Set Theory and Logic

May 14, 2008

This course is often found to be difficult, for two main reasons; logic involves a lot of new concepts, and the mathematical phenomena of set theory are also strange and new.

Going against the conventional wisdom on this subject, we will first concentrate on set theory, and then lead gently onto logic, which means we will actually follow the course schedules quite closely.

There are three main components to the course: first, mathematical sets (covering topics: ordinals, cardinals, the axiom of choice and Zorn's lemma, and fixed points and so on; these blend together rather than forming separate chapters), second logic (covering propositional logic, which is not actually that hard, and predicate logic. This latter we shall study quite seriously; it is still possible to say things even at this level which are not found in textbooks. We will consider the limitations of logic, and the distinction between informal and formal reasoning), and finally formal set theory (axioms, axiomatic development, and a nonexaminable section on independence and consistency)

Books

The closest thing to a course book is P.T. Johnstone's "Notes on Logic and Set Theory". A nice book recommended for the second section is D. van Dalen's "Logic and Structure"; for the first and particularly the third section, "Set Theory" by A. Heingel and P. Hamburger is quite nice. For "bedside reading" rather than the course, the lecturer recommends T Forser's "Logic Induction and Sets", which takes a somewhat computer-scientific viewpoint, and also a rather hard book with few proofs, H. Heinlich's "Axiom of Choice". Any other books on the schedules are recommended against by the lecturer.

1 Well-orderings and Ordinals

1.1 Motivation

For the natural numbers $\mathbb{N}$, which we take in this course to include 0, there are several nice properties; in partucular, we have proof by induction: if we have a set of statements $P(n)$ for each $n \in \mathbb{N}$ (e.g. $\sum_1^n r^2 = \frac{1}{6}n(n+1)(2n+1)$), then if we can show $P(0)$ and $\forall n, P(n) \Rightarrow P(n+1)$ then we have $\forall n P(n)$. We can also define functions by recursion, e.g. $0! = 1, (n+1)! = (n+1)n!$. The idea behind both of these is that if we have dealt with a (proper) initial segment of $\mathbb{N}$ then there is a "next" element of $\mathbb{N}$, and if we know how to "go on" to this next element in general, then we can go on "forever". We ask: how general is this property?

1.2 Totally ordered sets

Definition: A totally ordered set $(X, <)$ is a set X equipped with a binary relation $<$ such that: $x < y, y < z \Rightarrow x < z \forall x, y, z \in X$, $x \not< x \forall x \in X$, and (trichotomy) $\forall x, y \in X$ either $x < y$ or $x = y$ or $x > y$.

Note that this implies that only one possibility among the three for trichotomy holds; if $x < y$ and $x = y$ then $x < x$, a contradiction; if $x < y$ and $y < x$ we again have $x < x$.

Examples: $\mathbb{N}, \mathbb{Q}$ or $\mathbb{R}$ with the standard $<$. A non-example is $S^1 = \{(x, y) : x^2 + y^2 = 1\}$ with $(x_1, y_1) < (x_2, y_2)$ if $y_1 < y_2$; trichotomy does not hold for e.g. $(-1, 0), (1, 0)$.

Example: the lexicographic ordering: given a set (alphabet) Σ write $\Sigma^\star$ for the set of finite sequences (words) $\vec{a} = (a_1, \ldots, a_n)$ of elements of Σ (including the empty sequence). Suppose we have a total order ¡ on Σ. Then define $<$ on $\Sigma^\star$ by $\vec{a} = (a_1, \ldots, a_n) < \vec{b} = (b_1, \ldots, b_m)$ if $\exists i$ such that $a_j = b_j \forall j < i$ and either a_i "doesn't exist" while b_i does, or $a_i < b_i$. This is the order in which words appear in a dictionary, so it is plausible that it is a total order on $\Sigma^\star$ at least for finite Σ; in fact it is for general Σ: that $\vec{a} \not< \vec{a}$ is obvious, the reader may check $\vec{a} < \vec{b}, \vec{b} < \vec{c} \Rightarrow \vec{a} < \vec{c}$. We shall check trichotomy: suppose $\vec{a} \neq \vec{b}$, and let i be the first place where $\vec{a}, \vec{b}$ differ. Either one is defined at i and the other is not, in which case the latter is ¡ the former, or a_i, b_i both exist; then if $a_i < b_i$ then $\vec{a} < \vec{b}$, otherwise $b_i < a_i$ and $\vec{b} < \vec{a}$.

Notation: write $x \leq y$ to mean $x < y$ or $x = y$. Then we could equivalently define totally ordered sets to be $(X, \leq)$ such that $x \leq y$ and $y \leq z \Rightarrow x \leq z$, $x \leq y$ and $y \leq x \Rightarrow x = y$, and $\forall x, y$ either $x \leq y$ or $y \leq x$.

1.3 Well-orderings

Definition: a well ordered set $(A, <)$ is a totally ordered set for which if $\emptyset \neq X \subset A$ then X has a ¡-least-element, i.e. a $a_0 \in X$ such that $a_0 \leq a \forall a \in X$.

Examples: $(\mathbb{N}, <)$: we have the minimum principle (which is equivalent to mathematical induction).

Any finite totally ordered set is isomorphic to one of the form $\{0 < 1 < \cdots < n\}$ for some n.

Note $\emptyset$ is well ordered.

Non-examples are $\mathbb{Q}$ or $\mathbb{R}$ with the standard ¡; take $X = \{a : a > 0\}$.

Some examples as subsets of $\mathbb{Q}$ or $\mathbb{R}$: $\{1 - \frac{1}{n} : n \geq 1\} \cup \{1\}$; since distance is irrelevant this is order-isomorphic to $\mathbb{N}$ with an "extra" point "on top".

$\{1 - \frac{1}{n} : n \geq 1\} \cup \{2 - \frac{1}{n} : n \geq 1\}$

What about $\{m - \frac{1}{n} : n, m \geq 1\}$ [this is clearly also well ordered]? Is $\mathbb{N}^\star$ with the lexicographic ordering well-ordered? [Spoilers: no]

It is useful to have the equivalent formulation that a well-ordering is a total ordering for which $\forall 0 \neq X \subset A \exists a_0 : \forall a < a_0, a \notin X$.

Proposition: A total ordering $(A, <)$ is a well-ordering iff it satisfies the principle of $<$-induction: if [for some $P \subset A$] $\forall a, \forall b < a, b \in P \Rightarrow a \in P$, then $\forall a a \in P$ (i.e. $P = A$) (equivalently if $\forall a \forall b < a P(b) \Rightarrow P(a)$ then $\forall a P(a)$: suppose $(A, <)$ is well-ordered, take $P \subset A$ satisfying the condition. Suppose

$P \neq A$, then $A \setminus P$ is non-empty, so take $a \in A \setminus P$ ¡-minimal, then $\forall b < ab \in P$, but then by the induction condition $a \in P$, a contradiction. For the converse, suppose we have ¡-induction; take $X \subset A$ with no ¡-minimal element. Then $A \setminus X$ satisfies the induction condition: if $\forall b < a, b \notin X$ then $a \notin X$, as otherwise a is a ¡-minimal element of X. So by induction $A \setminus X = A$ so $X = \emptyset$.

If $(X, \leq)$ is a total order, an initial segment is $S \subset X$ such that $x \leq y \in S \Rightarrow x \in S$. If $a \in X$ then $\{x : x < a\} =: A_{<a}$ is an inital segment; for example, in $\mathbb{R}$ the $A_{<a}$ are the intervals $(-\infty, a)$. So we see that not all initial segments are of this form, e.g. $(-\infty, 0]$.

An initial segment S is proper if $S \neq A$. In a well-ordering $(A, <)$ all proper initial segments are of the form $A_{<a}$ for some a; this is [the lecturer claims] the best way to conceptialise a well-ordering: suppose S is a proper initial segment, $A \setminus S \neq \emptyset$ so take $a \in A \setminus S$ ¡-minimal; then $\forall b < a, b \in S$ so $A_{<a} \subset S$, and if $x \in S$ then $x \ngeq a$ (as otherwise $a \in S$ by the definition of S, and we have a contradiction) so $x < a$ and $S = A_{<a}$. So for each proper initial segment B of a well-order $(A, \leq)$ there is an $s(B) \in A$ such that $B = A_{<s(B)}$; the converse of this is an exercise.

Note that an initial segment of a well ordering will be a well ordering.

1.4 Order Isomorphisms

Let $(A, <), (B, <)$ be well orderings; an order-isomorphism from A to B is a bijection $A \to B$ such that $a < a' \Rightarrow f(a) < f(a') \forall a, a' \in A$ (note that by trichotomy the $\Rightarrow$ is an if and only if). For such an f, $B_{<f(a)} = f(A_{<a})$; thus $f(a) = s(B_{<f(a)}) = s(f(A_{<a}))$.

Lemma: If $f, g : A \to B$ are order isomorphisms then $f = g$, by induction: if $f = g$ on $A_{<a}$ then $f(a) = s(f(A_{<a})) = s(g(A_{<a})) = g(a)$.

Suppose $f_1 : A_1 \to B_1, f_2 : A_2 \to B_2$ are order isomorphisms between initial segments of A and B, then the restrictions of f_1, f_2 to the initial segment $A_1 \cap A_2$ are equal [in this case one of the A_i is a subset of the other, since we have a total order. But this style of proof generalises better].

Let $(f_i : A_i \to B_i)_{i \in I}$ be the family of all order isomorphisms between iniital segments of A and B. Then taking unions we have an order isomorphism $\bigcup_i A_i = A' \to B' = \bigcup_i B_i$. Suppose that both A', B' are proper; then we can extend f' to an order isomorphism $A' \cup \{s(A')\} \to B' \cup \{s(B')\}$ by $s(A') \mapsto s(B')$, a contradiction since the f_i are all order isomorphisms between initial segments of A, B. Thus we have proved:

Theorem: Let $(A, <), (B, <)$ be well orderings, then either A is (uniquely) order isomorphic to an inital segment of B or vice versa, since in the above either $A' = A$ or $B' = B$.

1.5 Ordinals as Order Types

We say $(A, <), (A', <)$ have the same order type, and so represent the same ordinal, if they are order isomorphic; we can define (following Frege) an ordinal as an equivance class of well-orderings under order isomorphism. Thus properties of ordinals are equivalent to properties of well-orderings which are invariant under order isomorphism.

Notation, due to Cantor: we shall use $\alpha, \beta, \ldots$ for our ordinals and write $\bar{A} = \alpha$ when A is a representative for α (there are canonical representatives for all ordinals, but this is irrelevant to us).

We have an ordering on ordinals: let $\alpha \leq \beta$ just when $\exists \bar{A} = \alpha, \bar{B} = \beta$ such that A is an initial segment of B; $\alpha < \beta$ when this is a proper initial segment. Since order isomorphisms between initial segments of well orderings are unique we have $\bar{A} \leq \bar{B}$ and $\bar{B} \leq \bar{A} \Rightarrow \bar{A} = \bar{B}$ (if $\bar{A} < \bar{B}$ and $\bar{B} < \bar{A}$ we have $\bar{A} < \bar{A}$, which could not happen as the unique order isomorphism from a set to itself is the identity). From above for any A, B either $\bar{A} \leq \bar{B}$ or $\bar{B} \leq \bar{A}$. So the collection of ordinals On is totally ordered by ¡.

The proper initial segments of a well order A are of the form $A_{<a}$ for some (unique) $a \in A$, so the set of proper initial segments of A is order isomorphic to A under the inclusion relation $\subset$, i.e. for any ordinal α, $\{\beta : \beta < \alpha\}$ is order isomorphic to α.

Suppose X is a nonempty subclass of On; take $\alpha \in X$. Either α is ¡-minimal in X or $\{\beta < \alpha : \beta \in X\}$ is nonempty, but then this set is a subset of one $\simeq \alpha$, so we can find a ¡-minimal element of it, β; thus either way we have a ¡-minimal element of X. So On is well-ordered by ¡.

Burali-Forti Paradox: Let Ω be the order type of On; then On is order isomorphic to $\{\alpha : \alpha < \Omega\} = On_{<\Omega}$, and On is isomorphic to initial segment of itself, a contradiction (compare this with Cantor's paradox: let V be the collection of all sets, then $P(V) = V$, a contradiction by Cantor's theorem, or the well known Russell's Paradox). The lecturer claims we should not be worried by this; the usual solution is to say that On is not a set.

1.6 Ordinal Arithmetic

The least ordinal 0 is the order type of $\emptyset$. For $\alpha = \bar{A}$ take $\infty \notin A$, and orrde $A \cup \{\infty\}$ by extending the order on A by $a < \infty \forall a \in A$; this gives a well ordering of order type $\alpha + 1$. Observe that if $\bar{A} = \alpha$ and A has a maximal element α_0 then $\beta = A_{<\alpha_0}$ is such that $\alpha = \beta + 1$; such an α is called a successor ordinals. The other nonzero (it is best to consider 0 separately, even though the reasons for such are not presently aparrent) ordinals are the limit ordinals, which are the nonzero ordirnals α such that $\beta < \alpha \Rightarrow \exists \gamma : \beta < \gamma < \alpha$.

Let $\alpha = \bar{A}, \beta = \bar{B}$; take $A + B = A \sqcup B$ (e.g. by $\{0\} \times A \cup \{1\} \times B$), and extend the order on A, B by setting $a < b \forall a \in A, b \in B$; this gives a well ordering. For $A \times B$ order anti-lexicographicly: $(a, b) < (a', b')$ iff $b < b'$ or $b = b'$ and $a < a'$. This is clearly a well ordering: for $X \subset A \times B$ non-emty take b_0 to be minimal such that $(a, b_0) \in X$ for some a, then set $X_{b_0} = \{a : (a, b_0) \in X\}$; this is a nonempty subset of A so has a minimal element a_0. Then (a_0, b_0) is minimal in X.

For subtraction, we set $\alpha - \beta = 0$ if $\beta \geq \alpha$. For $\beta \leq \alpha$ take $\bar{A} = \alpha$ and B an initial segment [of A] with $\bar{B} = \beta$, then $A \setminus B$ is a subset of A so well ordered; define $\alpha - \beta$ to be its order type.

Suprema of sets of ordinals: for a set X of ordinals, assume we have representatives A_β for $\beta \in X$ such that $\forall \beta \leq \gamma$ $(\beta, \gamma \in X)$ A_β is an initial segment of A_γ (we can do this by taking $On_{<\beta}$ as our representative for each β), then $A = \bigcup_{\beta \in X} A_\beta$ is well ordered by the union of the order relations (e.g. if $a < b < c \in A$ then $c \in A_\beta$ for some β so $a < c$; similarly for ¡-minimal elements) [We define the supremum of X to be the order type of this A].

The assumption was not really necessary, because we could take any representatives A_β and then quotient $\bigcup A_\beta$ by identifying elements which correspond under the unique order isomorphisms between initial segments of our A_β.

Note that we have a kind of "continuity" in the second argument of addition and multiplication: $\alpha + \sup_{i\in I} \beta_i = \sup_{i\in i} \alpha + \beta_i$ (provided $I \neq \emptyset$) and $\alpha_i \sup_{i\in I} \beta_i = \sup_{i\in I} \alpha\beta_i$.

Aside: A set X of ordinals could have a maximal element β_0, then $\sup X = \beta_0$ and $\beta_0 + 1$ is the least [ordinal] ¿ all the elements of X; if it does not, then $\sup X \notin X$ and $\sup X$ is the least [ordinal] ¿ all the elements of X.

