Logic and Set Theory

Mathematical Tripos Part II

June 5, 2023

1 Propositional Logic

1.1 Languages

We begin with *propositional logic*.

Definition 1.1 (Language)**.** Let P be a set of *primitive propositions*. Unless otherwise stated, $P = \{p_1, p_2, \dots\}$. The *language* L or $L(P)$ is defined inductively by

1. If $p \in P$, then $p \in L$.
2. $\perp \in L$ (where we read $\perp$ as 'false')
3. If $p, q \in L$ then $(p \Rightarrow q) \in L$.

Every *proposition* (member of L) is a finite string of symbols from the alphabet $\{(,), \Rightarrow, \perp, p_1, p_2, \dots\}$, satisfying some grammar.

The precise inductive definition is as follows. Let $L_1 = P \cup \{\perp\}$, and define $L_{n+1} = L_n \cup \{(p \Rightarrow q) \mid p, q \in L_n\}$. Then $L = \cup_{n=1}^{\infty} L_n$. Note there is exactly one way in which any element of the language can be constructed from the rules above.

We introduce the abbreviations $\neg$ (not), $\wedge$ (and), and $\vee$ (or) defined by

$$\neg p = (p \Rightarrow \perp); \quad p \wedge q = \neg(p \Rightarrow \neg q); \quad p \vee q = \neg p \Rightarrow q$$

1.2 Semantic Implication

We now assign some sort of 'true' or 'false' values to propositions.

Definition 1.2 (Valuation)**.** A *valuation* is a function $v : L \to \{0, 1\}$ (where we think of 0 as false and 1 as true) such that

1. $v(\perp) = 0$
2. $v(p \Rightarrow q) = 0$ if $v(p) = 1$ and $v(q) = 0$, and 1 otherwise.

Proposition 1.3 (Valuations Defined on Primitives)**.**

(i) Let $v, v' : L \to \{0, 1\}$ be valuations that agree on the primitives p_i. Then $v = v'$.

(ii) Any function $w : P \to \{0, 1\}$ extends to a valuation.

Proof. (i) Clearly v, v' agree on L_1. Then if they agree on p, q, they agree on $p \Rightarrow q$. So by induction they agree on L_n for all n, and hence on L.

(ii) Let $v(p) = w(p)$ for all $p \in P$, and $v(\perp) = 0$ to obtain v on the set L_1. Assuming v is defined on p, q we can define it at $p \Rightarrow q$ in the obvious way. This defines v on all of L. □

Definition 1.4 (Tautology)**.** A *tautology* is an element $t \in L$ such that $v(t) = 1$ for any valuation v. We write $\models t$.

Some examples of tautologies are $p \Rightarrow (q \Rightarrow p)$ and $\neg\neg p \Rightarrow p$.

Definition 1.5 (Semantic Implication)**.** Let $S \subseteq L$ and $t \in L$. We say S *entails* or *semantically implies* t, written $S \models t$, if $v(t) = 1$ whenever $v(s) = 1$ for all $s \in S$.

Definition 1.6 (Model)**.** We say that v is a *model* of S in L if $v(s) = 1$ for all $s \in S$.

So the statement $S \models t$ is the statement that every model of S is also a model of t. We also note that the notation $\models t$ is equivalent to $\emptyset \models t$.

1.3 Syntactic Implication

For a notion of proof, we require a system of axioms and deduction rules.

Axiom 1.7 (Axiom Scheme)**.**

1. $p \Rightarrow (q \Rightarrow p)$ for all $p, q \in L$;
2. $[p \Rightarrow (q \Rightarrow r)] \Rightarrow [(p \Rightarrow q) \Rightarrow (p \Rightarrow r)]$ for all $p, q \in L$;
3. $(\neg\neg p) \Rightarrow p$ for all $p \in L$.

For our deduction rules, we will only have *modus ponens*: from each p and $p \Rightarrow q$ we can deduce q.

