MATH655 LECTURE NOTES: PART 0

KAMERYN J. WILLIAMS

Set theory is the study of the mathematical concept of **set**, defined by Georg Cantor as a multiplicity considered as a unity. Despite the seeming simplicity of this concept, it is remarkably general and all of mathematics can be coded in terms of sets. Let me give an incomplete list of some of the themes of contemporary set theoretic research.

- **Large cardinals** are those cardinal numbers whose existence cannot be proven from the ordinary axioms of mathematics. They form a linearly ordered hierarchy which provide us a yardstick by which to measure the logical strength of principles which exceed the standard axioms in strength.
- **Descriptive set theory** concerns the structure of real line, and of Polish spaces more generally.
- **Forcing** was originally introduced by Paul Cohen to settle the consistency of the failure of the continuum hypothesis. It has become our primary tool for proving independence and consistency results. Starting with large cardinals, forcing is used to produce models which satisfy such and such principle, giving an upper bound to the logical strength of that principle.
- **Inner model theory** is, as the name suggests, the study of inner models, proper class-sized models of set theory which are as "narrow" as possible. The first inner model was Gödel's constructible universe L, and later work has produced other canonical inner models, especially canonical inner models which carry such and such large cardinal. Showing that a principle implies the existence of such inner models is a common strategy to give a lower bound to the logical strength of that principle.

In this course we will focus on two of these themes, namely large cardinals and forcing. My hope is to introduce you to these two themes and give you a baseline to understand contemporary research in these themes. In part 0 of this course we will set the groundwork to study these themes.

1. Cantorian set theory: Ordinals and cardinals

Before we get to the formal framework, let's consider some of the most important objects of set theory: ordinals and cardinals. We will start with a 'naive' treatment, and will revisit these objects after we talk about the axiomatic approach. Let us start with ordinals. First, we need the very important concept of a well-founded relation.

Definition 1. A relation R on a set A is *well-founded* if any nonempty $X \subseteq A$ has an R-minimal element. That is, there is $x \in X$ so that no $y \in X$ has $y \mathrel{R} x$.

Definition 2. A *well-order* is a well-founded linear order.

Example 3. Any finite linear order is a well-order. The usual order on $\mathbb{N}$ is a well-order. The standard orders on $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are not well-orders.

Date: January 8, 2019.

This lemma is very useful.

Lemma 4. *A relation R is ill-founded iff it has an infinite descending chain $x_0 \mathrel{\reflectbox{R}} x_1 \mathrel{\reflectbox{R}} x_2 \mathrel{\reflectbox{R}} \cdots$.*

Proof. ($\Leftarrow$) Clearly $\{x_0, x_1, \ldots\}$ does not have a minimal element.

($\Rightarrow$) Let $<_R$ denote the transitive closure of R, i.e. $x <_R y$ if there is a finite list

$$x = z_0 \, R \, z_1 \, R \cdots R \, z_n = y.$$

Pick $x_0 \in \operatorname{dom} R$ so that R restricted to $A_0 = \{y : y <_R x_0\}$, the predecessors of x_0, is ill-founded. Given x_n and A_n the predecessors of x_n, pick $x_{n+1} \, R \, x_n$ so that R restricted to $A_{n+1} = \{y : y <_R x_{n+1}\}$ is ill-founded. This choice can always be made, because if R is ill-founded on A_n this has to happen below some $y \, R \, x_n$. Then $x_0 \mathrel{\reflectbox{R}} x_1 \mathrel{\reflectbox{R}} x_2 \mathrel{\reflectbox{R}} \cdots$, as desired. □

Note that we appealed to choice to find the infinite descending chain. Indeed, this direction of the equivalence is equivalent to a weak form of the axiom of choice, known as the principle of dependent choices.

Definition 5 (Cantor)**.** An *ordinal* or *ordinal number* is the ordertype of a well-order. That is, an ordinal is an equivalence class of well-orders under the order isomorphism relation.

Ordinals will be denoted by Greek letters, usually though not exclusively toward the beginning of the alphabet $\alpha, \beta, \gamma, \ldots$.

Definition 6. ω is the ordertype of $\mathbb{N}$, with its usual ordering. 0 is the ordertype of the empty set and for natural numbers $n > 0$, n is the ordertype of $\{0, \ldots, n-1\}$, with its usual order.

We can do arithmetic with ordinals. Let us start by adding 1. If α is an ordinal, then $\alpha + 1$ is the ordertype of the order obtained from an taking something of ordertype α and adding a new element to the end.

Definition 7. Let α and β be ordinals. Let A and B be disjoint well-ordered sets whose ordertypes are, respectively, α and β.

(1) $\alpha + \beta$ is the ordertype of $A \cup B$ under the orders inherited from A and B and with $a < b$ for all $a \in A$ and $b \in B$. (Think: A followed by B.)
(2) $\alpha \cdot \beta$ is the ordertype of $A \times B$ under the *lexicographic order*: $(a, b) \leq (a', b')$ if $a < a'$ or ($a = a'$ and $b \leq b'$). (Think: B many copies of A.)

Exercise 8. Show these definitions are independent of the choice of A and B.

Remark 9. These operations work for ordertypes for linear orders in general, not just well-orders.

Exercise 10. Let η denote the ordertype of $\mathbb{Q}$. Show that $\eta = \eta + \eta = \eta + 1 + \eta$.

Proposition 11. *Let α and β be ordinals. Then $\alpha + \beta$ and $\alpha \cdot \beta$ are ordinals.*

For notational simplicity, let us hereon identify ordinals with representatives of the ordertype, rather than having to continually pick representatives and give them a new name, as we did above.

Proof. ($\alpha + \beta$) Let X be a nonempty subset of $\alpha + \beta$. If $X \cap \alpha \neq \emptyset$, then the α-least element of $X \cap \alpha$ is the $(\alpha + \beta)$-least element of X. If $X \cap \alpha = \emptyset$, then the β-least element of X is the $(\alpha + \beta)$-least element of X.

($\alpha \cdot \beta$) Let X be a nonempty subset of $\alpha \cdot \beta$, which we think of as the cartesian product $\alpha \times \beta$. Let a_0 be the α-least $a \in \alpha$ so that $(a, b) \in X$ for some b. Now let b_0 be the β-least $b \in \beta$ so that $(a_0, b) \in X$. Then (a_0, b_0) is the $(\alpha \cdot \beta)$-th least element of X. □

Exercise 12. Show that ordinal addition and multiplication are associative.

However, neither is commutative. For example, $2 + \omega = \omega \neq \omega + 2$. And $2 \cdot \omega = \omega \neq \omega \cdot 2$.

We will define ordinal exponentiation later, after we have had a chance to study transfinite induction in some detail.

Definition 13. Let α and β be ordinals. Then $\alpha = \beta$ if there is an order-isomorphism between α and β and $\alpha \leq \beta$ if there is an order-embedding of α into β. Then $\alpha < \beta$ if $\alpha \leq \beta$ and $\alpha \neq \beta$.

Clearly, $\alpha = \beta$ implies $\alpha \leq \beta$ and $\leq$ is transitive.

Theorem 14. *The ordinals are well-ordered by* $<$.

Proof. We will prove this as a series of lemmata, each of which is of independent interest.

Lemma 15. *Let α and β be ordinals. Then $\alpha \leq \beta$ iff α embeds onto an initial segment of β.*

Proof by transfinite recursion. The backward direction of the implication is immediate. For the forward direction, observe that we need only to see that if $A \subseteq \beta$ then A, inheriting β's order, is order-isomorphic to an initial segment of β. We construct this order-isomorphism in our first use of transfinite recursion. Here's the idea: we want to map the least element of A to the least element of β, the next element of A to the next element of β, and so on until we exhaust A. If A were finite or isomorphic to ω, then this would just be like ordinary recursion on $\mathbb{N}$. But it could be that we have to go further. We can indeed go further; because β is well-ordered, there is always a least element of β we have not already assigned to the range of the embedding. So this recursive definition is well-defined, and gives rise to an order-isomorphism of A onto an initial segment of β. To be explicit, the map is

$$f(a) = \min\{b \in \beta : \forall a' < a \; f(a') \neq b\}. \qquad \square$$

Here is an alternative argument, where the use of transfinite recursion is obfuscated.

Proof. Again, it is only the forward direction we must see, and it is enough to see that any $A \subseteq \beta$ embeds as an initial segment of β. Let X be the set of all $a \in A$ so that there is *no* embedding of $A \downarrow a = \{a' \in A : a' \leq a\}$ onto an initial segment of β. We want to see that X is empty, so suppose otherwise towards contradiction. Then because A is well-ordered X has a minimum element a. By minimality $\{a' \in A : a' < a\}$ embeds, say via f, onto an initial segment of β, call it I.

Claim 16. For all $a' < a$, $f(a') \leq a'$.

Proof sketch. Suppose this isn't true, and consider the least point where it fails. Show this leads to a contradiction. $\square$

As a consequence, I must be a strict initial segment of β. Otherwise, some $a' < a$ would get mapped to a, but then $f(a') > a'$, contradicting the claim. But now we can see how to embed $A \downarrow a$ onto an initial segment of β: for $a' < a$ use map a' to $f(a')$ and then map a to the least element of $\beta \setminus I$. This contradicts the definition of a, so it must be that X is empty. That is, every initial segment of A embeds onto an initial segment of β.

Claim 17. If I, J are initial segments of A and f, g are embeddings of, respectively, I, J onto initial segments of β, then f and g agree on the intersection of their domains.

In particular, any initial segment I of A embeds uniquely onto an initial segment of β.

Proof sketch. Suppose they don't agree, then consider the least point where they disagree. Show this leads to a contradiction. $\square$

We can now see how to embed all of A onto an initial segment of β: given $a \in A$ just map a to $f(a)$, where f embeds $A \downarrow a'$ for $a' \geq a$ onto an initial segment of β. This completes the proof. $\square$

Hopefully this demonstrates the utility of transfinite recursion. We could avoid directly appealing to this technique, but it came at the cost of having to repeatedly rely on the well-foundedness of β, with proofs by contradiction inside proofs by contradictions. (It's so bad, I made the inner arguments exercises, rather than proving them myself for you!) On the other hand, with transfinite recursion we get a straightforward construction of the embedding.

Let me introduce a bit of notation before the next lemma.

Definition 18. Let $f : X \to Y$ be a function and suppose $X_0 \subseteq X$. Then $f''X_0 = \{y \in Y : \exists x \in X_0\ y = f(x)\}$ is the *image* of X_0 under f. We use this notation, rather than $f(X_0)$ as you may have seen elsewhere, as later in this course we will see lots and lots of sets whose elements are also subsets. So we need different notation to distinguish the two concepts.[1]

Lemma 19. *Let α be an ordinal. Then $\alpha \not< \alpha$.*

Proof. Suppose otherwise. Then there is $f : \alpha \to \alpha$ embedding α onto an initial segment. Let a_0 be the least element of $A_0 = \alpha \setminus f''\alpha$. Given a_n the least element of A_n, set a_{n+1} to be the least element of $A_{n+1} = A_n \setminus f''A_n$. Then the sequence $\langle a_0, a_1, \ldots\rangle$ is an infinite descending sequence in α, contradicting that α is well-founded. $\square$

Lemma 20. *Any ordinal α is isomorphic to the set of ordinals $< \alpha$ under $<$.*

Proof. Let A be the set of ordinals $< \alpha$. By the previous lemma, for each $\beta < \alpha$ there is a unique $b_\beta \in \alpha$ so that β embeds onto $\{a \in \alpha : a < b_\beta\}$. Let $f(\beta) = b_\beta$. I claim that f is an isomorphism from α to A. $\square$

Hereon we will tacitly identify α with the set of ordinals $< \alpha$. Indeed, this identification will follow from our later, more formal treatment. Note that under this identification $\alpha < \beta$ iff $\alpha \in \beta$ iff $\alpha \subseteq \beta$!