Some ordinals: we have $0, 1, 2, \ldots$; the supremum of this is ω. Then we have $\omega + 1$ (note that $1 + \omega$ is simply ω), $\omega + 2, \ldots, \omega + \omega$ which we can calle $\omega 2$ (note that 2ω is simply ω. So we have $\omega 3, \ldots, \omega\omega = \omega^2, \ldots, \omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots, \epsilon_0$, which is the ordinal with the property that $\omega^{\epsilon_0} = \epsilon_0$, just as ω^ω is the ordinal with the property that $\omega\omega^\omega = \omega^\omega$, ω^2 is the ordinal with the property that $\omega + \omega^2 = \omega^2$, and so on.

Hartog's Lemma

This is important. Let X be a set, then there is a least ordinal $\gamma = \gamma(X)$ such that γ does not inject into X (i.e. such that $\gamma = \bar{C} \Rightarrow C$ does not inject into X): consider the set $W = \{R \subset X \cdot X : R \text{ is a well ordering of some subset of } X\}$ (possible R are possible relations). We have a function $W \to On$ by $R \mapsto \rho = \bar{R}$; the image Z is a set of ordinals, and as W is closed under initial segments, Z is an initial segment of the ordinals. The order type γ of Z is such that $Z = On_{<\gamma}$; $\gamma \notin Z$ and γ is the least such ordinal.

Recall that $\omega = \bar{\mathbb{N}}$; we may set $\omega_0 = \omega$. Then set $\omega_1 = \gamma(\mathbb{N})$, which will be the least uncountable ordinal; set $\omega_2 = \gamma(\omega_1)$ and so on.

1.7 The Recursion Theorem

Suppose $(A, <)$ is a well ordered set, X a set and g a function from the set of partial functions $A \to X$, $\mathrm{Ptl}(A, X)$, to X Then there is a unique function $f : A \to X$ such that $g(a) = g(f \mid_{A<a})$ $[\forall a \in A]$

Example of use: suppose $(A, <), (B, <)$ are well orderings (we use $X = B \cup \{\infty\}$ where $\infty \notin B$); then by this theorem there is a function $f : A \to B \cup \{\infty\}$ such that $f(a) = s\{f(a^{|}prime) : a' < a\}$ so long as the latter is a proper initial segment of B and ∞ otherwise (so this is another proof that either A is isomorphic to an initial segment of B or B is isomorphic to a proper initial segment of A).

Proof of the above theorem: Let an attempt be a map $\phi : A' \to X$ for some initial segment A' of A with $\phi(a) = g(\phi \mid_{A<a}) \forall a \in A^{|}prime$; by induction we have that if $\phi_1 : A_1 \to X, \phi_2 : A_2 \to X$ are attempts then $\phi_1 = \phi_2$ on their intersection. So take f to be the union of all attempts ϕ, i.e. $f(a) = x$ if $\phi(a) = x$ for some attempt ϕ. Then we certainly have $f(a) = g(f \mid_{A_{<a}}) \forall a \in$ the domain of f, some A'. But this is an initial segment so f itself is an attempt. If $A' \neq A$ then the domain of f is $A_{<a_0}$ for some a_0 and we can extend this to an attempt $\bar{f}$ by setting $\bar{f}(a_0) = g(f \mid_{A'})$, contradicting the definition of f. So the domain of f is A.

We now want to replace $(A, <)$ by the class $(On, <)$; write $\mathrm{Ptl}(On, V)$ for the set of ϕ defined on some subset (note set rather than class) of On with $\phi(\alpha)$ a set where defined (V here is the class of all sets):

Recursion Theorem: Given $G : \mathrm{Ptl}(On, V)L \to V$ there is a unique $F : On \to V$ such that $F(\alpha) = G(F\mid_{On_{<\alpha}})$.

Examples of applications: for fixed α we can define $\alpha + \beta$ by recursion on β: $\alpha + 0 = \alpha$, $\alpha + (\beta + 1) = (\alpha + \beta) + 1$ (where by $\delta + 1$ we mean the successor of δ)) (for successor ordinals) and $\alpha + \lambda = \sup_{\beta<\lambda} \alpha + \beta$ for limit ordinals γ. (Given a definition like this, we can define a suitable G by "brute force"; in this case, define $G(\phi)$ to be α if ϕ is everywhere undefined, $\phi(\beta_0) + 1$ if ϕ is defined on an initial segment of ordinals with greatest element β_0 and β_0 is an ordinal, sup(range(ϕ)) if ϕ is defined on a (nonempty) initial segment of ordinals with no greatest element and range$(\phi) \subset On$, 0 (or whatever you will; mathematicians joke about defining it to be the moon) for any other ϕ. Or in this particular case we can define $G(\phi) = \alpha \vee \{\phi(\beta) + 1 :$ β an ordinal in the domain of ϕ with $\phi(\beta)$ an ordinal$\}$, where $A \vee B$ means the supremum of $A \cup B$).

For fixed α we define $\alpha\beta$ by $\alpha 0 = 0, \alpha(\beta + 1) = \alpha\beta + \alpha, \alpha\lambda = \sup_{\beta<\lambda} \alpha\beta$ for λ a limit.

For fixed α we define α^β by $\alpha^0 = 1, \alpha^{\beta+1} = \alpha^\beta\alpha, \alpha^\lambda = \sup_{\beta<\lambda} \alpha^\beta$ for λ a limit; note that we can easily show by induction on β that for α, β countable so is α^β (which is not so for cardinals; see later).

2 Axiom of Choice

(Equivalents and consequences)

2.1 Statement (and some observations)

Axiom of Choice: Let X be a set of nonempty sets, then there is a function $c : X \to \bigcup X$ (i.e. $\bigcup_{x\in X} x$) such that $c(x) \in x \forall x \in X$. Some trivial alternative forms are for $\{x_i : i \in I\}$ an indexed family of nonempty sets, there is a function $c : I \to \bigcup_{i\in I} x_i$ such that $c(i) \in x_i \forall i \in I$; we will use this form for our observations below. As the final form, for any surjective map $p : Z \to I$ there is a maps$s : I \to Z$ such that $p \circ s : I \to I$ is the identity.

Observations: If I is finite then there is no problem: for $I = \{1, \ldots, n\}$ we take $c_i \in x_i \forall 1 \le i \le n$ and define $c(i) = c_i$; this is "definition by a finite number of cases", or "just do it". The axiom is almost no help if the sets x_i are finite, but even then it depends on the nature of the sets; as Russell famously said, we can pick members from an infinite collection of pairs of shoes (by e.g. picking the left shoe from each), but can we pick members from an infinite collection of pairs of socks?

2.2 Zorn's Lemma

A partially ordered set $(X, <)$ or $(<\le)$ is a set such that $x \le y \le z \Rightarrow x \le z$ and $x \le y$ and $y \le x \Rightarrow x = y$. A subset C of a partially ordered set $(X, \le)$ is a chain if it is totally ordered; note the empty set is a chain. An upper bound for any $Y \subset X$ where X is a partially ordered set is a $y_0 \in X$ with $y_0 \ge y \forall y \in Y$; a maximal element in a partially ordered set $(X, \le)$ is an x_0 such that $\forall x \in X$, $x \ge x_0 \Rightarrow x = x_0$.

Statement of Zorn's Lemma: Let $(X, \leq)$ be a partially ordered set in which every chain has an upper bound, then X has a maximal element. (Note that X is automatically nonempty since the empty chain has an upper bound; some earlier works prefer to require this explicitly).

Proof: Suppose X has no maximal element. For a chain C, we have an upper bound $\bar{u}_C$, which by assumption is not a maximal element, so we can take $u_C > \bar{u}_C$, and have a strict upper bound for any C; by the axiom of choice we can define u from the set of chains in X to X with $u(C)$ a strict upper bound for each C. Recall (Hartog) that there is an ordinal $\gamma(X)$ which does not embed into X. Define $f : On_{<\gamma(X)} \to X$ recursively by $f(\alpha) = u(f(On_{<\alpha}))$ for $f(On_{<\alpha})$ a chain, $u(\emptyset)$ otherwise (this will be seen to be irrelevant). We claim $\forall \alpha < \gamma(x)$ $f(On_{<\alpha})$ is a chain, by induction on α: suppose $\forall \beta < \alpha f(On_{<\beta})$ is a chain. In the case $\alpha = 0$ or α is a limit ordinal, $On_{<\alpha} = \bigcup_{\beta<\alpha} On_{<\beta}$ so $f(On_{<\alpha} = \bigcup_{\beta<\alpha} f(On_{<\beta})$, a nested union of chains so a chain; if $\alpha = \beta + 1$, $On_{<\alpha} = On_{<\beta} \cup \{\beta\}$ so $f(On_{<\alpha})/f(On_{<\beta}) \cup \{f(\beta)\}$ which is a chain together with a proper upper bound (by the definition of f) so a chain, so in either case $f(On_{<\alpha})$ is a chain, and so $f(On_{<\alpha})$ is a chain $\forall \alpha < \gamma(X)$. So the first condition in the definition of f always holds, so if $\beta < \alpha < \gamma(X)$ then $f(\alpha) = u(f(On_{<\alpha})) > f(\beta)$, so $f : On_{<\gamma(X)} \to X$ is injective, contradicting the definition of γ. Thus X has a maximal element.

Aside: we have actually used a modified principle of induction: suppose $P \subset On_{<\gamma})$ satisfies the modified induction condition that $P(0)$, $\forall \beta P(\beta) \Rightarrow P(\beta + 1)$, and for λ a limit, $\forall \beta < \lambda P(\beta) \Rightarrow P(\lambda)$, then $\forall \alpha < \gamma P(\alpha)$. This follows easily from the "official" principle of ¡-induction because the "official" induction condition follows easily from the modified one; in fact the two are equivalent.

2.3 Uses of Zorn's Lemma in Mathematics

We usually only use a simpler form, the observation below:

Definition: A supremum for $Y \subset$ a partially ordered set X is $y_0 \in X$ such that $y_0 \geq y \forall y \in Y$ and if $z \geq y \forall y \in Y$ then $z \geq y_0$ (i.e. y_0 is a least upper bound); clearly this is unique if it exists. $(X, \leq)$ is complete if all subsets have suprema, and chain complete if all chains have suprema.

Observation: If $(X, \leq)$ is chain complete then it has a maximal element.

Application: any vector space V has a basis: let $\mathcal{L}$ be the poset (partially ordered set) of linearly independent subsets of V under inclusion; we claim $\mathcal{L}$ is chain complete: suppose $\{L_i : i \in I\}$ is a chain in $\mathcal{L}$, then $\bigcup_{i \in I} L_i$ is linearly independent (for $\vec{e}_1, \ldots, \vec{e}_n \in \bigcup_i L_i$ with $\sum_i \lambda_i \vec{e}_i = \vec{0}$, each $\vec{e}_k$ lies in some L_{i_k}, and there must be a maximal element L of the finitely many chain elements $L_{i_1}, \ldots, L_{i_n}$, then $\vec{e}_1, \ldots, \vec{e}_n \in L$, a linearly independent set, so $\lambda_i = 0 \forall i$). So by ZL L has a maximal element L_0; were this not a basis, have $\vec{v} \in V \setminus \langle L_0 \rangle$ and then $L_0 \cup \{\vec{v}\}$ is linearly independent, a contradiction.

Application: Any ring R with $0 \neq 1$ has a maximal ideal: consider the collection of proper ideals $I \lhd R$; this is chain complete ($\emptyset$ has $\{0\} \lhd R$ as its supremum, a nonempty chain $(I_k : k \in K)$ has $\bigcup_{k \in K} I_k$ as its supremum), so by ZL we have a maximal eleement of the poset, which will be a maximal ideal.

Application: If R is a ring and $I \lhd R$ and $a \in R$ such that $a^0, a^1, a^2, \ldots \notin R$ then there is a prime ideal $P \lhd R$ with $a^n \notin P \forall n = 0, 1, \ldots$ (note this is automatically the case if $a \notin P$): take the poset of ideals J with $I \leq J$ and

$a^n \notin J \forall n$; this is chain complete (ordered by inclusion) as in the previous example, so by ZL there is a maximal element P. We have $1 = a^0 \notin P$. Suppose $b, c \notin P$, then $\langle P, b\rangle \ni a^r, \langle P, c\rangle \ni a^s$ since these ideals are bigger than P. So $a^r = p_0 + \lambda b, a^s = p_1 + \mu c$ so $a^{r+s} = (p_0 p_1 + \dots) + \lambda\mu bc$, so $\langle P, bc\rangle > P$ and $bc \notin P$. Thus P is prime as required.

2.4 The Well-Ordering Principle

Theorem: Any set can be wess ordered: for a set X consider the set W of well orderings of subsets of X, ordered by initial segments ($A \leq B$ if A is an initial segment of B). Consider a chain $\{A_i : i \in I\}$ in W; this is a set of well orderings such that among any two, one is an initial segment of the other, so $\bigcup_{i\in I} A_i$ is a well ordering; thus W is chain-complete, so by ZL it has a maximal element $(X', <')$. Were $X' \neq X$ take $x_0 \in X \setminus X'$, and order $X' \cup \{x_0\}$ by $x' < x_0 \forall x' \in X'$, which gives a well ordering with X' as a proper initial segment, contradicting the definition of X'. So $X = X'$ and X is well ordered.

The well-ordering principle WO is the statement that every set can be well ordered.

Proposition: WO⇒AC: suppose $(X_i : i \in I)$ is a family of non-empty sets; take a well-ordering of $\bigcup_{i\in I} X_i$ and define $c : I \to \bigcup_{i\in I} X_i$ by $c(i)$ is the ¡-least element in $X_i \subset \bigcup_{j\in I} X_j$.

Remarks: we thus have AC, WO, ZL are equivalent; it is quite hard to prove the reverses of the implications we have proved (e.g. to show $AC \Rightarrow WO$) directly. Although we have proven that e.g. a well-ordering of $\mathbb{R}$ must exists, it is hard to imagine what this would “look like”.

2.5 ZL via the Bourbake-Witt Theorem

The schedules suggest this is a preferred method of proving ZL, but the lecturer entirely disagrees.

Theorem: Let $(X, \leq)$ be a chain-complete poset and $h : X \to X$ an increasing function, i.e. $x \leq h(x) \forall x \in X$. Then h has a fixed point $x_0 \in X$.

Proof of this is easy using the ordinals $< \gamma(X)$; define $f : On_{<\gamma(X)} \to X$ recursively by $f(0) = \sup \emptyset$ (the bottom element, 1), $f(\beta+1) = h(f(\beta)), f(\lambda) = \sup_{\beta<\lambda} f(\beta)$ for λ a limit (so long as $\{f(\beta) : \beta < \lambda\}$ is a chain) (and $f(\alpha) = 1$ otherwise). By recursion $\{f(\beta) : \beta < \lambda\}$ is always a chain; then if h has no fixed point we have $f(\beta) < f(\beta+1) \forall \beta$ and if $\beta < \lambda$ with λ a limit then $f(\beta) < f(\lambda)$ (as otherwise $f(\beta) = f(\beta+1)$, so f embeds $On_{<\gamma(X)}$ into X, a contradiction; thus h has a fixed point.

Suppose $(X, \leq)$ is a chain complete poset; if X has no maximal elements then by AC we have a $h : X \to X$ with $x < h(x) \forall x \in X$, an increasing function with no fixed point, contradicting the theorem, so the theorem gives ZL for chain complete posets.

Now, to prove the stronger form of ZL, take $(X, \leq)$ a poset in which every chain has an upper bound. Consider the set of chains in X, ordered by inclusion; this is chain complete, so there is a maximal chain C; take x_0 an upper bound for C, then x_0 is maximal in X, as if there is $\bar{x} > x_0$ then $C \cup \{\bar{x}\}$ is a chain, contradicting the definition of C.