Definition 1.8 (Syntactic Implication/Proof)**.** For $S \subseteq L$, and $t \in S$ we say S *proves* or *syntactically implies* t, written $S \vdash t$, if there exists a sequence $t_1, \dots, t_n = t$ in L with every t_i is either an axiom, a member of S or q where $t_j = p$ and $t_k = p \Rightarrow q$ and $j, k < i$.

We say that S is the set of *premises* or *hypotheses*, and t is the *conclusion*.

Example 1.9. We will prove $\{p \Rightarrow q, q \Rightarrow r\} \vdash p \Rightarrow r$.

1. $q \Rightarrow r$ (Hyp)
2. $(q \Rightarrow r) \Rightarrow (p \Rightarrow (q \Rightarrow r))$ (Ax 1)
3. $p \Rightarrow (q \Rightarrow r)$ (MP on 2, 3)
4. $[p \Rightarrow (q \Rightarrow r)] \Rightarrow [(p \Rightarrow q) \Rightarrow (p \Rightarrow r)]$ (Ax 2)
5. $(p \Rightarrow q) \Rightarrow (p \Rightarrow r)$ (MP on 3, 4)
6. $p \Rightarrow q$ (Hyp)
7. $p \Rightarrow r$ (MP on 5, 6)

Definition 1.10 (Theorem)**.** If $\emptyset \vdash t$, we say that t is a *theorem*, written $\vdash t$.

Example 1.11. We will prove the theorem $\vdash (p \Rightarrow p)$.

1. $(p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p))$ (Ax 2)
2. $p \Rightarrow ((p \Rightarrow p) \Rightarrow p)$ (Ax 1)
3. $(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)$ (MP on 1, 2)
4. $p \Rightarrow (p \Rightarrow p)$ (Ax 1)
5. $p \Rightarrow p$ (MP on 3, 4)

1.4 The Deduction Theorem

Often, showing $S \vdash p$ is made easier by the idea that provability corresponds to the connective '$\Rightarrow$' in L.

Theorem 1.12 (Deduction Theorem)**.** *Let $S \subseteq L$ and $p, q \in L$. Then $S \vdash (p \Rightarrow q)$ if and only if $S \cup \{p\} \vdash q$.*

Proof. Given a proof of $p \Rightarrow q$ from S, add the line p by hypothesis and deduce q from modus ponens, to obtain a proof of q from $S \cup \{p\}$.

Conversely, suppose we have a proof of q from $S \cup \{p\}$. Let $t_1, \dots, t_n$ be the lines of the proof. We will prove that $S \vdash (p \Rightarrow t_i)$ for all i.

- If t_i is an axiom, we write t_i (Ax); $t_i \Rightarrow (p \Rightarrow t_i)$ (Ax 1); $p \Rightarrow t_i$ (MP).
- If $t_i \in S$, we write t_i (Hyp); $t_i \Rightarrow (p \Rightarrow t_i)$ (Ax 1); $p \Rightarrow t_i$ (MP).
- If $t_i = p$, we write the proof of $\vdash p \Rightarrow p$ given above.
- If t_i is obtained by MP from t_j and $t_k = t_j \Rightarrow t_i$, assume by induction that $S \vdash p \Rightarrow t_k$ and $S \vdash p \Rightarrow (t_j \Rightarrow t_i)$. Then write $(p \Rightarrow (t_j \Rightarrow t_i)) \Rightarrow ((p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i))$ (Ax 2); $(p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)$ (MP); $p \Rightarrow t_i$ (MP). □

Example 1.13 (Using the Deduction Theorem)**.** To show $\{p \Rightarrow q, q \Rightarrow r\} \vdash p \Rightarrow r$, the deduction theorem says that it's sufficient to show that $\{p, p \Rightarrow q, q \Rightarrow r\} \vdash r$, which is just modus ponens twice.