Corollary 21. *$<$ is well-founded on the ordinals.* $\square$

Lemma 22. *The ordinals are linearly ordered by $<$.*

Proof. Exercise. (Hint: use transfinite recursion.) $\square$

As we have seen that $<$ is a well-founded linear order, we are (finally!) done. $\square$

Definition 23. A *limit ordinal* is an ordinal which is the ordertype of a nonempty well-order without a maximum element. A *successor ordinal* is an ordinal which is the ordertype of a nonempty well-order with a maximum element.

Clearly, this divides the ordinals into three classes: limit ordinals, successor ordinals, and 0.

Exercise 24. Show that if α is a successor then there is β so that $\alpha = \beta + 1$.

Exercise 25. Write down a limit ordinal which is a limit of limit ordinals. Write down a limit ordinal which is a limit of limit ordinals which are themselves limits of limit ordinals. How much further can you go in this vein?

Theorem 26 (Burali-Forti theorem)**.** *There is no set of all ordinals.*

[1] One also sees the notation $f[X_0]$ for the image of X_0 under f in set theory texts.

Proof. Exercise. Hint: this is similar to Russell's theorem that there is no set of all sets. □

To distinguish collections which are too big to be sets, such as the collection of all ordinals or the collection of all sets, we introduce the term *class* to refer to collections which may or may not be sets. If a class is not a set we call it a *proper class*. When we regroup for an axiomatic approach to set theory we will see how classes are formally treated and paradox avoided. For now, we will proceed naively, with the remark that we are not allowed to talk about proper classes being elements of other classes, for danger of contradiction.

Definition 27. Let S be a *class* of ordinals. Then the *minimum* of S, written $\min S$, is the least element of S, which exists because S is well-ordered.

Let S be a *set* of ordinals. Then the *supremum* of S is

$$\sup S = \min\{\alpha : \forall\beta \in S\ \beta \leq \alpha\}.$$

If $\sup S \in S$, then $\sup S$ is also the *maximum* of S, written $\max S$. Note that $\max S$ always exists for finite S.

Exercise 28. Show that if S has no maximum element then $\sup S = \min\{\alpha : \forall\beta \in S\ \beta < \alpha\}$.

Definition 29. Given a set A, the *Hartog's number for* A, written $\aleph(A)$, is the least ordinal α so that there is no injection from α to A.

Proposition 30. $\aleph(A)$ *always exists.*

Proof. Because the ordinals are well-ordered we must only see there is some ordinal with this property. To see this: If α is the ordertype of some well-order of $\mathcal{P}(A)$ then by Cantor's theorem there is no bijection from A to α. □

Remark 31. This argument appealed to the axiom of choice, in the guise of assuming that $\mathcal{P}(A)$ could be well-ordered. In fact, one does not need Choice to prove $\aleph(A)$ always exists, as seen in the following exercise.

Exercise 32. Suppose that A is infinite. Show that $\aleph(A)$ is the supremum of the set of ordinals α for which there is an injection from α to A.

Definition 33. $\omega_1 = \aleph(\omega) = \aleph(\mathbb{N})$ is the least uncountable ordinal. Given any ordinal $\beta \geq 2$, ω_β is the least ordinal $> \aleph(\omega_\gamma)$ for $\gamma < \beta$.

Written differently: ω_1 is the supremum of the countable ordinals and ω_β is the supremum of the ordinals of cardinality ω_γ for $\gamma < \beta$.

One consequence of this definition, which should be made explicit, is that if $\beta < \omega_\alpha$ then there is no injection from ω_α to β. In particular, β is smaller in cardinality than ω_α. For this reason, the ω_αs are *initial ordinals*—ordinals so that each smaller ordinal is smaller in cardinality. Observe that the initial ordinals are the finite ordinals and the ω_αs

Proposition 34. *Every infinite ordinal is in bijection with some* ω_β *(or with* ω*).*

Proof. Let α be the least infinite ordinal without this property. Consider the set S of ordinals γ so that some $\beta < \alpha$ is in bijection with ω_γ. Let $\delta = \sup S$. Then $\alpha = \omega_\delta$, a contradiction. □

Exercise 35. Justify to yourself the assertion that S is a set.

We can think of ω_0 as another name for ω, so that the above proposition can be stated without the parenthetical. But no one ever actually writes ω_0.

Exercise 36. Show that if A is a final sequence of ω_α, i.e. $A = \{\beta : \gamma < \beta < \omega_\alpha\}$ for some γ, then the ordertype of A is ω_α.

To round out this discussion of ordinals before we move to cardinals, let us see how ordinal addition and multiplication interact with cardinality. First, let us see that addition does not make infinite ordinals larger in the sense of cardinality.

Exercise 37. Let α be an infinite ordinal. Construct a bijection between α and $\alpha + \alpha$. Conclude that if at least one of α and β is infinite then there is a bijection from $\max \alpha, \beta$ and $\alpha + \beta$.

The same is true for multiplication; one can try to directly construct a bijection between α and $\alpha \cdot \alpha$, but a better approach is to go through a pairing function for ordinals.

Definition 38. Let $\alpha, \beta, \gamma, \delta$ be ordinals. Say that $(\alpha, \beta) \lhd (\gamma, \delta)$ if

- $\max \alpha, \beta < \max \gamma, \delta$; or
- The maxima are equal and $\alpha < \gamma$; or
- The maxima are equal, $\alpha = \gamma$, and $\beta < \delta$.

It is clear that $\lhd$ is a well-order. Define the so-called Gödel pairing function $g(\alpha, \beta)$ to be the ordertype of $\lhd$ below the pair (α, β).

Exercise 39. Justify to yourself that the collection of $(\delta, \gamma) \lhd (\alpha, \beta)$ is a set, and so $g(\alpha, \beta)$ really is a map from pairs of ordinals to ordinals. And since $\lhd$ is a linear order, this immediately implies that g is an injective order homomorphism.

Exercise 40. Show that g is an isomorphism.

Proposition 41. *If α is an infinite ordinal then there is a bijection between α and $\alpha \cdot \alpha$.*

Proof. Because each infinite ordinal is in bijection with an ω_α, by composing bijections it is enough to show that this holds for the ω_αs. So it is enough to see that if $\beta, \gamma < \omega_\alpha$ then $g(\beta, \gamma) < \omega_\alpha$, as then g is a bijection between $\omega_\alpha \times \omega_\alpha$ and ω_α.

Suppose otherwise that this is not the case. Then let (β, γ) be least so that $g(\beta, \gamma) \geq \omega_\alpha$. Indeed, it must be that $g(\beta, \gamma) = \omega_\alpha$ and so $g(\beta, \delta) < \omega_\alpha$ for all $\delta < \gamma$. Note that by the definition of $\lhd$ that $g''(\{\beta\} \times \gamma)$ must be a final segment of ω_α. But then we get a bijection between ω_α and γ by sending $\xi < \omega_\alpha$ to the $\delta < \gamma$ which is the ξ-th element of $g''(\{\beta\} \times \gamma)$. This is a contradiction, because there is no bijection from ω_α to a smaller ordinal. $\square$

Let us now talk about cardinals.

Definition 42 (Cantor)**.** A *cardinal* or *cardinal number* is an equipotency class, that is, an equivalence class of sets under the relation $A \sim B$ iff there is a bijection from A to B.

Given a set A, its cardinal number or cardinality is denoted $|A|$. We will tend to use Greek letters from near the middle of the alphabet $\kappa, \lambda, \mu, \ldots$ for cardinals.

Given two cardinals κ and λ, say that $\kappa \leq \lambda$ if given any K of cardinality κ and L of cardinality λ there is an injection from K to L, and say that $\kappa = \lambda$ if there is a bijection from K to L. And $\kappa < \lambda$ if $\kappa \leq \lambda$ but $\kappa \neq \lambda$.

Exercise 43. Show that the order relations on cardinals are well-defined, and don't depend upon the choice of K and L.

Theorem 44 (Cantor)**.** *For any set A, we have $|A| < |\mathcal{P}(A)|$.*

Proof. It is clear that $|A| \leq |\mathcal{P}(A)|$, as witnessed by the injection $a \mapsto \{a\}$. To see they are not equal, suppose towards a contradiction that $f : A \to \mathcal{P}(A)$ is a bijection. Now set $D = \{a \in A : a \notin f(a)\}$. Let $d = f^{-1}(D)$. Then $d \in D$ iff $d \notin f(d) = D$, a contradiction. □

Let us see that $\leq$ and $=$ play nicely together.

Theorem 45 (Cantor–Schroeder–Bernstein)**.** *$\kappa \leq \lambda$ and $\lambda \leq \kappa$ iff $\kappa = \lambda$.*

Proof. We want to show that if there are injections $f : A \to B$ and $g : B \to A$ then there is a bijection between A and B. First, by instead considering the sets $(g \circ f)''A \subseteq g''B \subseteq A$, observe that it is enough to show that if $X \subseteq Y \subseteq Z$ and $|X| = |Z|$ then $|Y| = |Z|$

Lemma 46. *Suppose $F : \mathcal{P}(Z) \to \mathcal{P}(Z)$ is monotone—$X \subseteq Y$ implies $F(X) \subseteq F(Y)$. Then F has a fixed point, i.e. there is $P \in \mathcal{P}(X)$ so that $F(P) = P$.*

Proof. Let $T = \{X \subseteq Z : F(X) \subseteq X\}$. Clearly $Z \in T$, so $T \neq \emptyset$. Let $P = \bigcap T$. Let's see P is a fixed point. First, take any $X \in T$. Then $P \subseteq X$ so $F(P) \subseteq F(X) \subseteq X$. So $F(P) \subseteq \bigcap T = P$, so $P \in T$. And immediately $F(F(P)) \subseteq F(P)$, so $F(P) \in T$. But then $P \subseteq F(P)$, so $P = F(P)$. □

Now define $F : \mathcal{P}(Z) \to \mathcal{P}(Z)$ as $F(A) = (Z \setminus Y) \cup f''A$, where $f : Z \to X$ is a fixed bijection. It follows from f being one-to-one that F is monotone, so by the lemma let P be a fixed point of F. That is, $P = (Z \setminus Y) \cup f''P$. Then

$$g(x) = \begin{cases} f(x) & \text{if } x \in P \\ x & \text{if } x \notin P \end{cases}$$

is a bijection from Z to Y. □

Definition 47. $0 = |\emptyset|$; for natural numbers $n > 0$, $n = |\{0, \ldots n-1\}|$. For ordinals α, $\aleph_\alpha = |\omega_\alpha|$. In particular, $\aleph_0 = |\omega| = |\mathbb{N}|$.