Bourbake-Witt without ordinals etc. (Non-examinable sketch)

(It is somewhat disquieting that we needed the entire machinery of ordinals to make a relatively small proof; in fact it is possible to avoid doing so, though this proof is perilous, distasteful and hence nonexaminable)

Consider C defined as the intersection of all $A \subset X$ closed under suprema of chains in X and under h (i.e. $H(A) \subset A$). This has an induction principle! It is sufficient to prove C is a chain, as we can then take $c = \sup C$, which is $\in C$, so $c \leq h(c) \in C \Rightarrow h(c) \leq \sup C \Rightarrow h(c) = c$.

This proof only works by "magical" inspiration: we decide to prove that $(\dagger)\forall x \forall y x \leq y$ or $f(y) \leq x$, by induction on x. It is easy to see that the x such that † holds are closed under suprema of chains; to proove this set is closed under f is very messy to do directly or by immediate induction on y, so we again need inspiration: we proove $\forall y\ y \leq x$ or $f(x) \leq y$ by induction on y; the set of such y is closed under $\vee$s (suprema) as before; if $f(x) \leq y$ then $f(x) \leq f(y)$, if $y \leq x$ then either $x \leq y \Rightarrow x = y \Rightarrow f(x) \leq f(y)$, or $f(y) \leq x$, so in any case we are OK; the induction works, and easily implies † is closed under f; the first induction works, and C is a chain as required.

3 Cardinals and their arithmetic

3.1 Cardinals via equinumerosity

Informally, we want to consider the size (number of elements) of a set independently of its elements.

We say sets X, Y are equinumerous $X \approx Y$ if there is a bijection $X \xrightarrow{\sim} Y$ ($\xrightarrow{\sim}$ denotes a bijection). Then a cardinal or cardinal number is an equivalence class of sets under $\approx$ (which is clearly an equivalence relation); as before, operations on, properties of and propositions about cardinals are the same things about (representative) sets, invariant under $\approx$. We shall write our cardinals as $\vec{m}, \vec{n}, \ldots$; $|M|$ is the cardinal of a set M (Cantor wrote $\overline{\overline{M}}$). If $|M| = \vec{m}$ we say M is a representative of the cardinal $\vec{m}$.

The biggest difference between this and the case of ordinals is that order isomorphisms between well orderings are unique, wheras bijections between sets are hardly ever so.

If there is an injection $X \hookrightarrow Y$ we write $X \lesssim Y$; this is evidently invariant under $\approx$, so we can write $|X| \leq |Y|$. Since a composite $X \hookrightarrow Y \hookrightarrow Z$ of injections is injective, $|X| \leq |Y| \leq |Z| \Rightarrow |X| \leq |Z|$.

Schröder-Bernstein Theorem

Suppose $f : A \to B, g : B \to A$ are injections, then there is a bijection $A \xrightarrow{\sim} B$ (i.e. $|A| \leq |B|, |B| \leq |A| \Rightarrow |A| = |B|$); this gives that $\leq$ is a partial ordering on cardinals:

Set $A_{2n} = (gf)^n(A), A_{2n+1} = (gf)^n gB = g(fg)^n B, B_{2n} = (fg)^n B, B_{2n+1} = (fg)^n fA = f(gf)^n A$. So $f : A_{2n} \xrightarrow{\sim} B_{2n+1}, A_{2n+1} \xrightarrow{\sim} B_{2n+2}$, and so $f : A_{2n} \setminus A_{2n+1} \to B_{2n+1} \setminus B_{2n+2}$; similarly $g : B_{2n} \setminus B_{2n+1} \xrightarrow{\sim} A_{2n+1} \setminus A_{2n+2}$. Also $f^{-1}(\bigcap_{k \geq 0} B_k) = f^{-1}(\bigcap_{k \geq 1} B_k) = \bigcap_{k \geq 1} f^{-1}(B_k) = \bigcap_{k \geq 0} A_k$, so $f : \bigcap_{k \geq 0} A_k \xrightarrow{\sim} \bigcap_{k \geq 0} B_k$. We have $A = (A_0 \setminus A_1) \cup (A_1 \setminus A_2) \cup \cdots \cup \bigcap_k A_k, B =$

$(B_0 \setminus B_1) \cup (B_1 \setminus B_2) \cup \cdots \cup \bigcap_k B_k$, and these unions are disjoint; then we have an isomorphism by taking $g^{-1} : (A_{2n+1} \setminus A_{2n+2}) \xrightarrow{\sim} (B_{2n} \setminus B_{2n+1}), f : (A_{2n} \setminus A_{2n+1} \xrightarrow{\sim} (B_{2n+1} \setminus B_{2n+2})$, and either of these $\bigcap A_k \xrightarrow{\sim} \bigcap B_k$.

3.2 Tarski fixed point theorem

Recall that a poset $(X, \leq)$ is complete if all subsets $Y \sup X$ have suprema.

Example: For any set A consider the power set $P(A)$ ordered by $\subset$. This is a complete lattice (a lattice being something with finiet suprema and infina; see the example sheet for proof that one implies the other) with $\vee\{x_i : i \in I\} = \bigcup\{x_i : i \in I\}$. If $X \subset P(A)$ is closed under unions then $(X, \subset)$ is also a complete poset; in particular if (A, τ_A) is a topological space then $(\tau_A, \subset)$ is complete.

Definition: A map $f : (X, \leq) \to (Y, \leq)$ is order preserving if $x \leq y \Rightarrow f(x) \leq f(y)$.

Theorem: Let $(X, \leq)$ be complete and $f : X \to X$ order preserving, then f has a fixed point: let $a = \vee\{x : x \leq f(x)\}$. Then for $x \leq f(x)$ we have $x \leq a \Rightarrow f(x) \leq f(a) \Rightarrow x \leq f(a) \Rightarrow a \leq f(a)$, so $f(a) \leq f(f(a)) \Rightarrow f(a) \in \{x : x \leq f(x)\} \Rightarrow f(a) \leq a \Rightarrow a = f(a)$.

Application: Let $f : A \to B, g : B \to A$ be injections; define $F : P(A) \to P(A)$ by $F(x) = A \setminus g(B \setminus f(x))$. If $x \leq y$ then $f(x) \subset f(y) \Rightarrow B \setminus f(x) \supset B \setminus f(y) \Rightarrow g(B \setminus f(x)) \supset g(B \setminus f(y)) \Rightarrow F(x) = A \setminus g(B \setminus f(x)) \subset A \setminus g(B \setminus f(y)) = F(y)$, so F is order preserving. So we have a fixed point $\bar{A} \subset A$ with $\bar{A} = F(\bar{A}) = A \setminus g(B \setminus f(\bar{A}))$. Then $f : \bar{A} \xrightarrow{\sim} f(\bar{A}) =: \bar{B} \subset B$ and $g : (B \setminus \bar{B}) \xrightarrow{\sim} g(B \setminus \bar{B}) \subset A$, but $g(B \setminus \bar{B}) = A \setminus \bar{A}$, so $g^{-1} : A \setminus A \xrightarrow{\sim} B \setminus \bar{B}$ and we have a bijection $A \xrightarrow{\sim} B$.

3.3 Cardinal arithmetic

Take $|M| = \vec{m}, |N| = \vec{n}$¿ For addition set $\vec{m} + \vec{n}$ to be the cardinal of the disjoint union $M + N$ (i.e. $\{0\} \times M \cup \{1\} \times N$), for multiplication $\vec{m} \cdot \vec{n}$ is the cardinal of the product $M \times N$, and for exponentiation $\vec{n}^{\vec{m}}$ is the cardinal of the set N^M of all functions $f : M \to N$.

The elementary rules of arithmetic follow from (natural) isomorphisms between sets: addition is associative and commutative with unit $0 = |\emptyset|$, multiplication is associative and commutative with unit $1 = |1|$ where $1 = \{0\}$, multiplication distributes over addition $\vec{n} \cdot (\vec{m} + \vec{p}) = \vec{n} \cdot \vec{m} + \vec{n} \cdot \vec{p}$, and $(\vec{n} \cdot \vec{m})^{\vec{p}} = \vec{n}^{\vec{m}} \cdot \vec{m}^{\vec{p}}, \vec{n}^{\vec{m} \cdot \vec{p}} = (\vec{n}^{\vec{m}})^{\vec{p}}$.

Examples: Finite cardinals, the cardinals of finite sets and of finite ordinals (or well orderings), $0 = |\emptyset|, 1 = |\{0\}|, 2 = |\{0, 1\}|$ etc.

The cardinal of $\mathbb{N}$ is $\omega = \omega_0 = \aleph_0 = |\mathbb{N}| = |\{0, 1, \dots\}|$; this is the “denumerable” or “countably infinite” cardinal.

$\omega_1 = \aleph_1$ is the cardinal of the first innumerable ordinal (note that there is more than one infinite cardinal, essentially by Hartogs).

$2^\omega = 2^{\omega_0}$ is the cardinal of the continuum ($2^{\mathbb{N}} \simeq \mathbb{R}$, because we have injections in both directions: $2^{\mathbb{N}} \hookrightarrow \mathbb{R}$ either by $(a_0, a_1, a_2, \dots) \in 2^{\mathbb{N}} \mapsto \sum_{n=0}^{\infty} a_n (\frac{2}{3})^{n+1}$ onto the Cantor set (aside: the complement of the Cantor set has measure $\frac{1}{3} + \frac{2}{9} + \cdots = 1$, so the Cantor set is an uncountable (see later) set of measure 0), or by $2^{\mathbb{N}} \hookrightarrow \mathbb{N}^{\mathbb{N}} \hookrightarrow \mathbb{R}$ by the continued fraction expansion $(x_0, x_1, \dots) \mapsto x_0 + \frac{1}{x_1 + \frac{1}{\cdots}}$, the image of this second injection being the irrationals, and then there is an injection $\mathbb{R} \to 2^{\mathbb{N}}$ by e.g. composing $\mathbb{R} \to (0, \infty) \to (0, 1)$ by $x \mapsto e^x$ and then $y \mapsto \frac{y}{1+y}$, and then injecting $(0, 1)$ into $2^{\mathbb{N}}$ by mapping a real to its binary expansion, choosing the non-terminating one where we have a choice).

Interlude on Cantor's theorem: for clarity, define $\vec{n} \leq^\star \vec{m}$ to mean that whenever $|N| = \vec{n}, |M| = \vec{m}$ either $N = \emptyset$ or there is a surjection $M \to N$. Observe $\vec{n} \leq \vec{m} \Rightarrow \vec{n} \leq^\star \vec{m}$, for take $f : N \to M$ injective, then either $N = \emptyset$ and we are done, or take $a \in N$ and define $g : M \to N$ by $g(y) =$ the unique x such that $f(x) = y$ if y is in the image of f, and $g(y) = a$ otherwise; this is clearly surjective. If we have AC then $\vec{n} \leq^\star \vec{m} \Rightarrow \vec{n} \leq \vec{m}$; without it, $\leq^\star$ can be very bad to work with; it may not even be a partial order.

Theorem: We never have $2^{\vec{n}} \leq^\star \vec{n}$ (and so never have $2^{\vec{n}} \leq \vec{n}$): note $2^X \simeq P(X)$ by characteristic functions, and so is never empty. Suppose $g : X \to P(X)$ and consider $\{x \in X : x \notin g(x)\}$; this cannot be in the image of g, so g is not surjective.

Representative calculations: beware, we are working with cardinals, not ordinals throughout, and 2^ω the cardinal of the continuum is very different from 2^ω as an ordinal. $\omega + \omega = \omega$ for we have a bijection $\mathbb{N} + \mathbb{N} \to \mathbb{N}$; similarly $\omega \cdot \omega = \omega$ (for an explicit byjection, $(n, m) \mapsto \frac{1}{2}(n+m)(n+m+1)+n$, or just use that we have injections in both directions). $2^\omega \cdot 2^\omega = 2^{\omega+\omega} = 2^\omega$, the cardinality of the set $\mathbb{R}^{\mathbb{N}}$ of all real sequences is $(2^\omega)^\omega = 2^{\omega \cdot \omega} = 2^\omega$. The cardinality of the set $\mathbb{R}^{\mathbb{R}}$ of functions $\mathbb{R} \to \mathbb{R}$ is $(2^\omega)^{2^\omega} = 2^{\omega \cdot 2^\omega}$. Observe that $+, \cdot$ behave well wrt $\leq$, so $2^\omega = 2 \cdot 2^\omega \leq \omega \cdot 2^\omega \leq 2^\omega \cdot 2^\omega = 2^\omega$, so $\omega \cdot 2^\omega = 2^\omega$ and $|\mathbb{R}^{\mathbb{R}}| = 2^{(2^\omega)}$, which is bigger than 2^ω by Cantor.

3.4 The Hierarchy of Alephs

Here we will consider the cardinality of well-orderable sets (i.e. all sets if we assume AC): if X is well-orderable, then there is a minimal α such that X has an ordering of order type α. Ordinals κ which are of this form have $\forall \beta < \kappa$, $On_{<\beta} \not\simeq On_{<\kappa}$ so $|On_{<\beta}| \neq |On_{<\kappa}|$ (hereafter we will identify $On_{<\alpha}$ with α), so $\forall \beta < \kappa |\beta| \neq \kappa$; clearly if $\beta < \kappa$ then $|\beta| \leq |\kappa|$, so $\forall \beta < \kappa |\beta| < |\kappa|$.

We call ordinals of this kind initial ordinals; these are the canonical representatives of well-ordered cardinals. If κ, μ are initial ordinals we have at first sight two orderings: $\kappa \leq \mu$ as ordinals, or $|\kappa| \leq |\mu|$ as cardinals. However, clearly if $\kappa \leq \mu$ as ordinals then $|\kappa| \leq |\mu|$, and conversely suppose $|\kappa| \leq |\mu|$; were $\mu < \kappa$ then since κ is initial we would have $|\mu| < |\kappa|$, a contradiction, so $\kappa \leq \mu$ and the two orderings are the same.

What are these initial ordinals? So far we know the finite ordinals $0, 1, \ldots$, $\aleph_0 = \omega_0 = \omega$ the least infinite ordinal, and $\aleph_1 = \omega_1$ the least uncountable ordinal. But clearly we can define $\aleph_2 = \omega_2$, the least ordinal with cardinality not $\leq \omega_1$, and so on.

For any cardinal $\vec{m} = |M|$ define $\vec{m}^+$ to be $|\gamma(M)|$; from the definition $\gamma(M)$ is always initial, so we can write $|\gamma(M)| = \gamma(M)$. Define a function $On \to On$ $\alpha \mapsto \omega_\alpha = \aleph_\alpha$ by recursion, by $\omega_0 = \omega, \omega_{\beta+1} = \omega_\beta^+$ (we are about to show inductively that ω_α is inital $\forall \alpha$ (though this doesn't actually matter, $\vec{m}^+$ is always initial anyway)), and for λ a limit $\omega_\lambda = \sup_{\beta<\lambda} \omega_\beta$.

Some properties of this function: clearly $\beta \leq \gamma \Rightarrow \omega_\beta \leq \omega_\gamma$, by induction on γ. $\forall \alpha$ ω_α is initial, by induction: ω_0 is initial, for $\alpha = \beta + 1$ $\omega_\alpha = \omega_\beta^+$ which is always initial, for $\alpha = \lambda$ a limit for any $\rho < \omega_\lambda$ we have $\rho < \omega_\beta$ for some

$\beta < \lambda$, and ω_β is initial by the induction hypothesis, so $|\rho| < |\omega_\beta| \leq |\omega_\lambda|$, so ω_λ is initial. Observe that $\alpha \leq \omega_\alpha$ by induction on α.