1.5 The Completeness Theorem

We want to prove that $S \models t \iff S \vdash t$, the *completeness theorem*. This is made up of the *soundness* statement, $S \vdash t \Rightarrow S \models t$ ('our axioms and deduction rule are not silly'), and the *adequacy statement*, $S \models t \Rightarrow S \vdash t$ ('our axioms are strong enough to deduce from S every semantic consequence of S').

Proposition 1.14 (Soundness)**.** *Let $S \subseteq L$, $t \in L$. Then $S \vdash t$ implies that $S \models t$.*

Proof. We have a proof of t from S. We have $v(p) = 1$ for all $p \in S$ (as v is a model of S) and $v(p) = 1$ for every axiom p (as each axiom is a tautology), and if $v(p) = 1$ and $v(p \Rightarrow q) = 1$ then $v(q) = 1$. Hence each line t_i of our proof of t from S has $v(t_i) = 1$. □

Definition 1.15 (Consistent)**.** We say that S is *consistent* if $S \nvdash \perp$.

A special case of adequacy is $S \models \perp \Rightarrow S \vdash \perp$ (or taking the contrapositive, S is consistent implies S has a model).

This special case implies adequacy in general. Given $S \models t$, we have that $S \cup \{\neg t\}$ has no model, so we should know $S \cup \{\neg t\} \vdash \perp$ hence $S \vdash ((\neg t) \Rightarrow \perp)$ (by the deduction theorem), so $S \vdash (\neg\neg t)$, but axiom 3 gives $S \vdash ((\neg\neg t) \Rightarrow t)$ thus $S \vdash t$ by modus ponens.

Definition 1.16 (Deductive Closure)**.** We say a set $S \subseteq L$ is *deductively closed* if $p \in S$ whenever $S \vdash p$. Any set S has a *deductive closure*, which is the (deductively closed) set of statements $\{t \in L \mid S \vdash t\}$ that S proves.

Theorem 1.17 (Model Existence Lemma)**.** *Every consistent $S \subseteq L$ has a model.*

Proof. First we note that for any consistent $S \subseteq L$ and $p \in L$, $S \cup \{p\}$ or $S \cup \{\neg p\}$ is consistent[1]. Since L is countable, we can list it as $t_1, t_2, \dots$. Let $S_0 = S$, $S_1 = S_0 \cup \{t_1\}$ or $S_0 \cup \{\neg t_1\}$ (so that S_1 is consistent), and continue on. Set $\overline{S} = S_0 \cup S_1 \dots$. Then for all $t \in L$, either $t \in \overline{S}$ or $(\not t) \in \overline{S}$.

We can easily see inductively (since proofs are finite) that $\overline{S}$ is consistent, and also that $\overline{S}$ is deductively closed (as otherwise we'd have introduced an inconsistency). So we now define $v : L \to \{0, 1\}$ by

$$v(t) = \begin{cases} 1 & t \in \overline{S} \\ 0 & \text{otherwise} \end{cases}.$$

We then need only to check v is a valuation. To see $v(\perp) = 0$, we note $\overline{S}$ consistent implies that $\perp \notin \overline{S}$.

Now suppose $v(p) = 1, v(q) = 0$. Then $p \in \bar{S}$ and $q \notin \bar{S}$, and we want to show $(p \Rightarrow q) \notin \bar{S}$. If this were not the case, we would have $(p \Rightarrow q) \in \bar{S}$ and $p \in \bar{S}$, so $q \in \bar{S}$ as $\bar{S}$ is deductively closed.

Now suppose $v(q) = 1$, so $q \in \bar{S}$, and we need to show $(p \Rightarrow q) \in \bar{S}$. Then $\bar{S} \vdash q$, and by axiom 1, $\bar{S} \vdash q \Rightarrow (p \Rightarrow q)$. Therefore, since $\bar{S}$ is deductively closed, $(p \Rightarrow q) \in \bar{S}$.