Note that since every set can be well-ordered, every infinite set is in bijection with some $\aleph_\alpha$.

Proposition 48. *The cardinal numbers are well-ordered by $<$.*

Proof. Because the ordinal numbers are well-ordered. □

Definition 49. If κ is a cardinal, then κ^+ is the smallest cardinal $> \kappa$. Observe that ${\aleph_\alpha}^+ = \aleph_{\alpha+1}$.

Exercise 50. Show that cardinal trichotomy, the assertion that the cardinal numbers are linearly ordered, is equivalent to Choice. (Hint: the as-of-yet unproven direction follows once you know, without using Choice, that every set has a Hartog's number. This was an earlier exercise.)

Definition 51. A set is *finite* if its cardinal number is in $\mathbb{N}$. Otherwise, it is *infinite*.

Proposition 52. *If A is infinite, then $\aleph_0 \leq |A|$.*

Proof. Define that A is *Dedekind-infinite* if there is an injection from A to a proper subset of A. I claim that if A is Dedekind-infinite then $\mathbb{N}$ injects into A. To see this, let $f : A \to A$ be an injection witnessing that A is Dedekind-infinite, then pick $a_0 \in A$ to be a point not fixed by f. Such has to exist as if f fixes all points then ran $f = A$, contrary to assumption. Then set $a_1 = f(a_0)$. Given a_n, set $a_{n+1} = f(a_n)$. We have that $a_{n+1} \neq a_n$ because $a_n = f(a_{n-1})$, so if $f(a_n) = a_n$ then f would not be one-to-one. So $n \mapsto a_n$ injects $\mathbb{N}$ into A.

It remains only to see that A is infinite iff A is Dedekind-infinite. The contrapositive of the backward direction is the observation that no finite set is Dedekind-infinite. (Exercise!) For the forward direction, let $<$ be a well-order of A. Without loss we may assume that $<$ does not have a maximum, because every ordinal is in bijection with some ω_α, which is a limit ordinal. Define a function $f : A \to A$ by setting $f(a)$ to be the $<$-minimum element of $\{b \in A : a < b\}$. It is clear that f is an injection from A to a proper subset. □

Remark 53. That every Dedekind-infinite set is infinite is provable without the axiom of Choice. But it takes a small fragment of Choice to prove that every infinite set is Dedekind-infinite. For example, it is consistent with ZF that there are *amorphous sets*, i.e. infinite sets whose subsets are all finite or co-finite. These cannot be Dedekind-infinite.

Definition 54. An infinite set is called *countable* if its cardinal number is $\aleph_0$. Otherwise, it is called *uncountable*.

Remark 55. It's a matter of convention whether finite sets are called countable or if countable refers specifically to infinite sets. I personally am horribly inconsistent with my usage here, using either convention depending on which is more convenient in a given context. It should always be clear which is meant, however.

As befitting anything called "numbers", we can do arithmetic with cardinal numbers.

Definition 56. Let κ and λ be cardinals. Pick K of cardinality κ and L of cardinality λ.

- $\kappa + \lambda = |K \sqcup L|$, where $K \sqcup L = (K \times \{0\}) \cup (L \times \{1\})$ is the disjoint union of K and L.
- $\kappa \cdot \lambda = |K \times L|$.
- $\kappa^\lambda = \left|{}^L K\right|$, where ${}^L K$ is the set of functions from L to K.

Exercise 57. Show that these are well-defined and don't depend upon the choice of K and L.

Exercise 58. Check that the familiar rules for exponents $\kappa^\lambda \cdot \kappa^\mu = \kappa\lambda + \mu$ and $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$ hold for infinite cardinals.

An easy generalization of Cantor's theorem yields that $\kappa^\lambda > \lambda$ whenever $\kappa \geq 2$. The same is not true for addition and multiplication. Indeed, if at least one of κ and λ are infinite then $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$. To see this, it is enough to see that $\aleph_\alpha = \aleph_\alpha + \aleph_\alpha = \aleph_\alpha \cdot \aleph_\alpha$ for all $\aleph_\alpha$. To prove this, first observe that $\aleph_\alpha + \aleph_\alpha$ is the cardinality of $\omega_\alpha + \omega_\alpha$ (the first $+$ is cardinal addition and the second $+$ is ordinal addition) and $\aleph_\alpha \cdot \aleph_\alpha$ is the cardinality of $\omega_\alpha \cdot \omega_\alpha$ (ditto but for $\cdot$). This follows immediately from the definition of these arithmetic operations. But as we have already seen, ω_α is in bijection with $\omega_\alpha + \omega_\alpha$ and with $\omega_\alpha \cdot \omega_\alpha$.

But an important note: the proof this depends upon Choice. Indeed, Tarski showed that, over ZF, $\kappa \cdot \kappa = \kappa$ for all infinite cardinals κ iff Choice. Jan Mycielski relates the following amusing story about Tarski's theorem:[2] "Tarski told me the following story. He tried to publish his theorem (stated above) in the *Comptes Rendus Acad. Sci. Paris* but Fréchet and Lebesgue refused to present it. Fréchet wrote that an implication between two well known propositions is not a new result. Lebesgue wrote that an implication between two false propositions is of no interest. And Tarski said that after this misadventure he never tried to publish in the *Comptes Rendus*."

Project Idea 59. Cardinal arithmetic without Choice.

More generally, we can define cardinal addition and multiplication for arbitrary families of cardinals.

[2] See http://www.ams.org/notices/200602/fea-mycielski.pdf.

Definition 60. Let $\langle \kappa_i : i \in I \rangle$ be a sequence of cardinals, with disjoint sets K_i of cardinality κ_i. Then

$$\sum_{i\in I} \kappa_i = \left| \bigcup_{i \in I} K_i \right|$$
$$\prod_{i\in I} \kappa_i = \left| \prod_{i \in I} K_i \right|.$$

The product in the right half of the second line is the cartesian product.

Exercise 61. Show that $\sum_{i\in I} \kappa_i = \sup_{i \in I} \kappa_i$.

An important consequence of these cardinal arithmetic facts is the following.

Proposition 62. *Let $\kappa, \lambda < \mu$ be cardinals. Suppose that $\{A_i : i \in \lambda\}$ is a collection of sets each of cardinality $\leq \kappa$. Then*

$$\left| \bigcup_{i \in \lambda} A_i \right| < \mu.$$

In particular, a countable union of countable sets is countable.

Proof.
$$\left| \bigcup_{i \in \lambda} A_i \right| \leq \left| \bigcup_{i \in \lambda} A_i \times \{i\} \right| = \sum_{i \in \lambda} |A_i| \leq \sum_{i\in\lambda} \kappa = \kappa \cdot \lambda = \max\{\kappa, \lambda\} < \mu.$$ □

Once again, let me note an appeal to the axiom of choice. It's consistent with ZF that ω_1 is a countable union of countable sets. Which I personally find weird as hell. (But note that it takes only a weak fragment of choice to rule out this uneasy possibility.)

Let us end this section by stating a hypothesis of some importance.

Question 63 (Cantor)**.** *Because every set can be well-ordered, we know that $2^{\aleph_0}$ must be some $\aleph_\alpha$. The* continuum problem *is to settle this question. The* continuum hypothesis *(CH) is the assertion that $2^{\aleph_0} = \aleph_1$. Observe that CH is equivalent to the assertion that there is no cardinal intermediate between $\aleph_0$ and $2^{\aleph_0}$.*

The *generalized continuum hypothesis* (GCH), originally formulated by Hausdorff, is the assertion that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all α. It is equivalent to the assertion that for any infinite cardinal there is no cardinal intermediate between κ and 2^κ.

2. Transfinite induction and recursion

In this section we are introduced to one of our basic, but powerful tools. Let us begin by recalling the familiar case of induction and recursion on $\mathbb{N}$.

Fact 64 (Induction on $\mathbb{N}$)**.** *Suppose $S \subseteq \mathbb{N}$ has the property that for all $x \in \mathbb{N}$, if all $y < x$ are in S then $x \in S$. Then, $S = \mathbb{N}$.*

This can be put in the equivalent form: if $0 \in S$ and $x \in S$ implies $x + 1 \in S$, then $S = \mathbb{N}$.

Proof. Suppose that $S \neq \mathbb{N}$. By the well-foundedness of $\mathbb{N}$, let $x + 1$ be the least element of $\mathbb{N} \setminus S$. But then $x \in S$ and so $x + 1 \in S$, a contradiction. $\square$

This is a very important property of the natural numbers. Indeed, Dedekind, Peano, and others isolated it as *the* defining property of $\mathbb{N}$, from which most of our theorems about $\mathbb{N}$ can be proved. In general, this is how we think of induction—as a technique for proving things. On the flip side of things is a technique for constructing objects, namely recursion. Or to put on a picky formalist hat: induction is used to prove the existence of objects defined by recursion.

Example 65. Define the factorial function $n!$ on $\mathbb{N}$ as: $0! = 1$ and $(n+1)! = (n+1) \cdot n!$. Then $n!$ is defined for all $n \in \mathbb{N}$.

(Proof: we have seen that $0!$ is defined and that if $n!$ is defined then $(n + 1)!$ is defined. So by induction $n!$ is defined for all natural numbers.)

Recursion in general can be formalized as follows.

Fact 66. *Consider a function $g : \mathbb{N}^2 \to \mathbb{N}$ and let $c_0 \in \mathbb{N}$ be fixed. Then there is a function $f : \mathbb{N} \to \mathbb{N}$ so that $f(0) = c_0$ and for all n, we have $f(n + 1) = g(f(n), n + 1)$.*

Proof. Let $S \subseteq \mathbb{N}$ be the set on which the above definition of f is valid. Clearly $0 \in S$. Assume $n \in S$. But then $f(n+1)$ is defined, since $f(n+1) = g(f(n), n+1)$. So by induction $S = \mathbb{N}$. $\square$

Note that the key property about $\mathbb{N}$ which allows induction is that $\mathbb{N}$ is well-ordered. We can generalize induction to work along the ordinals. This is known as *transfinite induction.*

Definition 67. Ord is the class of all ordinals. We saw earlier that Ord is not a set.

Fact 68 (Transfinite induction)**.** *Suppose that S is a class of ordinals with the property that for all ordinals ξ if every $\eta < \xi$ is in S then $\xi \in S$. Then $S =$ Ord.*

Proof. Exercise. $\square$

There is an alternate formulation which is often useful.

Fact 69. *Suppose that S is a class of ordinals which satisfies the following three properties:*

(1) *$0 \in S$;*
(2) *If $\xi \in S$ then $\xi + 1 \in S$;*
(3) *If λ is limit and $\xi \in S$ for all $\xi < \lambda$, then $\lambda \in S$.*

Then $S = \alpha$.

Proof. Exercise. $\square$

Both versions of transfinite induction also work if we confine to a set of ordinals, rather than looking at the class of all ordinals. (Exercise: formulate this version of transfinite induction and prove that it is valid.)

Let us see an example of how to use transfinite induction.