Take κ an infinite initial ordinal; $\kappa \leq \omega_\kappa$ and so we can take α least such that $\kappa \leq \omega_\alpha$. We claim that then $\kappa = \omega_\alpha$: it is actually easiest to prove this by cases on α. If $\alpha = 0$ then $\kappa \leq \omega$ so $\kappa = \omega$ (since ω is the least infinite initial ordinal), if $\alpha = \beta + 1$ then $\kappa \leq \omega_\beta^+ = \omega_{\beta+1}$ but $\kappa \nleq \omega_\beta$ so $\omega_\beta < \kappa$; observe that there are no objections of κ into ω_β ($|\kappa| \nleq |\omega_\beta|$) so $\kappa \geq \omega_\beta^+$ by the definition of $\vec{m}^+$. So $\kappa = \omega_{\beta+1}$. Finally if $\alpha = \lambda$ a limit then $\kappa \leq \omega_\lambda = \sup_{\beta<\lambda} \omega_\beta$; if $\kappa < \omega_\lambda$ then $\kappa < \omega_\beta$ for some $\beta < \lambda$ contradicting the definition of α. So $\kappa = \omega_\lambda$. Thus the hierarchy ω_α of alephs enumerates the infinite initial ordinals. Thus, under AC, the ordering of cardinals is $\omega + On = On$.

3.5 Cardinal arithmetic with choice

Theorem: Suppose κ is an infinite well-ordered cardinal, then $\kappa \cdot \kappa = \kappa$; we proove this by induction. We need the base case $\omega \cdot \omega = \omega$, but we have done this already. Now, assume $\mu \cdot \mu = \mu \forall$ infinite $\mu \leq \kappa$. Define an ordering $\prec$ on $\kappa \times \kappa$ ($= \{(\alpha, \beta) : \alpha, \beta < \kappa\}$ by $(\alpha, \beta) \prec (\gamma, \delta)$ if either $\max(\alpha, \beta) < \max(\gamma, \delta)$ or $\max(\alpha, \beta) = \max(\gamma, \delta)$ and $\alpha < \gamma$ or $\max(\alpha, \beta) = \max(\gamma, \delta), \alpha = \gamma$ and $\beta < \delta$. This is a total ordering (the proof is easy but tedious, by cases), and a well-ordering: if $\emptyset \neq X \subset \kappa \times \kappa$ then take ρ the least ordinal such that $\rho = \max(\alpha, \beta)$ for some $(\alpha, \beta) \subset X$, α_0 least such that $\exists \beta : (\alpha_0, \beta) \in X_\rho$ the subset of X corresponding to ρ, and then β_0 least such that $(\alpha_0, \beta_0) \in X_{\rho, \alpha_0}$; then (α_0, β_0) is least in X. Finally, $(\kappa \times \kappa, \prec)$ is the union of the $(\rho \times \rho, \prec)$ over all ordinals $\rho < \kappa$ as initial segments; any (proper) initial segment of $\kappa \times \kappa$ is included in some $\rho \times \rho$. $|\rho| = \mu$ is a (well ordered) cardinal $< \kappa$ so $|\rho \times \rho| = \mu \cdot \mu = \mu$ by the induction hypothesis; thus all proper initial segments of $\kappa \times \kappa$ have cardinality $< \kappa$ and so order type $< \kappa$. So the order type of $(\kappa \times \kappa, \prec)$ is $\leq \kappa$, so as cardinals $\kappa \cdot \kappa \leq \kappa$, and $\kappa \cdot \kappa = \kappa$.

Thus, with AC, some cardinal arithmetic is trivial: if μ, κ are infinite cardinals (which are automatically well ordered if we assume AC) then $\mu + \kappa = \mu \cdot \kappa = \max(\mu, \kappa)$, for if we wlog assume $\mu \leq \kappa$ then $\kappa \leq \mu + \kappa \leq \kappa + \kappa = 2 \cdot \kappa \leq \mu \cdot \kappa \leq \kappa \cdot \kappa = \kappa$, [so all of these are equal] - addition and multiplication are boring. However, we have no hold at all on exponentiation; in particular, even with AC, there is no clear intuition as to what $2^{\aleph_0}$ is. The remainder of the course is, in large part, an attempt to set this fact/question into some reasonable context, and ask why it is so.

4 Propositional Logic

Much of the early part of this section is nonexaminable.

Aside: if $(X, \leq)$ is a (small) finite poset we can describe it via its Hasse diagram, effectively a directed acyclic graph; $x \leq y$ if there is a line between them and y is above x.

4.1 Boolean algebras as lattices

Finite suprema (joins) in a poset $(X, \leq)$: the binary case is $a \vee b$ determined by $a, b \leq a \vee b$ and if $a, b \leq x$ then $a \vee b \leq x$; the "zeroary" case is $\bot = 0$ determined by $\perp \leq x \forall x$. Then we inductively define $x b_1 \vee \cdots \vee x_n \forall n \geq 0$. Think of logical or, or set-theoretic union.

We always have 1) $\vee$ is associative and commutative with unit $\perp$ 2) the order can be recovered by $a \leq b \Leftrightarrow a \vee b = b$.

Similarly, finite infina (meets) (which dually are sups in $(X, \leq)^{op}$, the poset with order reversed) are defined by in the binary case $a \wedge b \leq a, b$ and if $x \leq a, b$ then $x \leq a \wedge b$. The "0-ary" infinum is T determined by $x \leq T \forall x$; think of logical and or set-theoretic intersection.

Definition: A lattice is a poset with finite meets and joins; a lattice is distributive if we have $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ and $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$ (in fact these two properties are equivalent). A complement of a in a lattice $(L, \leq)$ is $\bar{a} = \neg a \in L$ such that $a \vee \bar{a} = T, a \wedge \bar{a} = \perp$; note that these are not generally unique. However, in a distributive lattice complements are unique: suppose $\tilde{a}$ is another complement for a, then $\tilde{a} = \tilde{a} \wedge T = \tilde{a} \wedge (a \vee \bar{a}) = (\tilde{a} \wedge a) \vee (\tilde{a} \wedge \bar{a}) = \perp \vee (\tilde{a} \wedge \bar{a}) = \tilde{a} \wedge \bar{a}$ so $\tilde{a} \leq \bar{a}$, and similarly $\bar{a} \leq \tilde{a}$.

Definition: A Boolean algebra is a distributive lattice in which all elements have complements.

Example: $(P(X), \subset)$ is a boolean algebra; moreover if $L \subset P(X)$ is closed under intersections, unions and complements then L is a Boolean algebra, e.g. $L = \{x \in P(\mathbb{N}) : x \text{ or } \mathbb{N} \setminus x \text{ is finite}\}$ is a countable Boolean algebra, so not $\simeq (P(Y), \subset)$ for any Y.

4.2 Boolean algebra as algebra

We can give Boolean algebra by a set of operations $\{T, \wedge, \perp, \vee, \bar{()}\}$, so we can consider Boolean algebras constructed by generators and relations; in particular we have free Boolean algebras: for construction, take a set of generators (as primitive constants), construct terms in the operators, and quotient out by the equations. For Boolean algebras it is easy to see what form of elements we get: 1) apply the de Morgan laws $\neg(A \wedge b) = \neg a \vee \neg b, \neg(a \vee b) = \neg a \wedge \neg b$, and cancelling $\neg\neg a = a$; this brings all $\neg$s "to the centre"; they either act on a single primitive element by $\neg p$, or not at all. 2) Use $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ to bring $\vee$s "out" and $\wedge$s "in"; we get a "normal" form $(\dots) \vee (\dots) \vee \dots$, where the terms in the brackets are $\wedge$s of primitive elements and complements of primitive elements (some further simplifications are possible, but we will not bother with them here).

Free Boolean algebras: $B(0) = 2$ has two elements $T, \perp$. $B(1)$ has four elements, $\perp, p, \bar{p}, T$. $B(2)$ has $2^4 = 16$ elements given by unions of the 4 elements nearest the bottom, $p \wedge q, p \wedge \bar{q}, \bar{p} \wedge q, \bar{p} \wedge \bar{q}$. In general $B(n) = P(2^n)$ and has 2^{2^n} elements.

Aside: Let A be a kind of algebra; write $A(n)$ for the free [A-algebra] on n generators (we can think of this as the space of polynomials in these generators). Then for any other A-algebra R, each $[t(x_1, \dots, x_n)] \in A(n)$ gives us a map $R^n \to R$, so it gives a homomorphism of algebras $A(n) \to$ the ring of maps $R^n \to R$ with pointwise algebra structure. For Boolean algebras this is an isomorphism.

Example: Say we have a map $2^3 \to 2$ by $(T, \perp, T), (\perp, T, \perp), (T, T, \perp), (T, T, T)$ each $\mapsto T$ [the others $\mapsto \perp$]; then the equivalent Boolean expression is $(p \wedge \bar{q} \wedge r) \vee (\bar{p} \wedge q \wedge \bar{r}) \vee \dots$.

The propositional calculus is based on the free Boolean algebra on a countable set $\{p_0, p_1, \dots\} = B(\mathbb{N}) = B(\omega)$; we study it via homomorphisms $v : B(\omega) \to 2$; by freeness such a v is determined by the function $v : \{p_0, p_1, \dots\} \to 2 = \{T, \perp\}$ (and conversely any such function determines a HM). Such a v is called a valuation; we consider it as saying for each proposition p_i whether p_i is true ($\mapsto T$) or false ($\mapsto \perp$).

A HM $v : B \to 2$ is determined by either $v^{-1}(T)$, which is a prime filter, or $v^{-1}(\perp)$ which is a prime ideal (A filter is a $\Phi \subset B$ such that $t \in \Phi$, $a \wedge b \in \Phi \forall a, b \in \Phi$, and if $b \in \Phi$ and $a \geq b$ then $a \in \Phi$; such a Φ is prime iff $0 \in \Phi$ and if $a \vee b \in \Phi$ then $a \in \Phi$ or $b \in \Phi$).

The completeness theorem for propositional calculus amounts to: suppose $\Gamma \subset B(\omega)$ and $A \in B(\omega)$, then $A \in \mathrm{Fil}(\Gamma)$, the filter generated by Γ, iff whenever $v : B(\omega) \to 2$ is such that $v(C) = T \forall C \in \Gamma$, then $v(A) = T$: for the forward implication if $\Gamma \subset v^{-1}(T)$ then since the latter is a filter $\mathrm{Fil}(\Gamma) \subset v^{-1}(T)$, as a sketch of the reverse suppose $A \notin \mathrm{Fil}(\Gamma)$; by ZL take a maximal filter $\Phi \supset \Gamma$ with $A \notin \Phi$; this maximal filter will be prime, so corresponds to a $v : B(\omega) \to 2$ with $v(A) = \perp, v(C) = T \forall C \in \Gamma$, a contradiction.

4.3 Propositional calculus: semantic entailment

We start with a countable set $\{p_0, p_1, \dots\}$ of atomic propositions; from this we form a set Prop of all propositions: $\perp \in$ Prop ($\perp \neq p_i \forall i$, if $A, B \in$ Prop then there is another element $A \to B \in$ Prop (this will actually be the proposition that A implies B).

Given a valuation $v : \{p_0, p_1, \dots\} \to \{T, \perp\}$ we extend it to Prop by $v(\perp) = \perp$ and $v(A \to B) = \perp$ if $v(A) = T, v(B) = \perp$, and T otherwise.

Definition: For $\Gamma \subset$ Prop and $A \in$ Prop we write $\Gamma \models A$, "semantically entails", just when $\forall v$ if $v(c) = T \forall c \in \Gamma$ then $v(A) = T$. An A such that $\models A$ (i.e. $\emptyset \models A$) is called a tautology.

The traditional picture is to say → is determined by its truth table ; this works essentially because $B(2) \simeq$ the space of maps $2^2 \to 2$.

$A \to B$	$B = T$	$B = \perp$
$A = T$	T	$\perp$
$A = \perp$	T	T

From the Boolean algebra primitives we could define $a \to b$ by $\neg a \vee b$; conversely given $\to$ and $\perp$ we can define Boolean algebra by $\neg a = a \to \perp, a \vee b = \neg a \to b$ $(= (a \to \perp) \to b)$, $a \wedge b = \neg(\neg a \vee \neg b)$.

4.4 Proposition calculus: syntactic entailment

We want to define a relation $\Gamma \vdash A$ to mean "there is a proof of A from Γ"; we will in fact find $\Gamma \vdash A$ just when $A \in \mathrm{Fil}(\Gamma)$. We define this by giving 1) Axioms: $\Gamma \vdash$ each of these axioms: $A \to (B \to A)$, $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$, and $\neg\neg A \to A$ (the $\neg$ being just notation, by the above). 2) Natural stipulation: if $c \in \Gamma$ then $\Gamma \vdash c$ 3) Rule of inference (Modus Perens or MP): if $\Gamma \vdash A \to B$ and $\Gamma \vdash A$ then $\Gamma \vdash B$.

"Unravelling" this definition, $\Gamma \vdash A$ iff there is a sequence $A_0, A_1, \dots, A_n = A$ such that for each i either A_i is an axiom or in Γ, or there are $j, k < i$ such that $A_j = A_k \to A_i$ (it is possible to use this as the definition of a proof).

Lemma: $\vdash A \to A$: by axiom 2, $(A \to ((A \to A) \to A) \to ((A \to (A \to A)) \to (A \to A))$; by axiom 1, $A \to ((A \to A) \to A)$, so by MP $(A \to (A \to$

$A)) \to (A \to A)$; by axiom 1 $A \to (A \to A)$, so by MP $A \to A$.

Observe that if $\Gamma \vdash A \to B$ then $\Gamma, A \vdash A \to B$ and $\Gamma, A \vdash A$, so by MP $\Gamma, A \vdash B$ [Γ, A means Γ together with A; formally speaking $\Gamma \cup \{A\}$].

Deduction theorem: if $\Gamma, A \vdash B$ then $\Gamma \vdash A \to B$: induction on the (length of the) proof of $\Gamma, A \to B$: in the case B is an axiom or $\in \Gamma$ we have $\Gamma \vdash B$, but $\Gamma \vdash B \to (A \to B)$ by axiom 1, so by MP $\Gamma \vdash A \to B$. In the case B is A we apply the lemma $\Gamma \vdash A \to A$. The last case is where $\Gamma, A \vdash D \to B$ and $\Gamma, A \vdash D$ and we deduce $\Gamma, A \vdash B$; by the induction hypothesis $\Gamma \vdash A \to (D \to B)$ and $\Gamma \vdash A \to D$, but by axiom 2 $\Gamma \vdash (A \to (D \to B)) \to ((A \to D) \to (A \to B))$, and so applying MP twice we have $\Gamma \vdash A \to B$.

Evident property: If $\Gamma, A \vdash B$ and $\Gamma \vdash A$ then $\Gamma \vdash B$; we take a proof of A from Γ and append a proof of B from Γ and A (deleting the ocurrences of A in the latter, so that we don't have redundant repetition of A). A special case: if $A \vdash B, B \vdash C$ then $A \vdash C$.

Lemma: $\bot \vdash A \forall A$: $\vdash \bot \to (A \to \bot)$ i.e. $\vdash \bot \to \neg A$ by axiom 1; putting $\neg A$ for A we have $\vdash \bot \to \neg\neg A \forall A$, so $\bot \vdash \neg\neg A \forall A$; by axiom 3 $\neg\neg A \to A$ so $\neg\neg A \vdash A \forall A$, so $\bot \vdash A \forall A$.

4.5 Soundness

Observation: The rule MP is sound, in the sense that if $v(A \to B) = T$ and $v(A) = T$ then $v(B) = T$, directly from the truth table for $A \to B$.

Observation: Our axioms are sound, in that $v(A) = T \forall T$ for each axiom A: $v(\neg\neg A \to A) = T$ as $v(\neg\neg A) = v(A)$, and use the diagonal of the truth table. For $v(A \to (B \to A))$, were this $\bot$ we would have $v(A) = T$ and $v(B \to A) = \bot$ so $v(A) = \bot$, a contradiction; similarly for the other axiom.