Finally, suppose $v(p) = 0$, so $p \notin \bar{S}$. We want to show $(p \Rightarrow q) \in \bar{S}$. We know that $\neg p \in \bar{S}$, so it suffices to show that $p \Rightarrow \perp \vdash p \Rightarrow q$. By the deduction theorem, this is equivalent to proving $\{p, p \Rightarrow \perp\} \vdash q$, or equivalently, $\perp \vdash q$. But by axiom 1, $\perp \Rightarrow (\neg q \Rightarrow \perp)$ where $(\neg q \Rightarrow \perp) = \neg\neg q$, so the proof is complete by axiom 3. □

Corollary 1.18 (Adequacy)**.** *Let $S \subseteq L$, $t \in L$. Then $S \models t$ implies that $S \vdash t$.*

Putting our soundness and adequacy statements together, we get completeness.

Theorem 1.19 (Completeness Theorem for Propositional Logic)**.** *Let $S \subseteq L$ and $t \in L$. Then $S \models t$ if and only if $S \vdash t$.*

Directly from the completeness theorem we get a series of interesting consequences.

Theorem 1.20 (Compactness Theorem)**.** *Let $S \subseteq L$ and $t \in L$ with $S \models t$. Then there exists a finite subset $S' \subseteq S$ such that $S' \models t$.*

Proof. This follows directly from the completeness theorem, since proofs depend on only finitely many hypotheses in S. □

Corollary 1.21 (Compactness Theorem, Equivalent Form)**.** *Let $S \subseteq L$. Then if every finite subset $S' \subseteq S$ has a model, then S has a model.*

Theorem 1.22 (Decidability)**.** *Let $S \subseteq L$ and $t \in L$. Then there is an algorithm to decide (in finite time) if $S \vdash t$.*

Proof. Trivial by replacing $\vdash$ with $\models$, by drawing the relevant truth tables. □

2 Well-Ordering and Ordinals

2.1 Well-Orderings

We will now talk about orderings on sets.

Definition 2.1 (Total Order)**.** A *total order* is a pair $(X, <)$ where X is a set and $<$ is a relation on X that is

1. *irreflexive*: for all $x \in X$, $x \not< x$;
2. *transitive*: for all $x, y, z \in X$, $x < y$ and $y < z$ implies $x < z$;
3. *trichotomous*: for all $x, y \in X$, either $x < y$, $y < x$ or $x = y$.

We can instead have defined a total order in terms of $\leq$ in the obvious way (with reflexivity, transitivity, antisymmetry and trichotomy).

Definition 2.2 (Well-Ordering)**.** A total order $(X, <)$ is a *well-ordering* if every (non-empty) subset of X has a least element.

Proposition 2.3 (Decreasing Sequence Condition)**.** *$(X, <)$ is a well-ordering if and only if there does not exist a strictly decreasing sequence in X.*

Well-orderings allow us to perform (strong) induction on a set.

Proposition 2.4 (Principle of Induction)**.** *Let X be well-ordered and let $S \subseteq X$ be such that whenever $y \in S$ for all $y < x$ then $x \in S$. Then $S = X$.*

Proof. Suppose $S \neq X$. Then there is a least $x \in X \backslash S$. Then $y \in S$ for all $y < x$ but $x \notin S$ which is a contradiction. □

We will identify total orders as isomorphic in the natural way.

[1] If not, then $S \cup \{p\} \vdash \perp$ and $S \cup \{\neg p\} \vdash \perp$ so $S \vdash (p \Rightarrow \perp)$, that is, $S \vdash (\neg p)$. Hence from $S \cup \{\neg p\} \vdash \perp$ we obtain $S \vdash \perp$.

Definition 2.5 (Order Isomorphism)**.** We say that two total orders X and Y are *order isomorphic* if there exists a bijection $f : X \to Y$ such that $x < y \iff f(x) < f(y)$.