Proposition 70. *If $\alpha \leq \beta$ then $\alpha + \gamma \leq \beta + \gamma$ and $\alpha \cdot \gamma \leq \beta \cdot \gamma$.*

Proof. By induction on γ. We need to check the 0 case, the successor case, and the limit case. It's clearly true if $\gamma = 0$. If it's true for γ, then the embedding from $\alpha + \gamma$ to $\beta + \gamma$ is easily extended to an embedding from $\alpha + \gamma + 1$ to $\beta + \gamma + 1$: just send the new element to the least unused point in the target. Finally, suppose λ is limit and $\alpha + \gamma \leq \beta + \gamma$ for all $\gamma < \lambda$. If it were not the case that $\alpha + \lambda$ embeds into $\beta + \lambda$, then there would have to be a smallest point in $\alpha + \lambda$ so that the embedding cannot be extended to have that point in the domain. But since $\alpha + \lambda$ can be identified with the set of ordinals $< \alpha + \lambda$, this amounts to saying that there must be $\gamma < \lambda$ so that there is no embedding from $\alpha + \gamma$ to $\beta + \lambda$. This would contradict the existence of an embedding from $\alpha + \gamma$ to $\beta + \gamma$. $\square$

Just as induction on $\mathbb{N}$ allows definitions by recursion, so too does transfinite induction allow transfinite recursion. We will have to be a bit loose with the formulation, since we don't yet have a good formal explication of what a class versus a set is.

First, a bit of notation. If f is a function and $x \subseteq \operatorname{dom} f$ then $f \restriction x$ is the restriction of f to x. I use this notation because (1) it is standard among set theorists and (2) the other main notation for function restriction uses a vertical bar, which is overused.

Definition 71. The principle of *transfinite recursion* asserts: if G is a class function then there is a class function F with domain Ord defined as $F(\alpha) = G(F \restriction \alpha)$.

We will justify this principle later during the formal treatment. (We will need the axiom of Replacement.)

Quite likely, this is not at all an enlightening thing to see. So let's see some examples to get a handle on it. Let's start with alternative definitions of ordinal addition and multiplication. We will also be able to introduce ordinal exponentiation.

Definition 72. We define $\alpha + \beta$, $\alpha \cdot \beta$, and α^β by transfinite recursion. $\alpha + \beta$ is defined as:

- $\alpha + 0 = \alpha$;
- $\alpha + (\beta + 1) = (\alpha + \beta) + 1$; and
- If λ is limit, then $\alpha + \lambda = \sup_{\beta<\lambda} \alpha + \beta$.

$\alpha \cdot \beta$ is defined as:

- $\alpha \cdot 0 = 0$;
- $\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$; and
- If λ is limit, then $\alpha \cdot \lambda = \sup_{\beta<\lambda} \alpha \cdot \beta$.

α^β is defined as:

- $\alpha^0 = 1$;
- $\alpha^{\beta+1} = (\alpha^\beta) \cdot \alpha$; and
- If λ is limit, then $\alpha^\lambda = \sup_{\beta<\lambda} \alpha^\beta$.

Transfinite recursion says that each of these schemata defines a function on the ordinals.

Exercise 73. Show that this matches our previous definitions of ordinal addition and multiplication. (Hint: use transfinite induction.)

Warning! Caution! Achtung! Cardinal and ordinal exponentiation are very far apart. For instance, ω^ω is countable, while $\aleph_0^{\aleph_0}$ is uncountable.

Lemma 74. *Ordinal arithmetic has many familiar properties.*

(1) *If* $\beta < \gamma$ *then* $\alpha + \beta < \alpha + \gamma$.
(2) *If* $\alpha < \beta$ *then there is a unique* γ *so that* $\alpha + \gamma = \beta$.
(3) *If* $\beta < \gamma$ *and* $\alpha > 0$, *then* $\alpha \cdot \beta < \alpha \cdot \gamma$.
(4) *If* $\alpha > 0$ *and* β *is any ordinal, then there are unique* γ *and unique* $\delta < \alpha$ *so that* $\beta = \alpha \cdot \gamma + \delta$.
(5) *If* $\beta < \gamma$ *and* $\alpha > 1$ *then* $\alpha^\beta < \alpha^\gamma$.

Proof. (1) By transfinite induction on γ. Trivial for $\gamma = 0$. For the successor case, we need only to see that $\alpha + \gamma < \alpha + \gamma + 1$. But this is obvious. For the limit case, suppose it fails for λ limit. Then there must be a smallest $\beta < \lambda$ for which it fails, at which point it reduces to the successor case.

(2–5) Exercise. (Hint: (3) and (5) are also proved by transfinite induction.) □

Theorem 75 (Cantor Normal Form)**.** *Every ordinal* $\alpha > 0$ *can be uniquely represented in the form*

$$\alpha = \omega^{\eta_0} \cdot k_0 + \omega^{\eta_1} \cdot k_1 + \cdots + \omega^{\eta_n} \cdot k_n$$

where $n \geq 0$ *is finite,* $\alpha \geq \eta_0 > \eta_1 > \cdots > \eta_n$, *and* k_i *are nonzero finite ordinals.*

Proof. First we see that there is some such representation by induction on α. Here, the base case is $\alpha = 1$, where $1 = \omega^0 \cdot 1$. For larger α, let η be largest so that $\omega^\beta \leq \alpha$. Then $\alpha = \omega^\eta \cdot \gamma + \delta$. Necessarily γ must be finite, so this is $\alpha = \omega^\eta \cdot k$ for $k \in \omega$. And by inductive hypothesis we know that δ can be written in Cantor normal form, giving us Cantor normal for α.

Uniqueness is also proved by transfinite induction. (Do it!) □

Observe that it's possible to have $\alpha = \omega^\alpha$. For example, this is true whenever $\alpha = \omega_\beta$. The smallest ordinal with this property is denoted ε_0. (With ε_ξ being the ξth ordinal with this property.)

Exercise 76. Come up with an alternative characterization of ε_0.

3. Some applications of transfinite induction/recursion

I thought it might be nice to see a couple applications of transfinite recursion/induction, before we continue with the pure set theory. Both of these examples indicate one of the main uses of ordinals: as indices for the steps in transfinite constructions.

Definition 77. A set $P \subseteq \mathbb{R}$ is *perfect* if it contains all of its accumulation points.

Theorem 78 (Cantor–Bendixson)**.** *Every closed set $C \subseteq \mathbb{R}$ can be decomposed into the disjoint union of a perfect set and a countable set.*

Corollary 79. *The continuum hypothesis holds for closed sets. That is, any closed set of reals is either at most countable or else equipotent with the reals.*

Proof. Exercise! (Hint: show that the Cantor set can be injected into any nonempty perfect set.) $\square$

Proof of the Cantor–Bendixson theorem. Given a set $A \subseteq \mathbb{R}$, define the *Cantor–Bendixson derivative of A*, denoted A', to be the set of accumulation points of A. Thus, if A is closed, then A' is also closed. Now define the following transfinite sequence of sets, using a closed set A:

- $A^0 = A$;
- $A^{\alpha+1} = (A^\alpha)'$; and
- If λ is limit, then $A^\lambda = \bigcup_{\alpha<\lambda} A^\alpha$.

Make a couple observations: First, A^α is always closed. Second, if $A^\alpha = A^{\alpha+1}$ then A^α is perfect. In such a case, say that A *stabilizes by* α. For α least such this happens we say that A *stabilzes at* α.

Claim 80. $X \setminus X'$ is at most countable. In particular, $A^{\alpha+1} \setminus A^\alpha$ is at most countable.

Note that if $x \in X \setminus X'$ then there is an open neighborhood $U \ni x$ so that $U \cap (X \setminus \{x\}) = \emptyset$. So there is a family $\mathcal{U} = \{U\}$ of pairwise disjoint nonempty open sets so that each $x \in X \setminus X'$ is in exactly one $U \in \mathcal{U}$. It follows from the fact that $\mathbb{R}$ has the ccc (= *countable chain condition*, i.e. any collection of pairwise disjoint nonempty open sets must be countable) that $\mathcal{U}$ must be countable, and thus $X \setminus X'$ is countable.

Exercise 81. Show that any separable topological space has the ccc. In particular, $\mathbb{R}$ has the ccc. Show that a metric space has the ccc iff it is separable. Find a (necessarily non-metrizable) topological space which has the ccc but is not separable.

Claim 82. No matter what closed A we start with, it always stabilizes by some countable stage.

Suppose A stabilizes at α. Then $\langle \mathbb{R} \setminus A^\beta : \beta < \alpha \rangle$ is a strictly increasing sequence of open sets. Because $\mathbb{R}$ is separable, we thereby conclude that this sequence is countable. So α is countable, as desired.

These two claims together give us the theorem. If A stabilizes by countable α then $A = A^\alpha \cup (A \setminus A^\alpha)$. The former is perfect while the latter is a countable union of countable sets, hence countable. $\square$

Exercise 83. Show that for all $\alpha < \omega_1$ there is a closed set A so that A stabilizes at exactly α.

Now for a second application. Recall the following definitions.

Definition 84. A *sigma algebra* on a set X is a nontrivial collection of subsets of X closed under countable unions/intersections and complement.

Given a topological space X the *Borel sigma algebra* on X is the smallest sigma algebra on X which contains all the open sets.

This definition is okay, I guess. But it doesn't give you a handle on what the Borel sets are like. Alternatively, we can define the Borel sets as those appearing in a certain hierarchy. We will work in the context of perfect Polish spaces, as there is a nice theory in this context.

Definition 85. A *Polish space* is a complete, separable metric space. It is *perfect* if it has no isolated points.

A classic example of a perfect Polish space is the real numbers $\mathbb{R}$ under the standard topology. Another example is *Borel space* ${}^\omega\omega$. Points are ω-length sequences of finite ordinals. And the basic open sets are $N_s = \{x \in {}^\omega\omega : s \text{ is a subsequence of } x\}$, for finite length sequences of finite ordinals. Rounding out the trio of well-known perfect Polish spaces is *Cantor space* ${}^\omega 2$. Points are ω-length binary sequences and the basic open sets are $N_s = \{x \in {}^\omega 2 : s \text{ is a subsequence of } x\}$, for finite length binary sequences. (Note that Cantor space can be identified with $\mathcal{P}(\omega)$.)

Following usual set theoretic practice, we will refer to elements of Borel space or Cantor space as *reals*. That is, depending on context a real is an infinite binary sequence, an infinite sequence of natural numbers, or a set of natural numbers. When studying the Borel structure of the reals, it is equivalent to study the Borel structure of any perfect Polish space. This is due to the important theorem, which we will not prove, that between any two uncountable Polish spaces is a Borel isomorphism, a bijection which preserves the Borel structure. And Borel space and Cantor give us a nicer combinatorial handle on things. (For a basic example of this, contrast how easy it is to prove Cantor space is uncountable whereas showing that $\mathbb{R}$ is uncountable requires mucking about in unpleasant details.)

Exercise 86. Show that Borel space and Cantor space are perfect Polish spaces.

Definition 87. Let X be a Polish space. Inductively along the countable ordinals define the following hierarchy, the *Borel hierarchy on* X.

- The Σ^0_0 sets are the open sets;
- The Π^0_0 sets are the closed sets;
- The $\Sigma^0_{\alpha+1}$ sets are the countable unions of Π^0_α sets.
- The $\Pi^0_{\alpha+1}$ sets are the complements of $\Sigma^0_{\alpha+1}$ sets. Equivalently, they are the countable intersections of Σ^0_α sets.
- For λ limit, the Σ^0_λ sets are the countable unions of Π^0_α sets for $\alpha < \lambda$.
- And the Π^0_λ sets are the complements of Σ^0_λ sets.
- The Δ^0_α sets are those which are both Σ^0_α and Π^0_α.