Soundness theorem: If $\Gamma \vdash A$ then $\Gamma \models A$: take a valuation v such that $v(c) = T \forall c \in \Gamma$; then by induction on (length of) proofs, if $\Gamma \vdash A$ then $v(A) = T$: take a valuation v such that $v(c) = T \forall c \in \Gamma$: for the case A is an axiom, we are done by the second observation, for $A \in \Gamma$ the result is true by hypothesis, and if A has followed from an earlier B and $B \to A$, $v(B) = T$ and $v(B \to A) = T$ by the induction hypothesis, so by our first observation $v(A) = T$.

4.6 Completeness

Completeness theorem: If $\Gamma \models A$ then $\Gamma \vdash A$.

Definition: Γ is consistent if $\Gamma \nvdash \bot$.

Clear fact: If Γ is consist and $\Gamma \vdash A$ then Γ, A is also consistent, for if $\Gamma, A \vdash \bot$ then $\Gamma \vdash A \to \bot$, but then since $\Gamma \vdash A$, $\Gamma \vdash \bot$ by MP.

Model existence theorem: If Γ is consistent then there is a valuation v with $v(c) = T \forall c \in \Gamma$; we call such a v a model of Γ (Remarks: this is the result that Γ consistent $\Rightarrow$ Γ has a model, or that if $\Gamma \nvdash \bot$ then $\Gamma \nvDash \bot$): By ZL take Φ a maximal consistent set containing Γ. By maximality Φ is deductively closed, i.e. $\Phi \vdash A \Rightarrow A \in \Phi$. Define a valuation v by $v(p) = T$ if $p \in \Phi$, $\bot$ if $p \notin \Phi$. We claim that $\forall A$, $v(A) = T \Leftrightarrow A \in \Phi$: this is true for atomic propositions by the definition of v. It is true for $\bot$ since $v(\bot) = \bot$ and $\bot \notin \Phi$. So by induction on the structure of formulae, it is sufficient to show $A \to B \in \Phi \Leftrightarrow (A \in \Phi \Rightarrow B \in \Phi)$. For the forward implication suppose $A \to B \in \Phi$, then if $A \in \Phi$ then $\Phi \vdash B$ by MP so $B \in \Phi$; for the reverse, if $B \in \Phi$ we have $\vdash B \to (A \to B)$ and so by MP

$\Phi \vdash (A \to B)$ and $A \to B \in \Phi$; in the case $A \notin \Phi$ then $\Phi, A \vdash \perp$; recall $\perp \vdash B \forall B$ so $\Phi, A \vdash B$ and $\Phi \vdash A \to B$. Now since $\Gamma \subset \Phi$ we have $v(c) = T \forall c \in \Gamma$.

Proof of completeness: Suppose $\Gamma \nvdash A$. Then $\Gamma, \neg A$ is consistent (for if $\Gamma, \neg A \to \perp$ then $\Gamma \vdash \neg A \to \perp$ i.e. $\Gamma \vdash \neg\neg A$, but $\neg\neg A \vdash A$ so $\Gamma \vdash A$). So by model existence, $\Gamma, \neg A$ has a model, i.e. there is a valuation with $v(c) = T \forall c \in \Gamma$ but $v(A) = \perp$, so $\Gamma \nvDash A$.

Consequence: The question "is A provable" is decidable, for $\vdash A \Leftrightarrow \vDash A$, and we can check the latter by checking all valuations on the finite number of letters in A.

Compactness Theorem: if Γ is a set of propositions such that any finite $\Delta \subset \Gamma$ has a model, then Γ has a model: if Γ is inconsistent, i.e. $\Gamma \vdash \perp$, then for some finite $\Delta \subset \Gamma$, $\Delta \vdash \perp$ (as proofs are of finite length, so can use only finitely many elements of Γ as hypotheses). So if all [finite] $\Delta \subset \Gamma$ are consistent then Γ is consistent; if all [finite] $\Delta \subset \Gamma$ have models then they are consistent by soundness, so Γ is consistent, so Γ has a model by model existence.

Application: Let $(X, \leq)$ be a poset and take our atomic propositions to be p_{xy} for $x, y \in X$. Consider $\Gamma = \{p_{xy} : x \leq y \in X\} \cup \{\neg(p_{xy} \wedge p_{yx}) : x \neq y\} \cup \{p_{xy} \wedge p_{yz} \to p_{xz} : x, y, z\} \cup \{p_{xy} \vee p_{xy} : x, y\}$. Consider $\Delta \subset \Gamma$ finite; then there is $Y \subset X$ finite such that propositions $\in \Delta$ mention only elements of Y. Any finite partial order extends to a total order (by induction); applying this to $(Y, \leq)$ gives a model for Δ. So all finite subsets of Γ have a model, so Γ has a model, which gives a total order on X extending $\leq$.

An aside: why is compactness so called? Write $\Gamma \vDash \Delta$ for: whenever v is a valuation making all of Γ true, it makes one of Δ true. Note $\Gamma \vDash \Delta$ iff $\Gamma, \neg\Delta \vDash$ (by which we mean $\vDash \perp$). Equivalently $\Gamma \vDash \Delta$ iff $\vDash \neg\Gamma, \delta$. Now consider the valuations $v : \{p_0, \dots\} \to \{T, \perp\} = 2$ as the points of a space ($2^{\mathbb{N}}$); consider the propositions A as basic open sets in a topology where $v \in A$ iff $v(A) = T$. Then $\vDash \Gamma$ says that Γ covers the space, and the compactness theorem becomes: if $\vdash \Gamma$ then $\vdash \Delta$ for some finite $\Delta \subset \Gamma$, i.e. that the space is compact.

Another aside: we could have shown model existence by enumerating $p_0, p_1, \dots$ and adding them to Γ just when they are consistent with what we already have.

5 Predicate Calculus

5.1 Terms and equational logic

A signature Σ consists of a set of function simbols f with associated arities $\#f \in \mathbb{N}$; constants are of arity 0. E.g. for groups we have e (of $\# = 0$), $\cdot$ (of $\# = 2$) and $^{-1}$ (of $\# = 1$).

Take a (countable) set V of variables, then we define the set Terms(V) of terms in V from the signature Σ by recursion: each $x \in V$ is a term, and if $f \in \Sigma$ with $\#f = n$ and $t_1, \dots, t_n$ are terms then $f(t_1, \dots, t_n)$ is a term.

Write Terms($\vec{x}$) for the terms whose variables lie in $\vec{x} = x_1, \dots, x_n$. Terms($\emptyset$) are the closed terms; if there are no constants then there are no closed terms. A typical term might look like e.g. $((x \cdot y^{-1}) \cdot (zz^{-1})^{-1})^{-1}$ in groups.

A structure for a signature σ consists of a set A and, for each $f \in \Sigma$ with $\#f = n$, an n-ary function $[\![f(\vec{x})]\!] : A^n \to A$ (the reader is entitled to call this just f if they do not find doing so confusing [arguably I should have done this]). We can extend this, evidently, to an interpretation of terms $t \in$ Terms($\vec{x}$); then we have $[\![t(\vec{x})]\!] : A^n \to A$; we write this as $\vec{a} \in A^n \mapsto [\![t(\vec{a})]\!]$ (we tacitly extend the signature with constants $\forall a \in A$) (Note we allow "dummy" variables; t need not depend on all the variables $x_1, \ldots, x_n$).

Equational logic is concerned with deductions between equations. We take a relation symbol of $=$, and equations are $t = s$ where t, s are terms. Given a structure A we say $A \vDash t = s$, i.e. $t = s$ is true in A, just when $[\![t(\vec{x})]\!] = [\![s(\vec{x})]\!]$ as functions $A^n \to A$, where $\vec{x}$ includes all the variables in s, t (and possibly more; this is irrelevant).

Note: If there are no constants in Σ, then $A = \emptyset$ (with the inevitable choice of interpretation) is a valid structure. For it, $A \vDash t = s \forall t, s$.

Let Γ be a set of equations and $t = s$ an equation in Σ. We say $\Gamma \vDash t = s$ just when $\forall$ structures A for Σ, if $A \vDash u = v \forall u, v \in \Gamma$ then $A \vDash t = s$.

Aside: this is a logic of equations though of as universally quantified, i.e. $t = s$ really means $t = s \forall x_1, \ldots, x_n$, and so on.

Example: the familiar equations $x \cdot (y \cdot z) = (x \cdot y) \cdot z, x \cdot x^{-1} = e$; note that in the second we have different numbers of variables on both sides, so we must take a "dummy" variable on the right hand side.

Equational logic is given by the axiom $t = t$ and rule: if $t = s$ and $u(s) = v(s)$ then $u(t) = v(t)$ (the lecturer writes A and $B \Rightarrow C$ as $\frac{AB}{C}$. $u(s) = v(s)$ means $u(x)[\frac{s}{x}] = v(x)[\frac{s}{x}]$, (this is notation for "$u(x)$ with s substituted in place of x", etc.). Special cases: if $t = s$ and $s = s$ then $s = t$, since $s[\frac{s}{x}] = x[\frac{s}{x}]$. If $t = s$ and $s = r$ then $t = r$.

Inductively from the easy consequence that if $t_1 = s_1, \ldots, t_n = s_n$ then $f(t_1, \ldots, t_n) = f(s_1, \ldots, s_n)$ (†), we only actually need to assume the axiom for the case when t is a variable.

If Γ is a set of equations and $t = s$ an equation in some Σ, we say $\Gamma \vdash t = s$ just when $t = s$ follows from Γ using our axiom and rule.

Soundness is clear: if $\Gamma \vdash t = s$ then $\gamma \vDash t = s$, by induction on the length of the proof.

Completeness theorem: If $\Gamma \vDash t = s$ then $\Gamma \vdash t = s$: let $\vec{x}$ be the variables in $t = s$. Consider Terms$(\vec{x})$ factored out by the equivalence relation $u \sim v$ if $\Gamma \vdash u(\vec{x}) = v\vec{x})$. Define for each $f \in \Sigma$, $[\![f]\!]([t_1], \ldots, [t_n]) = [f(t_1, \ldots, t_n)]$ (the brackets here have their normal meaning of $[x]$ is the equivalence class of x); this is well defined by (†). Now in this structire all the equations of Γ hold, so if $\Gamma \vDash t = s$ then $t = s$ holds and so $\Gamma \vdash t = s$.

$\Gamma \vdash t = s, \Gamma \vDash t = s$ both hold just if $t = s$ is true when evaluated at the elements $[x_1], \ldots, [x_n]$ in the free model for Γ generated by $\{x_1, \ldots, x_n\}$.

Example: the group axioms together with $x^2 = e \vdash xy = yx$; this is the case because the free group on two generators a, b say, together with $x^2 = e$, is $\{e, a, b, ab\} = C_2 \times C_2$.

Some of the notation in the previous section confused some members of the lecture audience; $u[\frac{s}{x}]$ denotes the result of substituting s for x in u, $[\![f]\!] : A^n \to A_n$ is an interpretation of the function symbol f, $[\![t(\vec{x})]\!] : A^n \to A_n$ is an interpretation of the term t, where the variables $\vec{x} = x_1, \ldots, x_n$ include all the variables of t (but possibly also some dummy variables). We write $[\![t(\vec{x})]\!](\vec{a}) = [\![t(\vec{a})]\!]$ for $\vec{a} \in A^n$. Finally $[t(\vec{x})]$ is the equivalence class of a term with variables in the free structure for some equations Γ; we have $[\![f]\!]([t_1(\vec{x})], \ldots, [t_n(\vec{x})]) = [f(t_1, \ldots, t_n)(\vec{x})]$ (of course this is the only possible way to interpret what we have written on the LHS, but this point can be confusing).

5.2 The Language of the Predicate Calculus

A signature for a first order language consists of 1) a functional signature, i.e. a set of function symbols f with arities $\#f \in \{0, 1, 2, \dots\}$ 2) a relational signature, a set of relation symbols R with arities (we almost never see $\#R = 0$; see later) and 3) a special relation symbol $=$, of arity 2.

Examples: 1) No function symbols and one (non-$=$) relation symbol R, $\#R = 2$, e.g. $R =<$ for posets, $R = E(\cdot, \cdot)$ the edge relation for a graph, $R =\in$ for set theory (see later). 2) $0, 1, +, \times, <, \dots$, e.g. for arithmetic, algebra, etc.

Recall: the terms are defined recursively by: if $x \in V$ is a variable then x is a term, if $\#f = n$ and $t_1, \dots, t_n$ are terms then $f(t_1, \dots, t_n)$ is a term.

Formulae of the language: If R is a relation symbol (including $=$) of arity n and $t_1, \dots, t_n$ are terms then $R(t_1, \dots, t_n)$ is an atomic formula (Note that this depends very much on what the terms are; e.g. in example 1 the atomic formulae (up to change of variables) are $R(x, y)$, $R(x, x)$, $x = y$ and $x = x$; in example 2 e.g. $(1 + 1)(xy) < xx + yy$ is an atomic formula). The atomic formulae are all formulae, $\bot$ is a formula and if ϕ, ψ are formulae then $(\phi \to \psi)$ is too. Finally if x is a variable and ϕ a formula then $\forall x\phi$ is a formula.

The new feature here is $\forall x$; in $\forall x$ we say that x is bound. E.g. in a poset, $\forall xy \leq x$ says something about y, but says nothing about x (wheras $y \leq x$ says something about both x and y). This is the same as $\forall zy \leq z$; a change of bound variable (which is called an α-equivalence for historical reasons) makes no difference. But this is not the same as $\forall yy \leq y$; we can change bound variables so long as we avoid capture. $\forall yz \leq y$ says the same thing about z as $\forall xy \leq x$ says of y (a possibly useful analogy is with $\int f(x, y)dx$, which is a function of y and not of x; it $= \int f(z, y)dz$ and is the same function of y as $\int f(y, z)dy$ is of z.

We define the free variables $FV(\phi), FV(t)$ of formulae and terms: for terms, $FV(x) = \{x\}, FV(f(t_1, \dots, t_n)) = FV(t_1) \cup \dots \cup FV(t_n)$. For atomic formulae $FV(R(t_1, \dots, t_n)) = FV(t_1) \cup \dots \cup FV(t_n)$; for other formulae $FV(\bot) = \emptyset, FV(\phi \to \psi) = FV(\phi) \cup FV(\psi), FV(\forall x\phi) = FV(\phi) \setminus \{x\}$ (Examples: $FV(\forall x \bot) = FV(\forall xx = x) = \emptyset$, $FV(\forall xx < y) = \{y\}$, $FV(x < y \to \forall x\neg x = x) = \{x, y\}$ (we would usually write this last term as e.g. $x < y \to \forall z\neg z = z$ to be less confusing)).

Recall that we get all the Boolean operations $\wedge, \vee, \neg$ from $\to$ and $\bot$, and we have normal forms which arose from the de Morgan laws and the cancellation law (that $\neg\neg A, A$ are equivalent). In particular, $A \vee B$ is equivalent to $\neg(\neg A \wedge \neg B)$. By analogy we define $\exists x\phi(x)$ to mean $\neg(\forall x(\neg\phi(x)))$.

5.3 Modules and Satisfaction

Let $\mathcal{L}$ be a first order language. A structure $\mathcal{M}$ for $\mathcal{L}$ consits of a set $\mathcal{M}$ together with: i) for each function symbol f with arity n, an n-ary function $\llbracket f \rrbracket : \mathcal{M}^n \to \mathcal{M}$ ii) for each relation symbol R with arity n, an n-ary relation $\llbracket R \rrbracket \subset \mathcal{M}^n$ or $\llbracket R \rrbracket : \mathcal{M}^n \to \{T, \bot\}$ (Note we shall always take the $=$ symbol to mean "honest" equality in $\mathcal{M}$, i.e. $=$ is $\{(a, a) : a \in \mathcal{M}\} \subset \mathcal{M}^2$.