Proposition 2.6. *Let X, Y be isomorphic well-orderings. Then there exists a unique isomorphism.*

Proof. Let $f, g : X \to Y$ be two isomorphisms. Suppose that for some $x \in X$, we have $f(y) = g(y)$ for all $y < x$. Define $S = \{f(y) \mid y < x\}$, and note $S = \{g(y) \mid y < x\}$ by assumption. Then $Y \backslash S$ is non-empty as $f(x) \notin S$, so it has a least element say a. We must then have $f(x) = a$, and by the same logic $g(x) = f(x) = a$. So by induction $f(x) = g(x)$ for all $x \in X$. $\square$

2.2 Initial Segments

Given an ordered set, we can remove the end of the set and keep the beginning. What we are left with is an initial segment.

Definition 2.7 (Initial Segment)**.** A subset I of a totally ordered set X is a *initial segment* if $x \in I$ implies $y \in I$ for all $y < x$.

Proposition 2.8 (Initial Segments in Well-Orderings)**.** *Every initial segment of a well-ordered set X is of the form $I_x = \{y \in X \mid y < x\}$ for some x.*

Proof. We can see that I_x is clearly an initial segment for any $x \in X$, and if Y is any initial segment, we can take $x = \min X \backslash Y$ to get $Y = I_x$. $\square$

We now want to show that every subset of a well-ordering X is isomorphic to an initial segment of X. To do this we need to build a notion of recursion.

Theorem 2.9 (Definition by Recursion)**.** *Let X be a well-ordering and let Y be any set. Take $G : \mathcal{P}(X \times Y) \to Y$ (i.e a 'rule'). Then there exists a unique function $f : X \to Y$ such that $f(x) = G\left(f|_{I_x}\right)$ for all $x \in X$.*

Proof. Say that h is an 'attempt' if $h : I \to Y$ for some initial segment I of X, and for all $x \in I$ we have $h(x) = G\left(h|_{I_x}\right)$.

If we have two attempts h, h' both defined at x, then they must agree, by induction on x.

For all x, there also exists an attempt defined at x, by induction on x. Indeed, by induction we can assume there exists an attempt h_y defined at y for all $y < x$, and then we can define h to be the union of the h_y. This is an attempt with domain I_x, so the attempt $h' = h \cup \{x, G(h)\}$ is an attempt defined at x. Therefore, there is an attempt defined at each x, so we can define the function $f : X \to Y$ by setting $f(x)$ to be the value of $h(x)$ where h is some attempt defined at x.

For uniqueness, we apply induction on x. If f, f' agree below x, then they must agree at x since $f(x) = G\left(f|_{I_x}\right) = G\left(f'|_{I_x}\right) = f'(x)$. $\square$

We can now prove our result by recursively sending minimum elements of our subset to minimum elements of our set.

Proposition 2.10 (Subset Collapse)**.** *Any subset Y of a well-ordering X is isomorphic to a unique initial segment of X.*

Proof. For $f : X \to Y$ to be an order preserving bijection with an initial segment of x, we need to map x to the smallest thing not yet mapped to, that is, $f(x) = \min(X \backslash \{f(y) \mid y < x\})$. We can take this minimum since $f(z) < x$ for all $z < x$, and hence x is in this set. Then, by the recursion theorem, this function exists and is unique. $\square$

Corollary 2.11. *A well-ordered X can never be isomorphic to a proper initial segment of itself.*

Proof. Since X is isomorphic to itself by the identity function, and uniqueness shows that it cannot be isomorphic to another initial segment. $\square$

2.3 Relating Well-Orderings

We now want to be able to talk more about the structure of well-orderings.

Definition 2.12 (Comparing Well-Orderings)**.** For well-orderings X, Y, we will write $X \leq Y$ if X is isomorphic to an initial segment of Y.

By subset collapse, $X \leq Y$ if and only if X is isomorphic to some subset of Y.

Proposition 2.13 (Trichotomy)**.** *Let X, Y be well-orderings. Then either $X \leq Y$, $Y \leq X$ or both (in which case X and Y are isomorphic).*

Proof. Consider $f : X \to Y$ given by $f(x) = \min(Y \backslash \{f(y) \mid y < x\})$. If this is well-defined, then it is an isomorphism from X to an initial segment of Y. If it is not well-defined, there is some x such that $\{f(y) \mid y < x\} = Y$, but then f is a isomorphism between I_x and Y. Hence either $X \leq Y$ or $Y \leq X$ or both.

