder Bernstein, there is also a bijection $h\colon \alpha \to {}^\alpha 2$. Now use Cantor's diagonalization trick: define $f\colon \alpha \to 2$ by

$$f(\xi) = 1 - h(\xi)(\xi),$$

(note $h(\xi)$ is a function from α into 2). This is a contradiction, since $f \notin \operatorname{ran}(h)$.

Since (in concise functional notation)

$${}^\alpha 2 \subseteq {}^\alpha\alpha \subseteq \alpha \times \alpha,$$

the last statement follows from $\alpha \cdot \alpha = \alpha$ for $\alpha \notin \omega$. □

Today we know that almost nothing can be shown about the function $\xi \mapsto 2^\xi$ in ZFC. For example, assuming ZFC is consistent, and letting n be your favorite natural number, we can show also ZFC together with the statement $2^\omega = \aleph_n$ is consistent.[4] So ZFC does not suffice to give a specific value to 2^ω, and in fact apart from a few relatively obvious constraints, this value can be shown to be consistently *anything* you wish. The situation is analogous with 2^α for arbitrary α. Note that $2^\omega = \operatorname{card}(\mathbb{R})$ (this is left as an exercise), so this could be considered a serious shortcoming of ZFC.[5]

If, on the other hand, we consider β^α for arbitrary α and β, the situation becomes even more complex. While in the case of $\xi \mapsto 2^\xi$, we know exactly what is merely consistent with ZFC and what can be shown in ZFC, we do not know these things about α^β in general and there are many open questions–but a surprisingly rich structure theory has emerged within ZFC, called *pcf-theory.*

Very much at the beginning of the history of set theory, Cantor formulated his famous continuum hypothesis:

Definition 5.4. The continuum hypothesis, or CH is the statement $2^\omega = \omega_1$. The generalized continuum hypothesis, or GCH is the statement that for every cardinal α, $2^\alpha = \alpha^+$.

The proof of the following would take another block:

Theorem 5.5. *Both CH and its negation are consistent with ZFC; the same holds for GCH and its negation (assuming that ZFC is itself consistent).*

---

[4]Such so-called *relative consistency results* use a technique called *forcing.*

[5]On the other hand, we can consider adding axioms to ZFC: a very popular one called PFA proves that $2^\omega = \aleph_2$.