Recall that we extend the interpretation of a function symbol to an interpretation of terms with possibly dummy variables: $\llbracket t(\vec{x}) \rrbracket : \mathcal{M}^n \to \mathcal{M}$. We extend this to an interpretation $\llbracket \phi(\vec{x}) \rrbracket$ of a formula ϕ where the variables $\vec{x} = x_1, \dots, x_n$ include $FV(\phi)$ (and possibly some dummies) by $\vec{a} \in \llbracket R(t_1, \dots, t_k) \rrbracket$ iff

$(\llbracket t_1(\vec{a})\rrbracket,\dots,\llbracket t_k(\vec{a})\rrbracket) \in \llbracket R\rrbracket \subset \mathcal{M}^k$ (or $\llbracket R(t_1,\dots,t_k)\rrbracket = \llbracket R\rrbracket(\llbracket t_1(\vec{a})\rrbracket,\dots,\llbracket t_k(\vec{a})\rrbracket)$), $\llbracket\perp\rrbracket(\vec{a}) = \perp, \llbracket\phi\to\psi\rrbracket(\vec{a}) = T$ iff $\llbracket\phi(\vec{a}) = T$ implies $\llbracket\psi(\vec{a})\rrbracket = T$. $\llbracket\forall x\phi(x)\rrbracket(\vec{a}) = T$ iff $\forall c \in \mathcal{M}\llbracket\phi(c,\vec{a})\rrbracket = T$. This is sometimes called Tarski's definition of truth; in philisophy it is sometimes talked about as "the grass is green" is true if and only if the grass is green.

For a formula ϕ with $FV(\phi) \subset \vec{x}$ in a language $\mathcal{L}$ and a structure $\mathcal{M}$ for $\mathcal{L}$, we have an interpretation $\llbracket\phi(\vec{x})\rrbracket \subset \mathcal{M}^n$. We say that $\mathcal{M}$ satisfies $\phi(\vec{a})$, $\mathcal{M} \vDash \phi(\vec{a})$, just when $\vec{a} \in \llbracket\phi(\vec{x})\rrbracket$.

If Γ is a set of formulae with free variables $\subset \vec{x}$ then we say that $\vec{a} \in \mathcal{M}^n$ satisfies all Γ when $\mathcal{M} \vDash \gamma(\vec{a})\forall\gamma \in \Gamma$. We are interested particularly in the case where $\vec{x}$ is empty so Γ consists of sentences, i.e. $FV(\gamma) = \emptyset\forall\gamma \in \Gamma$.

Note: If γ is a sentence then $\llbracket\gamma\rrbracket$ is either T or $\perp$ (i.e. $T \sim \mathcal{M}^0 = 1, \perp\sim \emptyset \subset \mathcal{M}^0$).

We have a general notion of semantic entailment: Take Γ a set of forumlae and ϕ a formula with all the free variables $\subset \vec{x}$. Then Γ semantically entails ϕ $\Gamma \vDash \phi$ just when for any structure $\mathcal{M}$ and $\vec{a} \in \mathcal{M}^n$, if $\mathcal{M} \vDash \gamma(\vec{a})\forall\gamma \in \Gamma$ then $\mathcal{M} \vDash \phi(\vec{a})$. A nuance: unless $\vec{x}$ is empty, the empty structure (which exists just when there are no constants) need not be considered; however, in the case where "everything" is a sentence, which we are most interested in, the empty model does need to be considered.

Examples: 1) $\forall x,y,z x\cdot(y\cdot z) = (x\cdot y)\cdot z$ ($\forall x,y,z$ is just notation for $\forall x\forall y\forall z$), $\forall x x\cdot e = x \wedge e\cdot x = x$, $\forall x x\cdot x^{-1} = e \wedge x^{-1}\cdot x = e$: this holds exactly in a group. We call this set of sentences Groups. 2) Groups together with $x \neq e, x^2 \neq e, x^3 \neq e, \dots$ (x^n is just notation, with the obvious meaning). This will be interpreted in groups with an element of infinite order. Note that we are not saying $\exists x : x \neq e \wedge x^2 \neq e \wedge \dots$; we cannot do this, since we do not have infinite conjunctions. 3) $\forall x,y x \le y \wedge y \le x \Rightarrow x = y, \forall x,y,z x \le y \wedge y \le z \to x \le z$ - posets. 4) $\forall x\neg E(x,x), \forall x,y E(x,y) \to E(y,x)$ - graphs.

When we have a class of structures for $\mathcal{L}$ such that there is a set of sentences Γ such that this class is exactly the set of all structures in which Γ is true, we say the class is axiomatizable in first order logic.

Some structures in mathematics seem canonical: 1) $(\mathbb{N},0,1,S,+,\times)$. PA (Peano arithmetic) is the collection of sentences: $\forall x 0 \neq Sx, \forall x,y Sx = Sy \to x = y, \forall x x+0 = x, \forall x,y x+Sy = S(x+y), \forall x x\times 0 = 0, \forall x,y x\times Sy = x\times y+x$, and $\forall\vec{x} = x_1,\dots,x_n$ $((\phi(0,\vec{x}) \wedge \forall y\phi(y,\vec{x}) \to \phi(Sy,\vec{x})) \to \forall y\phi(y,\vec{x}))$. But PA (and its consequences) are not the collection of all things true in arithmetic; in particular PA is incomplete, i.e. there are ϕ such that PA neither entails ϕ nor $\neg\phi$.

5.4 Applications of completeness and compactness

There is a notion of proof of ϕ from Γ (finitely, as for the propositional calculus), and we have analagous theorems: $\Gamma \vdash \phi$ (disregarding empty model issues for now):

Completeness theorem: If $\Gamma \vDash \phi$ then $\Gamma \vdash \phi$

Model existence: If $\Gamma \nvdash \perp$ then Γ has a model

Compactness theorem: If Γ is such that all finite $\Delta \subset \Gamma$ have models then so does Γ

The completeness theorem extends to say that if $\mathcal{L}$ is countable then Γ will have a countable model - and if it is denumerable then there are models of arbitrarily large cardinality (see later).

Application: Completeness and decidability. Γ is complete iff $\forall$ sentences ϕ, $\Gamma \vdash \phi$ or $\Gamma \vdash \neg\phi$. If we can computably enumerate Γ then the notion of proof gives a way to generate consequences computationally, and so we can effectively decide whether ϕ follows from Γ or not: generate consequences until either ϕ or $\neg\phi$ "turns up" (of course this assumes Γ is consistent).

Example: Let Γ be $\{\forall x_1, \dots, x_n \exists y : x_1 \neq y \wedge \cdots \wedge x_n \neq y\}$ i.e. $\forall \vec{x} \exists y : \bigwedge_{i=1}^{n} x_i \neq y$ (call this proposition η_i for later). Suppose Γ is not complete. Then we have ϕ such that Γ, ϕ and $\Gamma, \neg\phi$ are both consistent. Each of them has a countable model, but any two such are isomorphic, so we cannot have ϕ true in one and false in the other.

Note: for any $\mathcal{M}$, $\{\phi | \mathcal{M} \vDash \phi\}$ is a complete theory.

Examples: Let DLO (dense linear orders) be total orders plus $\forall x, yx < y \to \exists z : x < z < y, \neg\exists x : \forall yx \geq y, \neg\exists x : \forall yy \leq x$. In fact this has (up to isomorphism) a unique countable model $(\mathbb{Q}, <)$.

ACF_0 is the theory of algebraicly closed fields of characteristic 0. Forc any $\kappa > \omega$ there is (up to isomorphism) a unique model of this cardinality. It follows that this theory is complete and so decidable.

Application: Axiomatizability: The class of finite groups is not axiomatizable: suppose Γ was such a theory. Consider $\Pi = \Gamma \cup \{\eta_n : n = 1, 2, \dots\}$ where η_i is as above. Any finite subset Δ of Π contains only finitely many η_n, so if $m >$ all the ns in question then C_m is a model of Δ. So by compactness Π has a model, an infinite finite group, a contradiction.

Application: Existence of non-standard models. Let Th($\mathbb{N}$) be the collection of all sentences true in $(\mathbb{N}, 0, 1, s, +, \times)$ (s being successor). It is complete and consistent, but very complicated - more so than PA, and even PA is not "decidable" - there is no algorithm for deciding $PA \vDash \phi$. Extend the language with a new constant ∞. Considre $\Gamma = \mathrm{Th}(\mathbb{N}) \cup \{\infty \neq 0, \infty \neq 1, \infty \neq 2, \dots\}$ (of course 2 is just notation for $s(1)$, etc.); any finite $\Delta \subset \Gamma$ contains finitely many statements "$\infty \neq n$", so if N is greater than all the n for which $\infty \neq n$ appears in Δ, Δ has a model (namely $\mathbb{N}$ where ∞ is interpreted to be N). So Γ has a model, a non-standard model of true arithmetic.

5.5 Proofs in the predicate calculus

Generally we have proofs with formulae with free variables. A way to think about this is to assume the variables have been "declared", e.g. "let p be a prime number". So there is a tacit assumption that our structure is non-empty. For this course (the notation is not standard) we will write $\Gamma \vdash \phi$ if ϕ follows from Γ assuming Γ non-empty, and $\Gamma \overset{0}{\vdash} \phi$ when we also consider the empty case (this only makes sense for sentences) (The formal notation for these would be $\overset{\vec{x}}{\vdash}$ and $\vdash$ respectively.

Axioms: $\phi \to (\psi \to \phi), (\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi)), \neg\neg\phi \to \phi$ as before. $\forall xx = x, \forall x, yx = y \to (\phi \to \phi[\frac{y}{x}])$ (where this is "substitution without capture"; if y appears elsewhere in ϕ we must change that before substituting). $\forall x(\phi \to \psi) \to \phi \to (\forall x\psi)$ so long as $x \notin FV(\phi)$. $\forall x\phi \to \phi[\frac{t}{x}]$ ([for any term t] again without capture). All the axioms other than the last are true for $\vdash$ and $\overset{0}{\vdash}$ wherever they make sense; the last is always true for $\vdash$ but only true for $\overset{0}{\vdash}$ as long as t and $\forall x\phi(x)$ are closed.

Rules of inference: $\frac{\phi\phi\rightarrow\psi}{\psi}$ (MP as before). Suppose that x is not free in any assumption in the proof of $\phi(x)$, then $\frac{\phi(x)}{\forall x\phi(x)}$ (this is called Generalization or Gen; it introduces $\overset{0}{\vdash}$ when it can). We say $\Gamma \vdash \phi$ (or $\Gamma \overset{0}{\vdash} \phi$) just when there is a sequevce $\phi_1, \ldots, \phi_n = \phi$ with each ϕ_1 either an axiom, in Γ or following from earlier ϕ_j, ϕ_k by the rules of inference.

Nuances: $\vdash \forall x \bot \rightarrow \bot$ $(= \bot\, [\frac{y}{x}])$, but this does not hold for $\overset{0}{\vdash}$. We say that Γ is consistent when $\overset{(0)}{\nvdash} \bot$ (i.e. $\Gamma \overset{0}{\nvdash} \bot$ if Γ [consists of] sentences, $\Gamma \nvdash \bot$ otherwise). $\Gamma \nvdash \bot$ means Γ is consistent with the assumption $\exists x : x = x$, so $\forall x x \neq x$ is consistent in a language with no constants.

We have the usual propositional calculus theorems. Write $\overset{\cdot}{\bot}$ to mean a theorem holds for both $\vdash$ and $\overset{\cdot}{\vdash}$.

Simple ones: Deduction theorem: $\Gamma, \phi \overset{\cdot}{\vdash} \phi \Rightarrow \Gamma \overset{\cdot}{\vdash} \phi \rightarrow \psi$. Soundness theorem: If $\Gamma \overset{\cdot}{\vdash} \phi$ then $\Gamma \overset{\cdot}{\models} \phi$.

Lemma: Suppose $\exists x\phi(x)$ is consistent with Γ (i.e. $\{\exists x\phi(x)\}\cup\Gamma$ is consistent). Then so is $\phi(\vec{c})$ where $\vec{c}$ is a new constant: suppose not. Then $\Gamma, \phi(\vec{c}) \vdash \bot$; replace c by some free (new) variable x, then $\Gamma, \phi(x) \vdash \bot$ so $\Gamma \vdash \neg\phi(x)$ (by the deduction theorem), so $\Gamma \overset{(0)}{\vdash} \forall x\neg\phi(x)$ so $\Gamma, \neg\forall x\neg\phi(x) \vdash \bot$ i.e. $\Gamma, \neg\exists x\phi(x) \vdash \bot$.

The completeness theorem: if $\Gamma \overset{(0)}{\models} \phi$ then $\Gamma \overset{(0)}{\vdash} \phi$. As in the propositional calculus, this follows from:

Model existence: If $\Gamma \overset{(0)}{\nvdash} \bot$ then Γ has a model (i.e. $\Gamma \nVdash \bot$ [presumably there should be a (0) there, but I don't care enough to check]), which also proves the compactness theorem.

The remainder of this section (5) is nonexaminable.

Background: A substructure $\mathcal{N} \hookrightarrow \mathcal{M}$ of a structure $\mathcal{M}$ (for some language) is given by an $N \subset M$ [some of my earlier $\mathcal{M}$ should have been M, but the lecturer was not terribly clear in his distinctions between them] closed under the defined functions (and so with corresponding interpretations $\llbracket f \rrbracket_{\mathcal{N}} = \llbracket f \rrbracket_{\mathcal{M}} |_N : N^n \rightarrow N$) and with the restricted interpretation of the relation symbols $\llbracket R \rrbracket_{\mathcal{N}} = \llbracket R \rrbracket_{\mathcal{M}} \cap N^n$.

1) Suppose $\phi(\vec{x})$ is a QF (quantifier free) formula. Then if $\mathcal{N} \hookrightarrow \mathcal{M}$ and $\vec{a} \in N^n$¡ then $\mathcal{N} \models \phi(\vec{a})$ iff $\mathcal{M} \models \phi(\vec{a})$ (aside: $\mathcal{N} \hookrightarrow \mathcal{M}$ is said to be an elementary embedding if this holds for all ϕ. This is terrible notation, but sadly well established). 2) Suppose $\forall\vec{y}\phi(\vec{x}, \vec{y})$ is a formula, ϕ QF. Then $\mathcal{M} \models \forall y\phi(\vec{a}, \vec{y}) \Rightarrow \mathcal{N} \models \forall y\phi(\vec{a}, \vec{y})$. 3) Suppose $\mathcal{M}$ is a structure for $\mathcal{L}$. Then $\mathcal{M}$ has a minimal substructure $\mathcal{M}_0$ by $\mathcal{M}_0 =$ the set of $\llbracket t \rrbracket_{\mathcal{M}}$ for closed terms t of $\mathcal{L}$. Examples: The minimal substructure of $\mathbb{R}$ as an ordered field is $\mathbb{Z}$ (multiplicative inversion is not a function symbol, since we can't apply it to 0), the minimal substructure of any graph is the empty graph. 4) Let $\mathcal{M}$ be a structure. Define the theory of $\mathcal{M}$ $\mathrm{Th}(\mathcal{M})$ to be the set of sentences ϕ with $\mathcal{M} \models \phi$. This is consistent, deductively closed, and complete; $\phi \rightarrow \psi \in \mathrm{Th}(\mathcal{M}) \Leftrightarrow (\phi \in \mathrm{Th}(\mathcal{M}) \Rightarrow \psi \in \mathrm{Th}(\mathcal{M}))$. 5) If in addition $\mathcal{M} = \mathcal{M}_0$ then $\forall x\phi(x) \in \mathrm{Th}(\mathcal{M})$ iff $\phi(t) \in \mathrm{Th}(\mathcal{M})\forall$ closed [terms] t (†). 6) If T is a collection of sentences in $\mathcal{L}$ satisfying † then there is a model for T, $\mathcal{M}$ (with $\mathcal{M} = \mathcal{M}_0$): set M to be the quotient of the set

of closed terms t by the relation $\sim$ where $t \sim s$ iff $t = s \in T$; then define $[\![f]\!]([t_1], \ldots, [t_n]) = [f(t_1, \ldots, t_n)]$ and similarly $[\![R]\!]$ (considered as a function $M^n \to \{T, \perp\}$).