If both hold, let $g : Y \to X$ be defined similarly and consider $g \circ f : X \to X$. This is an isomorphism from X to an initial segment of X, and hence all of X (as the initial segment can't be proper). For this to occur we must have that f and g are isomorphisms between X and Y. $\square$

2.4 Constructing Larger Well-Orderings

Given a well-ordering, we can extend it by exactly one element.

Definition 2.14 (Successor)**.** Given a well-ordering X, choose some $x \notin X$ and define a well-ordering on $X \cup \{x\}$ by setting $y < x$ for all $y \in X$. This is *successor* of X, written X^+.

We clearly have $X < X^+$. We can also 'stick a bunch of well-orderings together'.

Definition 2.15 (Extensions)**.** Given well-orderings $(X, <_X)$ and $(Y, <_Y)$ we say Y *extends* X if X is a proper initial segment of Y and $<_X, <_Y$ agree when both defined.

We say well-orderings $\{X_i \mid i \in I\}$ are *nested* if for all i, j one of X_i and X_j extends the other.

Proposition 2.16 (Extending Well-Orderings)**.** *Let $\{X_i \mid i \in I\}$ be a nested set of well-orderings. Then there exists a well-ordering X such that $X_i \leq X$ for all i.*

Proof. Let $X = \cup_i X_i$ with the ordering given by $<_X = \cup_i <_i$. That is, $x < y$ in X if there exists i such that $x, y \in X_i$ and $x <_i y$.

Given $S \subseteq X$ non-empty, we have $S \cap X_i$ non-empty for some $i \in I$. Let x be the least element in this (under $<_i$). Then x is the least element of S in X since X_i is an initial segment of X, by nestedness. So X is a well-ordering, and $X \geq X_i$ for all i. $\square$

2.5 Ordinals

We will now introduce a more convenient way of talking about well-orderings.

Definition 2.17 (Ordinal)**.** An *ordinal* is a well-ordered set, where we regard two ordinals as equal if they are isomorphic.

Definition 2.18 (Order Type)**.** If a well-ordering X has corresponding ordinal α, we say X has *order type* α, and write $\operatorname{otp}(X) = \alpha$.

The order type of the unique well-ordering on a collection of $k \in \mathbb{N}$ points is named k. The order type of $(\mathbb{N}, <)$ is named ω.

For ordinals α, β, write $\alpha \leq \beta$ if $X \leq Y$ for some X of order type α and Y of order type β. This does not depend on the choice of X and Y (since any two choices must be isomorphic).

Proposition 2.19 (I_α is Well-Ordered)**.** *For any ordinal α, $I_\alpha = \{\beta \mid \beta < \alpha\}$ form a well-ordered set of order type α.*

Proof. Let $\operatorname{otp}(X) = \alpha$. Then well-orderings $< X$ are precisely (up to isomorphism) the proper initial segments of X (by uniqueness of subset collapse). But these are the I_x for all $x \in X$, so we can biject X with the well-orderings $< X$ by $x \mapsto I_x$. $\square$

Proposition 2.20 (The Ordinals are Well-Ordered)**.** *Every non-empty set S of ordinals has a least element.*

Proof. Choose $\alpha \in S$. If it is minimal, done. If not, then $S \cap I_\alpha$ is non-empty, but I_α is well-ordered and hence has a least element, β. Then this is a minimal element of S. $\square$

However, the ordinals do not form a well-ordered *set* because of the following.