Remark 88. To be precise, this is actually the *boldface* Borel hierarchy.

Theorem 89. *The Borel sets on X are precisely the sets which are Σ^0_α (or Π^0_α) for some countable α.*

Proof. For one direction of the inclusion, take a countable collection $\{A_i : i \in \omega\}$ of sets, each of which is $\Sigma^0_{\alpha_i}$ for some countable α_i. Let $\alpha = \sup_i \alpha_i$, which is countable. Then $\bigcup_i A_i$ is Σ^0_α. And it is clear that this collection is closed under complement and contains the open sets, so it's a sigma algebra. Thus it must contain the Borel sets.

For the other direction of the inclusion, inductively show that the Borel sigma algebra must contain the Σ^0_α sets for all countable α. $\square$

Corollary 90. *Let $\mathcal{B}$ be the collection of Borel sets on a perfect Polish space X. Then $|\mathcal{B}| = 2^{\aleph_0}$.*

Proof. Because X is a perfect Polish space, there are $2^{\aleph_0}$ many open sets on X. Now inductively show that there are $2^{\aleph_0}$ many Σ^0_α sets on X for each countable α: at successor stages, there are at most $\aleph_0 \cdot 2^{\aleph_0} = 2^{\aleph_0}$ many Σ^0_α sets. And similarly at limit stages there are at most $\aleph_0 \cdot 2^{\aleph_0}$ many Σ^0_λ sets. But then there are $\aleph_1 \cdot 2^{\aleph_0} = 2^{\aleph_0}$ many Borel sets. $\square$

This is just a hint of *descriptive set theory*, the structure theory of the Borel sets of Polish spaces. This is an active and exciting area of study with connections to many different areas of mathematics. It is ripe with possible topics for projects.

4. Axioms of set theory

We now begin the formal set-up. We will have one primitive notion, namely set, and one primitive relation on sets, namely the membership relation $\in$. The axioms of Zermelo–Fraenkel set theory ZF describe the basic properties enjoyed by sets. The first axiom, **Extensionality**, defines equality on sets. Namely, two sets are equal if and only if they have the same elements. Formally:

$$\text{(Extensionality)} \qquad \forall x, y \; x = y \Leftrightarrow (\forall z \; z \in x \Leftrightarrow z \in y)$$

The axioms of **Pairing**, **Union**, and **Powerset** describe ways new sets can be built up from old sets. Pairing states that unordered pairs can be formed, Union states that unions can be formed, and Powerset states that powersets can be formed.

$$\text{(Pairing)} \qquad \forall x, y \exists z \; z = \{x, y\}$$

As you can see, we have already gone beyond our one primitive relation. For the sake of human readability, we will use many, many derived notions, usually without bothering to state them in our minimalist language. For example, $z = \{x, y\}$ is an abbreviation for $\forall w \; w \in z \Leftrightarrow (w = x \vee w = y)$.

$$\text{(Union)} \qquad \forall x \exists y \; y = \bigcup x$$

Here, $\bigcup x = \{a : \exists b \; a \in b \in x\}$ is the union of x.

$$\text{(Powerset)} \qquad \forall x \exists y \; (\forall z \; z \in y \Leftrightarrow z \subseteq x)$$

For any x, such y is called the *powerset* of x, which we write as $\mathcal{P}(x)$. The powerset of x is unique by the axiom of Extensionality.

Exercise 91. Write formulae, in the language with only $=$ and $\in$, which express:

- $x = \{a : \exists b \; a \in b \in x\}$;
- $y = \{z \in x : \varphi(z, a)\}$;
- $x = \{y : \varphi(y, a)\}$;
- $z \subseteq x$;
- $z = x \cup y$; and
- $x = \emptyset$.

Exercise 92. Show that Extensionality plus Union plus Pairing prove that $x \cup y$ exists for all sets x and y.

The axiom of **Separation** intuitively states that the universe of sets is closed under taking subsets. Formally, Separation is an infinite schema of axioms, each one stating that the subset of x consisting of those elements which satisfy φ always exists.

$$\text{(Separation)} \qquad \forall \bar{p} \; [\forall x \; \exists y \; y = \{z \in x : \varphi(z, \bar{p})\}]$$

Here, φ ranges over those formulae in the first-order language with $\in$ as the only non-logical symbol. (Recall that $=$ is considered a logical symbol; but even if we did not include it, it is definable from $\in$.)

Exercise 93. Show that Unrestricted Comprehension, the axiom schema with instances

$$\forall \bar{p} \; [\exists x \; x = \{y : \varphi(y, \bar{p})\}]$$

is inconsistent with Extensionality.

This exercise shows that in general $\{y : \varphi(y, \bar{p})\}$ is not a set. Nevertheless, it is often useful to speak of such collections. Our approach will be to treat the usage of these collections, known as *classes* as abbreviations for certain formula. That is, if $A = \{y : \varphi(y)\}$ then $x \in A$ is an abbreviation for $\varphi(x)$. For example,

$$\mathrm{Ord} = \{x : \forall y, y' \in x\ (y \in y' \vee y = y' \vee y' \in y) \wedge \forall y \in x\ \forall z \in y\ z \in x\}$$

is the class of all *ordinals*. So $\alpha \in \mathrm{Ord}$ is an abbreviation for the formula asserting that α is an ordinal.

Definition 94. $V = \{x : x = x\}$ is the class of all sets.

There is an alternative approach to classes, which goes by the names of class theory or second-order set theory, where classes are admitted as actual objects. Sets are then the classes which are elements of other classes, while proper classes are those which aren't members of any other class. Axioms can be formulated for this approach, and this approach is consistent if and only if ZF is consistent. We won't take this approach in this course, but I'm happy to talk about it outside of class—my dissertation work was on second-order set theory—and there are several good topics here for a project.

Project Idea 95. Gödels–Bernays set theory GB is conservative over ZF. Indeed, if we augment GB with the axiom of Global Choice to get GBC then GBC is conservative over ZFC, viz. ZF plus the axiom of Choice. But GB and GBC aren't the only second-order set theories which have been studied. Kelley–Morse set theory KM is not conservative over ZFC. And there is a hierarchy of theories intermediate in consistency strength.

The axiom of **Infinity** states that the natural numbers form a set. One way to characterize the natural numbers is that they are the smallest collection which contains 0 and is closed under successor. Since our only objects are sets, we must pick a set to be 0 and pick a function on sets to be the successor function. There are many possible choices, but the standard pick is to have 0 be the empty set and the successor of n to be $n \cup \{n\}$.

$$\exists x\ \emptyset \in x \wedge \forall y \in x\ (y \cup \{y\}) \in x \tag{Infinity}$$

Note that this is our first axiom which asserts the existence of a certain set, rather than asserting that such and such holds of all sets. Indeed, Infinity is the only axiom of ZFC which starts with an existential quantifier rather than a universal quantifier.

After we have defined what it means for a set to be infinite, we will see that the set whose existence Infinity asserts really is infinite.

The axioms we have seen so far—Extensionality, Pairing, Union, Powerset, Separation, plus Infinity—constitute Zermelo set theory Z.

Before we can state the next two axioms, we will need to be able to talk about functions and relations.

Definition 96 (Kuratowski)**.** The *ordered pair* (x, y) is the set $\{\{x\}, \{x, y\}\}$.

Exercise 97. Show that $(a, b) = (c, d)$ if and only if $a = b$ and $c = d$.

Definition 98. The *cartesian product* $x \times y$ of x and y is the set $\{(a, b) : a \in x \wedge a \in y\}$.

Exercise 99. Show, using the axioms we have seen so far, that $x \times y$ is always a set if x and y are sets.

Definition 100. A *(binary) relation* is a set of ordered pairs. We say R is a relation on A if $R \subseteq A \times A$. More generally, R is a relation between A and B if $R \subseteq A \times B$. The domain of R is $\operatorname{dom} R = \{a : \exists b \in \bigcup\bigcup R \ (a,b) \in R\}$ and the *range* of R is $\operatorname{ran} R = \{b : \exists a \in \bigcup\bigcup R \ (a,b) \in R\}$

A *function* f is a relation so that $(a,b),(a,b') \in f$ implies $b = b'$. For $a \in \operatorname{dom} f$ we write $f(a)$ for the unique b so that $(a,b) \in f$. We write $f : A \to B$ to express that $\operatorname{dom} f = A$ and $\operatorname{ran} f \subseteq B$.

Exercise 101. Suppose f, g are functions. Show that $f = g$ (as sets) if and only if f and g have the same domain and for every a in their domain we have $f(a) = g(a)$.

Exercise 102. Those of an algebraic bent often dislike this definition of a function, as it does not require a function to come along with a codomain. Give a alternative definition of a function with codomains, so that functions f and g are equal (as sets) if and only if they have the same domain and codomain and $f(a) = g(a)$ for all a in their domain.

One important class of functions are the sequences.

Definition 103. A *sequence* is a function whose domain is an ordinal. The length of the sequence is its domain. We often write $\langle x_i : i \in \alpha \rangle$ for the sequence $i \mapsto x_i$ with domain α.

From our definition of relations we can define partial orders, linear orders, well-orders, and so on in the usual way.

The axiom of **Choice** asserts that every set can be well-ordered.

$$\forall x \exists y \ y \text{ is a well-order on } x \tag{Choice}$$

Theorem 104 (Zermelo)**.** *Over Zermelo set theory, Choice is equivalent to the assertion that every nonempty set admits a choice function: for every $x \neq \emptyset$ there there is a function f so that for all $y \in x$ we have $f(y) \in y$.*

Exercise 105. Prove Zermelo's theorem.

Project Idea 106. Several other statements are also equivalent to Choice. Investigate some of them.

Project Idea 107. There is a zoo of choice principles which are weaker than the full axiom of Choice but are not implied by ZF. For example: countable choice, the principle of dependent choices, the boolean prime ideal theorem. Investigate some of these principles.

Several mathematicians independently observed that Zermelo set theory, even with the addition of Choice, is insufficient for set theoretic purposes. We are not yet in a position to see what those insufficiencies are, but we can see the axioms that provides the fix. **Replacement** asserts that the image of a set under a definable class function is a set. Like Separation, Replacement is actually a schema of assertions, one for each formula φ of appropriate arity. Below, $\exists! y$ expresses that there is a unique y such that blah blah.

$$\forall \bar{p} \ [(\forall x \exists! y \ \varphi(x,y,\bar{p})) \Rightarrow \forall a \exists b \ \forall x \in a \exists y \in b \ \varphi(x,y,\bar{p})] \tag{Replacement}$$

The utility of Replacement will become increasingly clear as this course progresses.