Given Γ consistent in $\mathcal{L}$, we extend to $\bar{\Gamma}$ with the properties in 6), in a language $\bar{\mathcal{L}}$, with the property that $\forall x\phi(x) \in \bar{\Gamma}$ iff $\phi(\vec{c}) \in \bar{\Gamma}\forall$ constants c (equivalently, $\exists x\phi(x) \in \bar{\Gamma}$ if $\phi(c) \in \bar{\Gamma}$ for some c: first we take a maximal consistent extension Γ_1 of Γ, then for all sentences $\exists x\phi(x) \in \Gamma_1$ add a constant c_ϕ and $\phi(c_\phi)$ giving $\bar{\Gamma}_1$ in a new language $\mathcal{L}_1$; this $\bar{\Gamma}$ is still consistent by an earlier lemma. Repeat, obtaining $\Gamma_2, \bar{\Gamma}_2$ in $\mathcal{L}_2$ and so on, then let $\bar{\Gamma} = \bigcup_i \bar{\Gamma}_i$ in $\bar{\mathcal{L}} = \bigcup_i \mathcal{L}_i$ and we are done.

6 Formal set theory

6.1 The hierarchy V_α of pure sets

Define $V : On \to$"Sets" by recursion: $V_0 = 0$ (the empty set $\emptyset$), $V_{\alpha+1} = P(V_\alpha$, $V_\lambda = \bigcup_{\beta<\lambda} V_\beta$ for λ a limit.

Early stages: $V_0 = 0, V_1 = \{0\}, V_2 = \{0, \{0\}\}, V_3 = \{0, \{0\}, \{\{0\}\}, \{0, \{0\}\}\}$. The cardinalities are 0 for V_0, $2^0 = 1$ for V_1, $2^1 = 2$ for V_2, $2^2 = 4$ for V_3, $2^4/16$ for V_4, $2^{16} = 65536$ for V_5 and so on.

Every set is a set of sets; this is what is meant by pure sets.

Definition (important): A (pure) set x is transitive just when $z \in y \in x \Rightarrow z \in x$ $(\forall y, z)$.

Two equivalent interpretations: 1) $\{z : \exists y \in x : z \in y\} = \cup\{y : y \in x\} = \cup x$. So x is transitive iff $\cup x \subset x$.

Lemma: Suppose $\{x_i : i \in I\}$ is an indexed family of transitive sets. Then $\bigcup_i \{x_i : i \in I\}$ is transitive, either obviously, or with our bare hands, or because $\cup \bigcup_i \{x_i : i \in I\} = \bigcup_i \{\cup x_i : i \in I\} \subset \bigcup_i \{x_i : i \in I\}$.

2) $z \in y \in x \Rightarrow z \in x$ is the statement that $y \in x \Rightarrow y \subset x$. So x is transitive iff $x \subset Px$.

Lemma: If x is transitive so is Px: if $x \subset Px$ then $Px \subset PPx$.

Proposition: For all α, V_α is transitive, by induction on α: 0 is transitive, and the two lemmas give the induction.

Proposition: For $\alpha \leq \gamma$ we have $V_\alpha \subset V_\gamma$, by induction on $\gamma \geq \alpha$: for $\gamma = \alpha$, $V_\alpha \subset V_\alpha$. For $\gamma = \lambda > \alpha$ a limit, $V_\alpha \subset \bigcup_{\beta<\lambda} V_\beta = V_\lambda$. If $V_\alpha \subset V_\gamma$, then as V_γ is transitive, $V_\alpha \subset V_\gamma \subset P(V_\gamma) = V_{\gamma+1}$.

Let $V = \bigcup_{\alpha \in On} V_\alpha$. Just as On is not itself an ordinal, V is not itself a pure set (= element of some V_α).

We shall axiomatize the properties of $(V, \in)$. Sets=members of V; if $x \in V$ then it represents $\{y \in V : V \models y \in x\}$. We also consider classes, subcollections of V defined (with parameters from V).

6.2 Axioms for Set Theory

The official language has (in addition to =) just one binary relation symbol $\in$; for practical purposes we need "definitional extensions": 1) Suppose $\mathcal{L}$ is a language, T a theory and $\phi(\vec{x})$ a formula. Add to $\mathcal{L}$ a relation symbol R, and to T $\forall \vec{x} R(\vec{x}) \leftrightarrow \phi(\vec{x})$ ($x \leftrightarrow y$ being notation for $x \to y \wedge y \to x$), forming $\mathcal{L}'$ and T'. Then (†) T' proves for any formula in $\mathcal{L}'$ that it is equivalent to a formula in $\mathcal{L}$, and T' proves in $\mathcal{L}$ just what T does, so R is harmless notation. 2) Suppose $\mathcal{L}T$ are as above and $\phi(\vec{x}, y)$ a formula such that $T \vdash \forall x \exists ! y \phi(\vec{x}, y)$, i.e. $T \vdash \forall \vec{x} \exists y \phi(\vec{x}, y) \wedge \forall \vec{x}, y, y'(\phi(\vec{x}, y) \wedge \phi(\vec{x}, y') \rightarrow y = y')$. Add to $\mathcal{L}$ a function symbol $f(\vec{x})$ and to T $\forall \vec{x}, y(f(\vec{x}) = y \leftrightarrow \phi(\vec{x}, y))$. Then † again holds.

"Basic axiom": sets are determined by their members: Extensionality: $\forall x, y(\forall zz \in x \leftrightarrow z \in y) \rightarrow x = y$.

Set existence axioms: Empty set: $\exists z : \forall y \neg y \in z$. By extensionality, such a z is unique and so we can introduce a constant 0 and the axiom becomes $\forall y \neg y \in 0$. Pairing: $\forall x, y \exists z : \forall ww \in z \leftrightarrow w = x \vee w = y$; by extensionality for fixed x, y such a z is unique and so we can introduce a function symbol $\{,\}$ and the axiom becomes $\forall x, y \forall ww \in \{x, y\} \leftrightarrow w = x \vee w = y$.

Pairing enables us to code the notion of an ordered pair: we set $(x, y) = \{\{x\}, \{x, y\}\}$ where $\{x\} = \{x, x\}$. The lecturer claims that of course an ordered pair is still a mathematical primitive, and "by rights" should be axiomatized; this is just a method for coding the notion in set theory; to say this is what an ordered pair "really is" is a bit silly.

Lemma: $(x, y) = (u, v)$ iff $x = u$ and $y = v$; the reverse implication is trivial. For the forward, we have $\{\{x\}, \{x, y\}\} = \{\{u\}, \{u, v\}\}$; either $\{x\} = \{u, v\}$ in which case $x = u = v$ so the RHS is $\{\{u\}\}$ so $\{x, y\} = \{u\}$ so $y = u = v$ and $x = u, y = v$ as required, or $\{x, y\} = \{u\}$ and the same argument holds, or $\{x\} = \{u\}$ and $\{x, y\} = \{u, v\}$, so $x = u$ and either $y = u$ and we argue as before, or $y = v$ and we have the result.

Unions: $\forall x \exists z : \forall ww \in z \leftrightarrow \exists y : w \in y \wedge y \in x$; by extensionality as always we can introduce a function symbol $\cup$ and $\forall x \forall ww \in \cup x \leftrightarrow \exists yw \in y \wedge y \in x$. We can then define $x \cup y = \cup\{x, y\}$.

(Our axioms thus far are absolute, not that we are defining the notion in this course. The next is "less safe")

Power set: $\forall x \exists z \forall yy \in z \leftrightarrow (\forall ww \in y \rightarrow w \in x)$. Introduce a relation symbol $y \subset x$ for $\forall ww \in y \rightarrow w \in x$, then a function symbol P and our axiom is $\forall x(\forall yy \in Px \leftrightarrow y \subset x)$.

Separation: let $\phi(y, \vec{w})$ be a formula. $\forall \vec{w} \forall x \exists z \forall yy \in z \leftrightarrow y \in x \wedge \phi(y, \vec{w})$. By extensionality we can introduce a function symbol $\{y \in x : \phi(y, \vec{w})\}$ (this notation is slightly confusing as y "isn't really there"; it's not acted on by the function, but is rather being used as a dummy), and our axiom becomes $\forall \vec{w}, xy \in \{y \in x : \phi(y, \vec{w})\} \leftrightarrow y \in x \wedge \phi(y, \vec{w})$.

Conceptual explanation: let $\phi(y, \vec{w})$ be a formula and suppose $\vec{a}$ are sets i.e. in V. Then $\{y : \phi(y, \vec{a})\}$ is a <u>class</u>; it might be (represented by) a set, e.g. $\{y : y \neq y\}$ "$=$" 0, and it might not, e.g. $\{y : y = y\} = V$. $\{y : \phi(y, \vec{w})\}$ is a (paramaterized) class. Separation says that if x is a set and A a class then $x \cap A$ is a set.

Example: For x, y sets, $x \times y = \{(a, b) : a \in x, b \in y\}$ is a class. Consider $x \cup y = \cup\{x, y\}$, which has elements of the form a or b. Then $P(x \cup y)$ contains elements of the form $\{a\}, \{a, b\}, \ldots$, so $PP(x \cup y)$ contains elements like $\{\{a\}, \{a, b\}\}$. So $x \times y = PP(x \cup y) \cap x \times y$, which is a set by separation.

Axiom of infinity: set $Sx = x \cup \{x\}$ $(= \cup\{x, \{x\}\})$. $\exists z : 0 \in z \wedge \forall yy \in z \rightarrow Sy \in z$. This is not unique; however, with separation, we can find a unique minimal such: say that z is a $(0, S)$-algebra if $0 \in z \wedge \forall y \in zSy \in z$ ($\forall x \in u\psi$ is notation for $\forall xx \in u \rightarrow \psi$; $\exists x \in u\psi$ means $\exists xx \in u \wedge \psi$). Consider $A = \cap\{z : z \text{ is an } (0, S) \text{ algebra}\}$; this is a class. The axiom gives us some $(0, S)$ algebra z_0; we can consider by separation $z_0 \cap A$, which is A, and A is a $(0, S)$

algebra since all the properties hold for intersections. So A is the minimal $(0, S)$ algebra.

What is A? $0 \in A$, $S0 = 0 \cup \{0\} = \{0\} =: 1 \in A$, $S1 = \{0\} \cup \{1\} = \{0, 1\} =: 2 \in A$, $S2 = \{0, 1, 2\} =: 3 \in A$ and so on, so A is ω. With separation the axiom becomes: $0 \in \omega \wedge \forall y \in \omega Sy \in \omega, \forall z 0 \in z \wedge \forall y \in z Sy \in z \to \omega \subset z$; this second half is the principle of mathematical induction.

Axiom of replacement: intuitively, the image of a set under a definable function (with parameters), which we shall call a Function, is a set. The statement (actually an axiom scheme) is: for any formula $\phi(x, y, z) \forall z \forall x \exists! y \phi(x, y, z) \to \forall u \exists v \forall y y \in v \leftrightarrow \exists x \in u \phi(x, y, z)$. Eqivalently, for any $\phi(x, y, z) \forall z \forall u \forall x \in u \exists! y \phi(x, y, z) \to \exists v \forall y y \in v \leftrightarrow \exists x \in u \phi(x, y, z)$. An alternative formulation: let F be a class. (F is a relation if $\forall z \in F \exists x, y : z = (x, y)$) F is a Function just when $\forall x \exists! y : (x, y) \in F$ (We may insist that F be a relation as well; this is "neater", but not actually relevant). F is a Function on u if $\forall x \in u \exists! y (x, y) \in F$. The Axiom is then $(\forall \vec{z})$ $\forall u F$ a function on $u \to \mathrm{Im} F$ is a set.

ZF (Zermelo-Fraenkel Set Theory) is the full axiom system; Z (Zermelo Set Theory) is the same but without replacement.

6.3 Sets and Classes

In ZF, the variables denote sets; there are no variables for classes, we can talk directly only about sets; the theory refers to classes indirectly and "class by class".

Remarks: 1) If x, y are sets then so is $x \times y$. Then the set of relations $\mathrm{Rel}(x, y) = P(x \times y)$ is a set, and the sets of functions $\mathrm{Fun}(x, y)$ and partial functions $\mathrm{Ptl}(x, y)$ are definable subcollections of Rel so sets by separation. The collection of partial orders or well orders on a set x are definable subcollections of $\mathrm{Rel}(x, x)$, $PO(x)$ and $WO(x)$, and we could define these as function symbols. A partially ordered or well-ordered set is a pair $(x, <)$ where $< \in PO(x)$ or $WO(x)$. The collections of these (i.e. the collection of posets and the collection of well ordered sets) are classes. 2) If X, Y are classes (with parameters, as always) then $X \times Y = \{(a, b) : a \in X, b \in Y\}$ is also a class. But we cannot continue as in 1; $\{B : B \subset A\}$ makes no sense in the language, we can't use a variable B to represent a class. We handle relations from X to Y relation by relation: if R is a class, $\forall z \in R \exists x \in X, y \in Y : z = (x, y)$ says that R is a relation from X to Y. If F is a relation from X to Y, $\forall x \in X \exists! y \in Y : (x, y) \in F$ says F is a Function from X to Y. We can also take R a relation from X to X and say that R well-orders X if (R is a total order and) $\forall x \neq \emptyset x \subset X \exists a \subset x : a$ is R-minimal. This is a "safe" definition just when R is local on X in the sense that $\forall a \in X \{b \in X : bRa\}$ is a set. 3) If A is a class then $P_S(A) = \{x : x \subset A\}$ ($x \subset A$ meaning $\forall y \in x y \in A$ is a class, the class of subsets of A (the S is for "small", since this only includes sets, not classes) (Aside: $P_S(V) = V$, and the argument for Russell's Paradox shows that $\{x : x \notin x\}$ is not in $P_S(V) = V$, i.e. it is not a set). In the same spirit, for X, Y classes we have classes $\mathrm{Rel}_S(X \times Y) = P_S(X \times Y), \mathrm{Ptl}_S(X, Y) = \{\phi \in \mathrm{Rel}_S(X \times Y) : \forall a \in \mathrm{dom}\phi \exists! b : (a, b) \in \phi\}$. Note that if X is a set then we have $\mathrm{Fun}(X, Y)$ a class.

6.4 Recursion theorems

Theorem 1: Let $(x, <)$ be a well ordered set, $g : x \times \mathrm{Ptl}(x, y) \to y$. Then there is a unique $f : x \to y$ such that $\forall a \in x f(a) = g(a, f \mid_{x<a})$. This is a single provable sentence in Zermelo set theory (i.e. it does not use replacement).

Theorem 2: Let $(x, <)$ be a well ordered set, $G : x \times \mathrm{Ptl}_S(x, V) \to V$ a Function. Then there is a unique $f : x \to V$ (a set $f \in \mathrm{Fun}(x, V)$) such that for all $a \in x$ $f(a) = G(a, f \mid_{x<a})$ $(\star)$. G is "reall" a formula with paramaters $\vec{z}$ (so this is a schema: for any G we can prove the result). The theorem is then that $\forall \vec{z} \forall (x, <)(x, <)$ well ordered $\to \exists! f : \forall a \in x \star$ holds.