Theorem 2.21 (Burali-Forti Paradox)**.** *The ordinals do not form a set.*

Proof. Suppose X was the set of all ordinals. Then since it's well-ordered it has an order type say α. Thus X is order-isomorphic to I_α, so X is order-isomorphic to a proper initial subset of itself, which is a contradiction. $\square$

Definition 2.22 (Supremum)**.** Let $S = \{\alpha_i \mid i \in I\}$ be a set of ordinals. We define $\sup S$ to be the *supremum* or *least upper bound* of S.

We note that the least upper bound exists since an upper bound exists (taking α arising from extending the nested family of well orderings I_{α_i}), and I_α is a set.

2.6 Examples of Ordinals

We already know the ordinals $\{0, 1, 2, 3, \dots, \omega\}$. Writing $\alpha + 1$ for the successor α^+ of α, this along with the supremum allows us to generate a lot of ordinals.

$$\begin{array}{cccccc}
0 & \omega\cdot 2+1 & \omega^2+1 & \omega^2\cdot 3 & \omega^{\omega+2} & \varepsilon_0+1 \\
1 & \omega\cdot 2+2 & \omega^2+2 & \omega^2\cdot 4 & \vdots & \vdots \\
2 & \omega\cdot 2+3 & \omega^2+3 & \omega^2\cdot 5 & \omega^{\omega\cdot 2} & \varepsilon_0\cdot 2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\omega & \omega\cdot 3 & \omega^2+\omega & \omega^3 & \omega^{\omega^2} & \varepsilon_0^2 \\
\omega+1 & \omega\cdot 4 & \vdots & \vdots & \vdots & \vdots \\
\omega+2 & \omega\cdot 5 & \omega^2+\omega\cdot 2 & \omega^\omega & \omega^{\omega^\omega} & \varepsilon_0^{\varepsilon_0} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\omega+\omega & \omega\cdot\omega & \omega^2\cdot 2 & \omega^{\omega+1} & \omega^{\omega^{\cdot^{\cdot^{\cdot}}}} & \varepsilon_0^{\varepsilon_0^{\cdot^{\cdot^{\cdot}}}} \\
=\omega\cdot 2 & =\omega^2 & & & =\varepsilon_0 & =\varepsilon_1
\end{array}$$

We have used some new notation, for example $\omega + 1 = \omega^+$ and $\omega \cdot 2 = \sup\{\omega, \omega + 1, \cdots$. We will formally define this later.

Each of the ordinals above are countable, as at each step we are only adding one element and taking countable unions. This question is equivalent to asking can we well order $\mathbb{R}$, and we find that we can though we can't write it down.

Theorem 2.23 (Uncountable Ordinal Existence)**.** *There is an uncountable ordinal.*

Proof. Let $A = \{R \in \mathcal{P}(\omega \times \omega) \mid R$ is a well ordering of a subset of $\omega\}$. Then $B = \{\operatorname{otp}(R) \mid R \in A\}$ is the set of all countable ordinals.

Let $\omega_1 = \sup B$. If ω_1 was countable, it would be in B. But then $\omega_1 < \omega_1^+ \in B$ is a contradiction, so ω_1 is uncountable.

Alternatively we could say that B isn't all ordinals since they don't form a set (Burali-Forti), forcing there to be an uncountable ordinal. $\square$

We note ω_1 is the *least* uncountable ordinal by the definition of B.

The ordering ω_1 has some remarkable properties. For example all of the proper initial segments of ω_1 are countable but ω_1 is not. Also any sequence $\alpha_1, \alpha_2, \dots$ in I_{ω_1} is bounded, namely by $\sup\{\alpha_1, \alpha_2, \dots\}$ which is countable as a countable union of countable sets.

The same argument allows us to find ordinals that don't inject into any given set.

Theorem 2.24 (Hartogs' Lemma)**.** *For every set X, there exists an ordinal α that does not inject into X. We call the least such ordinal $\gamma(X)$ (read 'Hartogs' of X')*

So $\gamma(\omega) = \omega_1$.

2.7 Limits and Successors

In general we can divide ordinals into two categories.