You can also phrase Replacement using the language of classes. In this formulation, Replacement is a schema over the class functions F asserting

$$\forall a \exists b \ \forall x \in a \ F(x) \in b.$$

The final axiom of ZFC, the axiom of **Foundation**, gives a restriction on which sets can exist. Abraham Fraenkel proposed that there should be an axiom of restriction—"*das Axiom der Beschränktheit*"—which asserts that the only sets which exist are those which must be in every model of the axioms. Those of you with a background in logic will know that the Löwenheim–Skolem theorem rules out the possibility of such an axiom. Nevertheless, a weak version of this axiom, originally due to John von Neumann, can be expressed. Foundation asserts that the membership relation is well-founded. This rules out, for example, the existence of a Quine atom, viz. a set a so that $a = \{a\}$.

(Foundation) $$\forall x \ (x \neq \emptyset \Rightarrow \exists y \in x \ y \cap x = \emptyset)$$

Together, these axioms make up ZFC. If we don't include Choice, we get ZF.

4.1. **Believing the axioms.** I won't spend any class time attempting to justify these axioms to you, though we will see why they are useful as the course progresses. You're welcome to accept them as stipulations if you are so inclined. But if you're like me and you think mathematics has meaning and isn't mere symbol pushing, then you'll want to know why you should believe the axioms. Penelope Maddy has an excellent paper about that question, appropriately titled "Believing the axioms, Part I". (Part II is also good, but is focused on axioms which go beyond ZFC.) I'm also happy to recommend further papers or talk to you outside of class. But my advice would be to first focus on understanding the mathematics, before delving into the philosophical side.

5. Ordinals, cardinals, and transfinite recursion: take two

We return to these topics to give them a formal basis in ZFC.

Definition 108. A set x is *transitive* if every element of x is a subset of x. That is, x is transitive if $\forall y \in x \forall z \in y\ z \in x$.

Definition 109 (von Neumann)**.** An *ordinal* is a transitive set which is linearly ordered by $\in$. The order relation $<$ on the ordinals is then $\in$.

Observe that, due to Foundation, a set is linearly ordered by $\in$ if and only if it is well-ordered by $\in$. So the von Neumann definition of an ordinal picks out a canonical representative from each ordertype for a well-order. As such, all our previously proven results about ordinals immediately apply to this new definition.[3]

But first we have to see that every well-order is isomorphic to some (von Neumann) ordinal. This follows from the Mostowski collapse lemma. But before we can prove it we must revisit transfinite recursion and put it on a firm foundation.

Definition 110. A class relation R is *set-like* if for every x the collection $\{y : y\,R\,x\}$ of predecessors of x is a set.

For example, the class of ordinals, under its usual order, is set-like. There are a proper class of ordinals, but beneath any particular ordinal there is only a set of ordinals. Compare to the natural numbers—there are infinitely many natural numbers, but each natural number has only finitely many predecessors.

Theorem Schema 111 (Transfinite Recursion)**.** *Suppose that G is a class function and R is a well-founded, set-like class relation. Then there is a class function F so that for each $x \in \operatorname{dom} R$ we have $F(x) = G(F \upharpoonright x)$, where $F \upharpoonright x$ is the restriction of F to the set of R-predecessors of x.*

Before proving this, I want to remark on the formal underpinning. Formally, this is not a single theorem but rather a schema for producing theorems. If we take formulae γ and ρ defining appropriate classes G and R then apply the argument we get formula φ which defines the class F. The reason for needing this to be a schema is that classes are metatheoretic objects—recall, they are just shorthand for certain formulae—so we cannot directly quantify over them. But it's cleaner and more intuitive to speak about classes rather than formulae. A rule of thumb: it's okay to talk about individual classes, whether via a schema or a specific class such as the class of ordinals or the class of groups. But it's not okay to quantify over classes. (At least in the framework of ZFC. There are alternative frameworks which do allow this.)

Also, note that it is harmless to assume that the domain of G is V, the class of all sets. Often, we are only interested in G being defined on certain sets, e.g. on the ordinals. But we can always extend the definition of G to cover all sets. Clearly, the definition of F is unaffected by what G does on the side.

[3]Why use this definition? First off, it's pretty. Second, this gives us a logically simple definition of an ordinal. Cantor's definition requires quantifying over the whole universe of sets—to say such and such is well-founded is an unbounded universal quantifier. But von Neumann's definition only requires quantifying over x itself. This technical fact is of supreme importance to set theoretic practice. In particular, we will eventually use it to see that being well-founded is absolute—whether a relation is well-founded doesn't depend on what universe you evaluate the sentence in. On the other hand, we will see that, for example, being a cardinal is not absolute; different universes of sets can disagree about what sets are cardinals.

We are still a long way off from this, however.

Moreover, I want to highlight that this schema is often applied in the case where G and R are sets. This follows from the version stated above by extending G arbitrarily outside of the set of interest.

Proof. We define F as follows: $F(x) = y$ if there is a set-sized function f whose domain includes $\{z : z \leq_R x\}$, where $<_R$ is the transitive closure of R which obeys the recursive definition of F on its domain. (That is, $f(x) = G(f \restriction x)$ for $x \in \operatorname{dom} f$.) We must see two things: that $F(x)$ is defined for all $x \in \operatorname{dom} R$ and that $F(x)$ is uniquely determined.

First, suppose it were the case that $F(x)$ were not defined for all $x \in \operatorname{dom} R$. Then there would be a minimal $x \in \operatorname{dom} R$ so that $F(x)$ is not defined. In particular, for each $z <_R x$ we have that $F(x)$ is defined. Since R is set-like, by Replacement there is a set-sized function f so that for all $z <_R x$ we have $f(z) = F(z)$ satisfies the recursive definition for F. But then we can extend f by defining $f(x) = G(f \restriction x)$. But then $F(x)$ is defined, a contradiction.

To see that $F(x)$ is uniquely determined, suppose x were a minimal counterexample to uniqueness and produce a contradiction, similar to before. (Exercise: do it!) $\square$

Notice that we used Replacement in order to define F. It is natural to ask whether we actually need Replacement to prove transfinite recursion.

Exercise 112. Show that over the axioms of ZFC without Replacement that the schema of Replacement is equivalent to the schema of transfinite recursion. (Hint: consider an instance of Replacement, i.e. you have a set a and a class function F. You want to use transfinite recursion to define a set function f so that $f = F \restriction a$. You need a well-founded relation to do recursion on, so start by well-ordering a.)

As an example of the use of transfinite recursion, let's prove the following very important lemma.

Theorem 113 (Mostowski collapse lemma)**.** *Let R be a well-founded relation on a set D. Then there is a unique transitive set T and an epimorphism $e : (D, R) \to (T, \in)$. We call T the* collapse *of R and e the* collapse map*. If R is moreover extensional—if $x, y \in D$ have the same set of R-predecessors then $x = y$—then e is moreover an isomorphism. In particular, if R is a well-order then it collapses to a unique ordinal.*

This also works if R is a set-like well-founded class relation on a class D.

Proof. Recursively define e according to the rule $e(x) = \{e(y) : y \, R \, x\}$. Then e exists by transfinite recursion. By construction $y \, R \, x$ implies $e(y) \in e(x)$. And if R is extensional we can reverse the implication. $\square$

To end out this section, let us give a new definition of cardinals. We will define them as certain ordinals.

Definition 114. A *cardinal* is an initial ordinal. That is, an ordinal α is a cardinal if there is no injection from α to a smaller ordinal.

Definition 115. Let α be an ordinal. Then $\aleph_\alpha = \omega_\alpha$.

Observe that our definition of ω_α immediately carries over from the Cantorian context. For example, ω_1 is the least (von Neumann) ordinal which is uncountable. Equivalently, ω_1 is the set of countable ordinals.

There is one infelicity in defining the cardinals as certain ordinals. Namely, there is now an ambiguity in the arithmetic operations. When we say $\alpha + \beta$, do we mean ordinal addition or cardinal addition? Following standard practice, we will not make notational distinction between the two. Which is meant should be clear from context. And in case it is not clear, we will explicitly say what is meant.

One way to distinguish the two for initial ordinals is by using $\aleph_\alpha$ when we want to think of the set as a cardinal, and using ω_α when we want to think of it as an ordinal. But this convention is not universally followed, and many set theorists will use ω_α for both uses.

6. The cumulative hierarchy

We define by transfinite recursion on the ordinals the *von Neumann hierarchy*:

$$\begin{aligned} V_0 &= \emptyset; \\ V_{\alpha+1} &= \mathcal{P}(V_\alpha); \\ V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha. \end{aligned}$$

Proposition 116. *Each V_α is transitive.*

Proof. We prove this by transfinite induction. It is vacuously true for V_0. Suppose V_α is transitive and take $x \in V_{\alpha+1}$. Then each element of x is in V_α, by definition. Because V_α is transitive we get that each element of x is a subset of V_α, i.e. each element of x is an element of $V_{\alpha+1}$. And so $x \subseteq V_{\alpha+1}$. Finally, take $x \in V_\lambda$ for limit λ, where we suppose V_α is transitive for $\alpha < \lambda$. Then there is $\alpha < \lambda$ so that $x \in V_\alpha$. But then $x \subseteq V_\alpha \subseteq V_\lambda$, and we are done. $\square$

Corollary 117. *If $\alpha < \beta$ then $V_\alpha \subseteq V_\beta$.*

Proof. It's enough to see that $V_\alpha \subseteq V_{\alpha+1}$, since the general result follows from an easy induction. But this is an immediate consequence of V_α being transitive: every element of V_α is also an element of the powerset of V_α. $\square$

In general, if t is transitive then $\mathcal{P}(t)$ is also transitive.

Theorem 118. $V = \bigcup_{\alpha\in\mathrm{Ord}} V_\alpha$. *That is, every set appears as an element of some member of the von Neumann hierarchy.*

Proof. We prove this by induction on $\in$. Since $\in$ is well-founded, by Foundation, and set-like, this is valid. Consider a set x and suppose that for every $y \in x$ we have that $y \in V_{\alpha_y}$. By Replacement, we can consider the set $\{\alpha_y : y \in x\}$. Set $\alpha = \sup_{y\in x} \alpha_y$. Then $y \in V_\alpha$ for each $y \in x$. But then $x \subseteq V_\alpha$, so $x \in V_{\alpha+1}$. $\square$

Definition 119. The *rank* of a set x, denoted $\operatorname{rank} x$, is defined recursively as:

$$\operatorname{rank} x = \sup\{\operatorname{rank} y + 1 : y \in x\}.$$

Here, we make the convention that $\sup\emptyset = 0$ and so $\operatorname{rank}\emptyset = 0$.

Exercise 120. Show that $\operatorname{rank}\alpha = \alpha$.

Exercise 121. Show that $\operatorname{rank} x = \alpha$ iff $V_{\alpha+1}$ is the least place in the von Neumann hierarchy which has x as an element.

Combined with the previous exercise, this shows that α first appears in $V_{\alpha+1}$. In other words, $\alpha \subseteq V_\alpha$ but $\alpha + 1 \not\subseteq V_\alpha$. So $\alpha = \mathrm{Ord} \cap V_\alpha$.

We can also use the von Neumann hierarchy to show that adding Foundation to the axioms of ZFC doesn't raise the consistency strength of the axioms.

Theorem 122. *Inside any model of* ZFC − *Foundation the von Neumann hierarchy* $\bigcup_{\alpha\in\mathrm{Ord}} V_\alpha$ *is a model of* ZFC.