Theorem 3: Let $(A, <)$ be a local well-ordered class. Let $G : A \times \mathrm{Ptl}_S(A, V) \to V$ be a Function. Then there is a unique Function $F : A \to V$ such that $\forall a \in AF(a) = G(a, F \mid_{A_{<a}})$ $(\dagger)$, i.e. we have a map from Formulae(A, G) to a Formula for F such that $\forall \vec{z}((A, <)$ local well-ordered$) \wedge (G : A \times \mathrm{Ptl}_S(A, V) \to V) \to (F : A \to V) \wedge \dagger$. Uniqueness here means that if we have two such functions, it is a provable theorem in ZF that they are equal.

Proof of theorem 3: Let $\phi \in \mathrm{Ptl}_S(A, V)$ be an attempt if $\mathrm{dom}\phi \subset A$ is a $<$-initial segment and $\forall a \in \mathrm{dom}\phi$ $\phi = G(a, \phi \mid_{A_{<a}})$. We can show by induction, if ϕ, ψ are attempts then $\forall a \in \mathrm{dom}\phi \cap \mathrm{dom}\psi$, $\phi(a) = \psi(a)$. So define F by $F(a) = b$ (i.e. $(a, b) \in F$) if $\exists \phi$: ϕ an attempt and $\phi(a) = b$. It remains to prove that F is defined on all A: if not, take a least where $F(a)$ is not defined. $F \mid_{A_{<a}}$ is an attempt and can be extended to an attempt with domain $A_{\leq a}$ by setting $a \mapsto G(a, F \mid_{A_{<a}})$, so we have a contradiction.

6.5 von Neumann ordinals

Recall the idea that an ordinal is canonically represented by $On_{<a}$; we will make this happen.

Take $(x, <)$ a well-ordered set and define by recursion (using theorem 2) $f : x \to V$: $f(a) = \{f(b) : b < a \text{ in } x\}$. Consider $\mathrm{Im} f$: 1) if $y \in f(a) \in \mathrm{Im} f$ then $y = f(b)$ for some $b < a \in x$ and so $y \in \mathrm{Im} f$. Thus $\mathrm{Im} f$ is a transitive set. 2) $b < a$ in x iff $f(b) \in f(a)$ in $\mathrm{Im} f$. So $f : (x, <) \to (\mathrm{Im} f, \in)$ is an order isomorphism, and $\mathrm{Im} f$ is well ordered by $\in$

Definition: $\mathrm{On} = \{\alpha : \alpha \text{ is transitive and well ordered by } \in\}$.

Observations: 1) If $\gamma \subset \alpha$ is an initial segment of $\alpha \in On$ then $\gamma \in On$. If $\beta \in \alpha$ then $\alpha_{<\beta}$ is $\alpha_{\in \beta} = \{c \in \alpha : c \in \beta\} = \{c : c \in \beta\} = \beta$. Thus every member of α is an ordinal. 2) Put $(\alpha, \in)$ (for $\alpha \in On$) in the above recursion: $f(a) = \{f(b) : b < a \text{ in } \alpha\} = \{f(b) : b \in a\}$. Inductively, $f(a) = \{b : b \in a\} = a$. The same thing happens if we consider $f : \alpha \to \beta$ which is an order isomorphism to an initial segment: $a \in b$ in α iff $f(a) \in f(b)$ in $\mathrm{Im} f$. Then we deduce inductively $f(a) = \{f(b) : b \in a\} = \{b : b \in a\} = a$, so order isomorphisms are just inclusions. Recall that for any two well-ordered sets one is order isomorphic (uniquely) to an initial segment of the other, so if $\alpha\beta \in On$ either $\alpha \subset \beta$ or $\beta \subset \alpha$. If $\alpha \subset \beta$ either it is a proper initial segment and $\beta_{<\alpha} = \alpha$ and $\alpha \in \beta$, or $\alpha = \beta$. So for $\alpha, \beta \in On$, $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$. Thus On is totally ordered by $\in$.

Claim: On is a local well-ordered class under $\in$. For locality, $On_{<\alpha} = \alpha$, a set. Let $X \subset On$ be a nonempty class; pick $\alpha \in X$; either α is $\in$-minimal $(\{\beta \in \alpha : \beta \in X\} = \emptyset)$, or the set $\alpha \cap X$ is $\neq \emptyset$, and by the well orderedness of α we can take β $\in$-minimal in $\alpha \cap X$ which is then $\in$-minimal in X.

6.6 The Axiom of Foundation

E.g. by T3 we have justifications for the definitions of $\alpha+\beta, \alpha\cdot\beta, \alpha^\beta$ by recursion on $\beta \in On$. Equally, we have justified the recursion $V_0 = 0, V_{\alpha+1} = P(V_\alpha), V_\lambda = \bigcup_{\beta<\lambda} V_\beta$ for λ a limit. So $V : On \to (V)$ the universe of sets is a Function; there is a formula which "says" $y = V_\alpha$; there is also a formula $x \in V_\alpha$. Propositions such as $\forall\alpha V_\alpha$ is transitive, $\forall\beta \leq \alpha V_\beta \subset V_\alpha$ are then all theorems of ZF.

Our idea was that every set should be in $V = \bigcup_\alpha V_\alpha$. We need an axiom to make this happen.

Axiom of Foundation: Suppose $A \neq 0$ is a nonempty class (with paramaters, as always). Then A contains an $\in$-minimal element, i.e. $\forall\vec{z}(\exists aa \in A) \to \exists aa \in A \wedge \forall x \in ax \notin A$ (or being more cute we can write the RHS of this as $\exists a \in Aa \cap A = 0$). This says the class V is <u>well-founded</u>; see later. An equivalent formulation is the axiom of $\in$-induction: $(\forall x((\forall y \in x\phi(y)) \to \phi(x))) \to \forall x\phi(x)$; the proof of equivalence is easy but nonexaminable, as it is really dull. Full ZF is usually the system with foundation.

Theorem: It follows from the axioms that $\forall x\exists\alpha x \in V_\alpha$: suppose not, i.e. $A = \{x : \neg\exists\alpha x \in V_\alpha\} \neq 0$. By foundation take $a \in A$ $\in$-minimal, then $\forall x \in ax \in V_\alpha$ for some α, so there is a Function $x \mapsto$ the least α such that $x \in V_\alpha$ which is functional on a. By replacement it follows that there is a set $z \subset On$ such that $\forall x \in a\exists\alpha \in z : x \in V_\alpha$. Let $\gamma = \sup z$, then $\forall x \in ax \in V_\gamma$, i.e. $a \subset V_\gamma$ and so $a \in V_{\gamma+1}$, a contradiction.

Aside: the least α such that $x \in V_\alpha$ is always a successor ordinal. So it is more interesting to consider the previous ordinal.

Definition: The rank $\mathrm{rk}(x)$ of a (pure) set x is the least α such that $x \subset V_\alpha$.

Lemma: $\mathrm{rk}(0) = 0$.

Lemma: Suppose $(x_i : i \in I)$ is a family of pure sets. Then $\mathrm{rk}(\bigcup_i x_i) = \sup_i \mathrm{rk}(x_i) = \bigcup_i \mathrm{rk}(x_i)$ (recall the $rk(x_i)$ are ordinals): $x_i \subset \bigcup_{i\in I} x_i \subset V_{\mathrm{rk}(\bigcup_i x_i)}$ so $\mathrm{rk}(x_i) \leq \mathrm{rk}(\bigcup_i x_i)\forall i$ and $\sup \mathrm{rk}(x_i) \leq \mathrm{rk}(\bigcup_i x_i)$. But also $x_i \subset V_{\mathrm{rk}(x_i)} \subset V_{\sup \mathrm{rk}(x_i)}$ so $\bigcup_i x_i \subset V_{\sup \mathrm{rk}(x_i)}$ so $\mathrm{rk}(\bigcup_i x_i) \leq \sup \mathrm{rk}(x_i)$.

Recall that $S(x) = x \cup \{x\}$ and note that if $\alpha \in On$ then $S(\alpha) = \alpha \cup \{\alpha\} = \alpha + 1$.

Lemma: $\mathrm{rk}(S(x)) = \mathrm{rk}(x) + 1$: $x \subset V_{\mathrm{rk}(x)}$ so $x \in V_{\mathrm{rk}(x)+1}$ and both $x, \{x\} \subset V_{\mathrm{rk}(x)+1}$ and so $S(x) \subset V_{\mathrm{rk}(x)+1}$ and $\mathrm{rk}S(x) \leq \mathrm{rk}(x) + 1$. Also $x \cup \{x\} = Sx \subset V_{\mathrm{rk}S(x)}$ and so $x \in V_{\mathrm{rk}(S(x))}$; it follows (by the definition of V_α) than $x \subset V_\beta$ for some $\beta < \mathrm{rk}(S(x))$ so $\mathrm{rk}(x) < \mathrm{rk}(s(x))$ and $\mathrm{rk}(x)+1 \leq \mathrm{rk}S(x)$ and we are done.

Proposition: $\mathrm{rk}(\alpha) = \alpha$, by induction on On using these lemmas.

6.7 The Recursion Theorem for V

We have a principle of $\in$-induction for V, and so expect to have a corresponding recursion theorem.

The Transitive Closure of a set x

Aside: for pure sets $x \in V_\alpha$ some α and V_α is transitive, and so $x \in V_\alpha \cap \bigcap\{y : y$ transitive,$y \supset x\}$ and this RHS is automatically the largest transitive set t containing x. But for historical reasons we shall generalise (this is allegedly more instructive):

For a set x, define by recursion (using theorem 2) on ω, $T_x : \omega \to V$ by $T_x(0) = x, T_x(n+1) = T_x(n) \cup \bigcup T_x(n)$. By replacement $\{T_x(n) : n \in \omega\}$ is a set, and taking unions, $TC(x) := \bigcup_{n\in\omega} T_x(n)$ is a set.

Claim: $TC(x)$ is the least transitive set containing x. $TC(x)$ is transitive, since for any $z \in y \in TC(x)$, $y \in T_x(n)$ for some n so $z \in T_x(n+1) \subset TC(x)$. For the least part, we need:

Lemma: Let t be transitive, $t \supset z$. Then $\cup z \subset \cup t = t$, so $z \cup (\cup z) \subset t$.

Let $t \supset x$ be transitive, then by induction $t \supset T_x(n)$: $t \supset x = T_x(0)$, and suppose $t \supset T_x(n)$, then by the lemma $t \supset T_x(n) \cup \cup T_x(n) = T_x(N+1)$. So we have the claim.

Theorem: Suppose $G : V \otimes \mathrm{Ptl}_S(V, V) \to V$ is a Function. Then there is a unique $F : V \to V$ such that $\forall a \in V$, $F(a) = G(a, F\mid_{\{b:b\in a\}}) = G(a, F\mid_a)$: say ϕ is an attempt if $\phi \in \mathrm{Ptl}_S(V, V) \wedge \mathrm{dom}\phi$ is transitive $\wedge\ \forall a \in \mathrm{dom}\phi\phi(a) = G(a, \phi\mid_a)$. As ever, any two attempts ϕ, ψ agree on their common domain $\mathrm{dom}\phi \cap \mathrm{dom}\psi$ (by induction, or else take a $\in$-least in $\mathrm{dom}\phi \cap \mathrm{dom}\psi$ where they disagree, then $\phi(a) = G(a, \phi\mid_a) = G(a, \psi\mid_a) = \psi(a)$, a contradiction. So we can define F functional by $F(a) = b$ (i.e. $(a, b) \in F$) iff there is an attempt ϕ with $\phi(a) = b$. Note $\mathrm{dom}F = \cup\mathrm{dom}\phi$ so is transitive. It remains to prove that $\mathrm{dom}F = V$; suppose not, then take a $\in$-least such that $a \notin \mathrm{dom}F$. Then $\forall b \in a$, $b \in \mathrm{dom}F$, i.e. $a \subset \mathrm{dom}F$, a transitive set, so $TC(a) \subset \mathrm{dom}F$. (By replacement, the image, $\mathrm{Im}F\mid_{TC(a)}$ is a set and so) $F\mid_{TC(a)} := \phi$ say, is a set. Clearly ϕ is an attempt, so we can exted ϕ to $\bar\phi$ defined on $TC(\{a\})$ by setting $\bar\phi(a) = G(a, \phi\mid_a)$. Then $a \in \mathrm{dom}\bar\phi$ for an attempt $\bar\phi$ so $a \in \mathrm{dom}F$, a contradiction.

Asides: 1) If we are dealing with transitive relations we don't need to fuss about transitive closure. 2) We do have $\forall b \in a \exists\phi : \phi$ is an attempt with $b \in \mathrm{dom}\phi$, but we would need to use AC to pick a ϕ for each $b \in a$, which we don't want to do.

Essentially the same argument proves: suppose A is a class, R a Relation on A which is local in the sense that $\forall a \in A$, $\{b \in A : bRa\}$ is a set, and well-founded in the sense that for any $0 \neq X \subset A$ or $0 \neq x \subset A$ (i.e. for both classes and sets) there is an R-minimal element, i.e. $\exists a \in X$ or x such that $\forall b \in AbRa \Rightarrow b \in X$.

Theorem: Given $G : A \times \mathrm{Ptl}_S(A, V) \to V$ a Function, there is a unique $F : A \to V$ with $\forall a \in A, F(a) = G(a, F\mid_{\{b\in A:bRa\}})$.

Appendix: We can code mathematics in ZF set theory. To do the work on cardinals or on models of the predicate calculus we should add an Axiom of Choice: $\forall xy \forall y : x \to y \forall w \in y \exists z \in x f(x) = w \to \exists s : y \to x : \forall w \in y f(s(w)) = w$. ZF with AC is usually written ZFC.

7 Models and Consistency

7.1 The Mostowski Collapse

Let A be a class and E a local well-founded relation on A. Then there is a Function $F : A \to V$ such that $F(a) = \{F(b) : bEa\}$. Then $\mathrm{Im}F$ is a transitive class. Suppose further that (A, E) satisfies extensionality, in the sense that $\forall a, b \in A(\forall c \in AcEa \leftrightarrow cEb) \to a = b$. Then F is injective, and $(\mathrm{Im}F, E) \simeq (A, E)$. NOw imagin that we are considering models of (the language) ZF in mathematics coded in ZF(C). Then if (x, E) is such a model and E is "really" well-founded (which is not quite the same thing as being well-founded in the sense above, but we lack the time to explore the difference) and the model satisfies extensionality, then $(x, E) \simeq (z, \in)$. So we can reasonably focus attention on models of the language which are themselves transitive sets.

7.2 Models for fragments of ZF

Example: Consider $V\omega$ (with the standard $\in$). What axioms are true? 1) Extensionality and Foundation are automatic 2) $0, \{\}, \cup$ are absolute; we just need $V\omega$ to be closed under them, which it is. 3) Power set P is not absolute; we have to check what the power set means from the model's point of view. For $x \in V\omega$, Px in $V\omega$ is $\{y \subset x : y \in V\omega\}$; in this case this is (the usual) Px and is in $V\omega$, 4) Replacement has a similar subtlety, but is OK, 5) There are clearly no infinite members of $V\omega$; infinity fails.

Observation: Suppose infinity held. Then $(V\omega, \in)$ would be a model of ZF, so ZF is consistent in ZF. This contradicts Gödel's second incompleteness theorem, assuming ZF is consistent.

[This ends the course; an excellent supplementary lecture covering the incompleteness theorems was also given, but will not be included in these notes].