Corollary 123. ZFC *and* ZFC − *Foundation are equiconsistent.*

Proof. Clearly if ZFC is consistent then so is ZFC − Foundation. For the other direction, suppose ZFC − Foundation is consistent. Then, by Gödel's completeness theorem, there is a model M of ZFC − Foundation. By the theorem, we can define inside M a model of ZFC. So ZFC has a model, and is thus consistent. □

Proof of theorem. Work in ZFC − Foundation. We want to see that $\bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$ satisfies all the axioms of ZFC. Extensionality is free, since it's true in the ambient universe. Pairing and Union are exercises.

Exercise: Show that if $x, y \in V_\alpha$ then $\{x, y\} \in V_{\alpha+1}$. Show that if $x \in V_\alpha$ then $\bigcup x \in V_\alpha$. (Indeed, if $x \in V_{\alpha+1}$ then $\bigcup x \in V_\alpha$.)

Infinity holds because $\omega \in V_{\omega+1}$. Choice holds because Choice holds in the ambient universe and any binary relation, including a well-order, on a set $x \in V_\alpha$ is in $V_{\alpha+2}$. (Exercise: check this!) Separation holds because the von Neumann hierarchy is closed under subset. Powerset holds because if $x \in V_\alpha$ then $\mathcal{P}(x) \in V_{\alpha+1}$. Foundation holds by essentially our argument in ZFC that $V = \bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$. Finally, Replacement holds because if F is a class function taking sets from $\bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$ to sets in $\bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$, then the image of any set under F is also in $\bigcup_{\alpha \in \mathrm{Ord}} V_\alpha$. □

A few remarks on this proof. First, it was kinda exhausting to check all those axioms. But it could be much worse. This is where a spartan axiomatization is nice; the more axioms we use, the more work it is to check these sorts of things. And if we had more primitative notions then just 'set' and 'membership', we'd have to check even more. Note, however, that we can lessen the work by checking some basic cases in a general fashion; see the exercise below.

Second, this is the first of a so-called inner model construction. Namely, we start from the whole universe of sets, assumed to satisfy such and such axioms. Then we consider a restriction of the universe to only certain sets, and show that this inner model satisfies more than what we started out assuming. This gives us a way to prove relative consistency results, similar to our proof that ZFC is consistent iff ZFC − Foundation is. For example, Gödel constructed the inner model L, the smallest inner model containing all the ordinals, to show that $\mathsf{ZFC} + \mathsf{GCH}$ is equiconsistent with ZF. (But we won't have time in this course to delve into inner model theory.)

Exercise 124. In this exercise you will show that levels of the von Neumann hierarchy satisfy lots of the axioms of ZFC. For all of the below, you want to consider V_α as a structure with the restriction of the membership relation. This structure has appropriate signature to ask whether it satisfies the axioms of set theory. (Hint: most of these are one line arguments!)

- Show that each V_α satisfies Extensionality.
- Show that each V_α satisfies Union.
- Show that if α is a limit ordinal then V_α satisfies Pairing.
- Show that each V_α satisfies Separation.
- Show that V_α satisfies Infinity if $\alpha \geq \omega + 1$.
- Show that each V_α satisfies Foundation.
- Show that if α is a limit ordinal then V_α satisfies Choice.
- Show that if α is a limit ordinal then V_α satisfies Powerset.

In short, if α is a limit ordinal $> \omega$ then V_α satisfies all the axioms of ZFC except possibly Replacement.

Note that as a corollary of this exercise we can deduce that ZFC proves the consistency of ZFC − Replacement. This is because ZFC proves that $V_{\omega+\omega}$ exists and that it satisfies each axiom of ZFC − Replacement. So unlike with Foundation, adding Replacement ups the consistency strength of our theory. We can also conclude from this that we need Replacement to prove the existence of cardinals $\geq \aleph_\omega$.

Let's take this moment to see natural models of a different weakening of ZFC, where we exclude Powerset instead of Foundation.

Definition 125. Let κ be an infinite cardinal. Then H_κ is the collection of all sets hereditarily of cardinality $< \kappa$, where x is hereditarily of cardinality $< \kappa$ if the transitive closure of x has cardinality $< \kappa$.

H_ω is also referred to as the collection of *hereditarily finite* sets and H_{ω_1} is also referred to as the collection of *hereditarily countable* sets.

Exercise 126. Let κ be an infinite cardinal. Show that x is hereditarily of cardinality $< \kappa$ iff $|x| < \kappa$ and each element of x is hereditarily of cardinality $< \kappa$. In other words, we could equivalently define this notion by induction on $\in$.

Exercise 127.

- Show that H_{ω_1} satisfies all axioms of ZFC except for Powerset.
- Let κ be an uncountable cardinal. Show that H_κ satisfies all axioms of ZFC except Powerset and possibly Replacement.

We will see soon that if κ is *regular* then H_κ also satisfies Replacement.

Definition 128. If t is a transitive set, φ is a formula in the language of set theory, and $\bar{a}$ are elements of t, we write $t \models \varphi(\bar{a})$ if $\varphi(\bar{a})$ is true in t, with the symbol for the membership relation being interpreted as $\in$ restricted to t.[4]

More formally, this *satisfaction relation* is defined according to the following Tarskian recursion:

(1) $t \models a = b$ iff $a = b$;
(2) $t \models a \in b$ iff $a \in b$;
(3) $t \models \varphi \wedge \psi$ iff $t \models \varphi$ and $t \models \psi$;
(4) $t \models \varphi \vee \psi$ iff $t \models \varphi$ or $t \models \psi$;
(5) $t \models \neg\varphi$ iff it's not the case that $t \models \varphi$;
(6) $t \models \forall x \varphi(x)$ iff for all $a \in t$ we have $t \models \varphi(a)$;
(7) $t \models \exists x \varphi(x)$ iff there is $a \in t$ so that $t \models \varphi(a)$;

It follows from transfinite recursion that this relation is well and uniquely defined for any transitive set t, and so this really defines a relation between finite sequences of elements from t and formulae in the language of set theory.

Exercise 129. Prove Tarski's theorem on the undefinability of truth: the class relation $V \models \varphi(\bar{a})$ is not definable. Why can we not use the principle of transfinite recursion to prove that such a class is definable? After all, each step in the recursion is itself definable.

An important consequence of the existence of the von Neumann hierarchy is the principle of Reflection.

Theorem Schema 130 (Lévy–Montague Reflection Principle)**.** *Let $\varphi(\bar{x}, \bar{y})$ be a formula in the language of set theory, and let $\bar{a}$ be sets. Then, there is an ordinal α with $\bar{a} \in V_\alpha$ so that for all $\bar{y} \in V_\alpha$ we have $V_\alpha \models \varphi(\bar{a}, \bar{y})$ iff $\varphi(\bar{a}, \bar{y})$ is true in the whole universe of sets.*

[4]For those with a background in logic, this is the usual model theoretic satisfaction relation for the structure $(t, \in)$.

Proof. Without loss we may assume that $\forall$ does not appear in φ, by replacing any $\forall$ with $\neg\exists\neg$.

For a formula φ and a transitive set t, say that a transitive set u is *closed under existential subformulae of φ with parameters from t* if for any subformula of φ of the form $\exists\bar{x}\psi(\bar{x},\bar{b})$, with $\bar{b} \in t$, if there are any $\bar{x}$ so that $\psi(\bar{x},\bar{b})$ hold, then there are such $\bar{x}$ in u. In brief, u is large enough to see witnesses for all true existential subformulae of φ.

We now recursively define a sequence $\langle \alpha_i : i \in \omega \rangle$ of ordinals as follows. Start off with α_0 least so that $\bar{a} \in V_{\alpha_0}$. Now given α_i, let α_{i+1} be least so that $V_{\alpha_{i+1}}$ is closed under existential subformulae of φ with parameters from V_{α_i}. Finally, set $\alpha = \sup_i \alpha_i$.

Let's see that $V_\alpha \models \varphi(\bar{a},\bar{y})$ iff $\varphi(\bar{a},\bar{y})$. We prove this by induction on formulae, working on subformulae of φ. I will do the argument in full since this is the first time in this class we see an induction on formulae.

Start with atomic formulae. By the definition of the satisfaction relation, $V_\alpha \models a = b$ iff $a = b$, and similarly for $a \in b$. Now assume $V_\alpha \models \psi$ and $V_\alpha \models \theta$ iff they hold, respectively, in V. Then $V_\alpha \models \psi \wedge \theta$ iff $V_\alpha \models \psi$ and $V_\alpha \models \theta$ iff ψ and θ (by inductive hypothesis) iff $\psi \wedge \theta$. A similar argument works for negation and disjunction.

By our assumption on φ, we only have to consider existential quantifiers. So assume $V_\alpha \models \psi(\bar{a},\bar{y})$ iff $\psi(\bar{a},\bar{y})$ for all $\bar{y} \in V_\alpha$. If $V_\alpha \models \exists x\, \psi(\bar{a},x,\bar{y})$ then there is such a witness $b \in V_\alpha$, but then $\psi(\bar{a},b,\bar{y})$ holds by inductive hypothesis. So $\exists x\, \psi(\bar{a},x,\bar{y})$ holds. For the other direction, suppose $\exists x\, \psi(\bar{a},x,\bar{y})$ holds. Then there is i so that $\bar{a},\bar{y} \in V_{\alpha_i}$. By construction there is therefore $b \in V_{\alpha_{i+1}} \subseteq V_\alpha$ so that $\psi(\bar{a},b,\bar{y})$. But by inductive hypothesis, we then get that $V_\alpha \models \psi(\bar{a},b,\bar{y})$, and so $V_\alpha \models \exists x\, \psi(\bar{a},x,\bar{y})$. This is the last step in the induction, and we are done. □

I want to remark that formally, this induction took place in the metatheory. φ has only finitely many subformulae, so really what I gave was a recipe for producing a formula proof of this for each φ, based upon expanding it out into its subformulae and proving them first.

Here is a hella important corollary to reflection.

Corollary 131 (Corollary Schema)**.** *If T is a finite collection of axioms of* ZFC*, then* ZFC *proves the consistency of T. In particular,* ZFC *(if consistent) cannot be axiomatized by finitely many axioms.*

Proof. Let φ be the conjunction of the axioms in T. By reflection, there is α so that $V_\alpha \models \varphi$ iff φ. But φ holds, because they are axioms. So $V_\alpha \models \varphi$. Since T has a model—V_α—it must be consistent, by the Soundness Theorem. □

Exercise 132. Convincingly explain to yourself why this does not contradict Gödel's second incompleteness, which implies that ZFC does not prove the consistency of ZFC.

I want to end this section by discussing a certain view as to what sets are, the so-called cumulative hierarchy view. I will give a sketch of the view, handwaving over some philosophical disagreements.

Under this view, sets are built up in transfinite stages. You start with the empty set, whose existence is uncontroversial, and you go up from there. As you go up a set, the new sets are all collections of old sets. In other words, to go from stage α to stage $\alpha + 1$ you take the powerset of the previous stage. This process is iterated along the ordinals, with the sets being the objects that appear at some stage in the process. It's clear that the stages in this construction are just the V_αs, since we start with $V_0 = \emptyset$ and take powersets at each successor. The statement that this process exhausts the sets is just the theorem $V = \bigcup_\alpha V_\alpha$.

Missing so far is the important question: how high? This view seems to preassume the totality of ordinals, which is quite a tall order. So we must say something about how many ordinals there are. One popular answer, with antecedents back to Cantor, is that the totality of ordinals should be ineffable—there should be no distinguishing property of Ord that separates it from an initial segment. So, for example, the universe cannot be $V_{\omega+\omega}$ because then the class of ordinals is describable—it is just $\omega + \omega$.

This ineffability is formalized through reflection principles, which assert that there are lots of αs so that V_α resembles V. We have already seen the most basic of these, Lévy–Montague Reflection which is reflection for properties expressible by a first-order property. Stronger reflection principles motivate strong axioms of infinity, which go beyond ZFC in strength. For example, this is one way to motivate *inaccessible cardinals*, which we will see soonish.

ZFC can be seen as an attempt to axiomatize this view of sets, similar to how Peano arithmetic is an attempt to axiomatize a view of what natural numbers are. Note that we can be confidant that this view avoids the paradoxes of size—Russell's paradox, Burali–Forti's paradox, and so forth—without needing to write down any axioms. For example, there is no set of all ordinals (and hence the Burali–Forti paradox is avoided) because there is no stage at which we have collected all of the ordinals.

Exercise 133 (Open-ended, kinda philosophical)*.* Which axioms of ZFC can be argued for from this view? Are there any which aren't settled by the cumulative hierarchy view?

Exercise 134*.* Show that Replacement is equivalent to the Reflection schema, over the other axioms of ZFC. (Hint: the only remaining step is to show that you can prove instances of Replacement from corresponding instances of Reflection.)

7. Small large cardinals

To round out part 0 of this course, I want to discuss the small large cardinals, by which I mean those which go beyond what you've likely seen in other classes, while still having their existence be provable from ZFC.[5] Let's start by dividing the infinite cardinals into two classes.

Definition 135. If $(L,<)$ is a linear order then the cofinality of L, written $\operatorname{cof} L$, is the least ordinal α so that there is an order embedding $e : \alpha \to L$ which is cofinal in L: for any $\ell \in L$ there is $\beta < \alpha$ so that $e(\beta) >_L \ell$.

For example, the cofinality of $(\mathbb{R},<)$ is ω. There is a countable cofinal sequence of reals, but no finite sequence can be cofinal. Also note that $\operatorname{cof} L$ is finite iff L has a maximum iff $\operatorname{cof} L = 1$. So this is really only interesting in case L lacks a maximum iff $\operatorname{cof} L$ is infinite.

We are mostly interested in the case where L is an ordinal.

Exercise 136. For any linear order $(L,<)$, we have that $\operatorname{cof} L$ is a cardinal.

In particular, this exercise shows that if $\operatorname{cof} \alpha = \alpha$ then α is a cardinal. (Here, we think of α as having its canonical order $\in$ attached.) So the following distinction wouldn't be meaningful applied to the non-cardinals.

Definition 137. An infinite cardinal κ is *regular* if $\operatorname{cof} \kappa = \kappa$. Otherwise, κ is called *singular*.

Observe that $\aleph_0$, $\aleph_1$, and in general $\aleph_n$ for $n \in \omega$ are all regular. The smallest singular cardinal is $\aleph_\omega$. It has cofinality ω, as witnessed by the sequence $\langle \aleph_i : i \in \omega \rangle$.

Proposition 138. *Let κ be an infinite cardinal. Then $\operatorname{cof}(\operatorname{cof} \kappa) = \operatorname{cof} \kappa$. In particular, $\operatorname{cof} \kappa$ is regular.*

Proof. Suppose $\langle \alpha_i : i \in \operatorname{cof}(\operatorname{cof} \kappa) \rangle$ is cofinal in $\operatorname{cof} \kappa$. Let the sequence $\langle \beta_i : i \in \operatorname{cof} \kappa \rangle$ witness that κ has cofinality at most $\operatorname{cof} \kappa$. Now consider the sequence $\langle \beta_{\alpha_i} : i \in \operatorname{cof}(\operatorname{cof} \kappa) \rangle$. This sequence is cofinal in κ. So $\operatorname{cof} \kappa \leq \operatorname{cof}(\operatorname{cof} \kappa)$. But since $\operatorname{cof} \lambda \leq \lambda$, we get $\operatorname{cof} \kappa = \operatorname{cof}(\operatorname{cof} \kappa)$. □

Definition 139. An infinite cardinal κ is a *successor cardinal* if $\kappa = \lambda^+$ for some cardinal λ. Equivalently, κ is a successor iff $\kappa = \aleph_{\alpha+1}$ for some α. Otherwise, κ is a *limit cardinal*, which happens iff $\kappa = \aleph_\gamma$ for a limit ordinal γ.

Proposition 140. *Every successor cardinal is regular.*

Proof. Suppose towards a contradiction that $\kappa = \lambda^+$ were singular. Then there would be a cofinal sequence $\langle \alpha_i : i \in \mu \rangle$ of ordinals in κ where $\mu < \kappa$. So $\kappa = \bigcup_{i \in \mu} \alpha_i$. Since $|\alpha_i| \leq \lambda$, we would thus get that

$$\kappa \leq \sum_{i\in\mu} |\alpha_i| = \sum_{i \in \mu} \lambda = \lambda \cdot \mu = \lambda,$$

a contradiction. □

It follows that any singular cardinal must be a limit cardinal.

Question 141. *Is there a regular limit cardinal?*

[5] Strictly speaking, this is not quite correct. Except for the cardinals below defined in terms of the cardinal exponentiation map $\kappa \mapsto 2^\kappa$, for each of them it is consistent with ZFC that they be smaller than $2^{\aleph_0}$. For example, it's consistent with ZFC that $2^{\aleph_0}$ is larger than the least aleph fixed point, defined below. We will prove these consistency assertions in part 2 of this course.

We will come back to this question in part 1.

Definition 142. A cardinal κ is an *aleph fixed point* if $\kappa = \aleph_\kappa$.

Proposition 143. *For any cardinal κ there is an aleph fixed point $> \kappa$.*

Proof. Recursively define an increasing sequence $\langle \lambda_i : i \in \omega \rangle$ of cardinals $\geq \kappa$. Start with $\lambda_0 = \kappa$. Given λ_i set $\lambda_{i+1} = \aleph_{\lambda_i}$. Set $\lambda = \bigcup_i \lambda_i$. Let us see that λ is an aleph fixed point. Suppose towards a contradiction that $\lambda = \aleph_\alpha$ for $\alpha < \lambda$. Then $\alpha < \lambda_i$ for some i. But then

$$\lambda = \aleph_\alpha < \aleph_{\lambda_i} = \lambda_{i+1} < \lambda,$$

which is a contradiction. □

The aleph fixed point we constructed is clearly singular, having cofinality ω. Are there regular aleph fixed points? Since aleph fixed points are limit cardinals (check it!) answering this question would answer the previous question. So we will come back to it in part 1.

Exercise 144. Fix your favorite regular cardinal ρ. Modify the construction of the above proof to show that above any cardinal κ there is an aleph fixed point of cofinality ρ. (Hint: it should be clear from your construction that the cofinality of your aleph fixed point is at most ρ. You must also show that the cofinality cannot be $< \rho$.)

We should also consider how cardinal exponention plays with the cardinals.

Definition 145. The *beth numbers* are defined as follows.

- $\beth_0 = \aleph_0 = \omega$;
- $\beth_{\alpha+1} = 2^{\beth_\alpha}$; and
- $\beth_\gamma = \sup_{\alpha<\gamma} \beth_\alpha$ for γ limit.

We can also relativize this, starting at a cardinal κ.

- $\beth_0(\kappa) = \kappa$;
- $\beth_{\alpha+1}(\kappa) = 2^{\beth_\alpha(\kappa)}$; and
- $\beth_\gamma(\kappa) = \sup_{\alpha<\gamma} \beth_\alpha(\kappa)$ for γ limit.

Using this terminology, we can state Cantor's continuum hypothesis as the equality

$$\aleph_1 = \beth_1.$$

We can also state the *Generalized Continuum Hypothesis* in the form:

$$\forall \alpha \in \mathrm{Ord}\ \aleph_\alpha = \beth_\alpha.$$

Exercise 146. Prove that for all α that $\aleph_\alpha \leq \beth_\alpha$.

Exercise 147. Define what it means for a cardinal to be a beth fixed point. Show that for every cardinal κ there is a beth fixed point $> \kappa$.

Definition 148. An infinite cardinal κ is a *strong limit* cardinal if $2^\lambda < \kappa$ for all $\lambda < \kappa$.

Exercise 149. Show that $\beth_\gamma$ is strong limit for all limit ordinals γ.

Question 150. *Is there a regular strong limit ordinal?*

We will return to this question in the next part.

To finish off this section, let us prove an important theorem about cardinal arithmetic.

Theorem 151 (Gyula Kőnig)**.** *Let I be a set and suppose $\langle \kappa_i : i \in I\rangle$ and $\langle \lambda_i : i \in I\rangle$ are sequences of cardinals so that $\kappa_i < \lambda_i$ for each $i \in I$. Then,*

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i.$$

Proof. Let A_i, for $i \in I$, be pairwise disjoint sets of cardinality, respectively, κ_i. We want to show that there is no surjective function $f : \bigcup_{i\in I} A_i \to \prod_{i \in I} \lambda_i$. To see this, take a function f with appropriate domain and range. For each $i \in I$ let f_i be the composition of f with the projection map $\prod_{j \in I} \lambda_j \to \lambda_i$. Because $|A_i| < \lambda_i$ we can pick, for each $i \in I$, $b_i \in \lambda_i \setminus \operatorname{ran} f_i$. But then $(b_i : i \in I)$ is not in $\operatorname{ran} f$, so f is not surjective onto the product. □

Note that we used Choice to prove this theorem. In fact, it is equivalent to Choice over ZF.

Corollary 152. *Let κ be an infinite cardinal. Then $\kappa < \kappa^{\operatorname{cof} \kappa}$.*

Proof. By applying Kőnig's theorem with $I = \operatorname{cof} \kappa$, $\langle \kappa_i : i \in \operatorname{cof} \kappa\rangle$ cofinal in κ, and $\lambda_i = \kappa$. Then

$$\kappa = \sum_{i \in \operatorname{cof} \kappa} \kappa_i < \prod_{i \in \operatorname{cof} \kappa} \kappa = \kappa^{\operatorname{cof} \kappa}.$$ □

Corollary 153. *Let κ be an infinite cardinal and $\lambda \geq 2$. Then $\kappa < \operatorname{cof} \lambda^\kappa$.*

Proof. Otherwise if $\kappa = \operatorname{cof} \lambda^\kappa$ then

$$\lambda^\kappa < (\lambda^\kappa)^\kappa = \lambda^{\kappa \cdot \kappa} = \lambda^\kappa,$$

a contradiction □

In particular, we know that $2^{\aleph_0} \neq \aleph_\omega$.

Exercise 154. Show that if κ is regular then H_κ satisfies Replacement.

(Kameryn J. Williams) University of Hawaiʻi at Mānoa, Department of Mathematics, 2565 McCarthy Mall, Keller 401A, Honolulu, HI 96822, USA
E-mail address: kamerynw@hawaii.edu
URL: http://kamerynjw.net