Definition 2.25 (Limits and Successors)**.** We say α is a *successor* if there exists β such that $\alpha = \beta^+$. Otherwise we say that α is a *limit*.

An α has a greatest element if and only if it is a successor. So α is a limit if and only if it has no greatest element. We typically denote limit ordinals by λ.

2.8 Ordinal Arithmetic

We will now make formal sense out of our arithmetic notation such as $\omega + \omega$ used earlier.

Definition 2.26 (Ordinal Addition – Inductive)**.** We define $\alpha + \beta$ by recursion on β (keeping α fixed). We take

$$\alpha + 0 = \alpha,$$
$$\alpha + \beta^+ = (\alpha + \beta)^+,$$
$$\alpha + \lambda = \sup\{\alpha + \gamma \mid \gamma < \lambda\},$$

for non-zero limit λ.

Proposition 2.27 (Non-Commutative Addition)**.** *Ordinal addition is not commutative*[2]*. In particular, $\omega + 1 \neq 1 + \omega$.*

Proof. We have $\omega + 1 = \omega + 0^+ = \omega^+$, and $1 + \omega = \sup\{1 + n \mid n \leq \omega\} = \sup\{1, 2, 3, \dots\} = \omega$. $\square$

Proposition 2.28 (Associative Addition)**.** *For all ordinals α, β, γ, we have $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.*

[2] Arises from asymmetry in decision to recurse on the right

Proof. We induct on γ, keeping α, β fixed. If $\gamma = 0$, then $\alpha + (\beta + 0) = \alpha + \beta = (\alpha + \beta) + 0$.

If $\gamma = \delta^+$ is a successor, then

$$\begin{aligned} \alpha + (\beta + \delta^+) &= \alpha + (\beta + \delta)^+ \\ &= [\alpha + (\beta + \delta)]^+ = [(\alpha + \beta) + \delta]^+ \\ &= (\alpha + \beta) + \delta^+, \end{aligned}$$

as required.

If λ is a non-zero limit ordinal, we have

$$\begin{aligned} (\alpha + \beta) + \lambda &= \sup\{(\alpha + \beta) + \gamma : \gamma < \lambda\} \\ &= \sup\{\alpha + (\beta + \gamma) : \gamma < \lambda\}. \end{aligned}$$

We claim that $\beta + \lambda$ is a limit. Indeed, we have $\beta + \lambda = \sup\{\beta + \gamma : \gamma < \lambda\}$. But λ is a limit, so for every $\gamma < \lambda$, we can find a γ' with $\gamma < \gamma' < \lambda$. So $\beta + \gamma' > \beta + \gamma$, so $\beta + \gamma$ cannot be the greatest element.

Now $\alpha + (\beta + \lambda) = \sup\{\alpha + \delta \mid \delta < \beta + \lambda\}$. We need to show that

$$\sup\{\alpha + \delta \mid \delta < \beta + \lambda\} = \sup\{\alpha + (\beta + \gamma) \mid \gamma < \lambda\}.$$

Now each element on the right is an element of the left, so we get that the left is $\geq$ the right. Also, for $\delta < \beta + \lambda$ we have $\delta < \sup\{\beta + \gamma \mid \gamma < \lambda\}$, so $\delta < \beta + \gamma$ for some $\gamma < \lambda$. Hence $\alpha + \delta < \alpha + (\beta + \gamma)$. Thus the left is $\leq$ the right too, and thus these supremums must be equal. □

We can give an alternative definition of addition based on actual well-orders, intuitively by writing all of the elements of α followed by all the elements of β.

$$\alpha + \beta = \underbrace{\quad \alpha \quad}\underbrace{\quad \beta \quad}$$

Definition 2.29 (Ordinal Addition – Synthetic)**.** $\alpha + \beta$ is the order type of $\alpha \sqcup \beta$ with all α before all of β.

Proposition 2.30 (Addition Notion Equivalence)**.** *The inductive and synthetic definitions of addition coincide.*

Proof. We write $+$ for the inductively defined one, and $+'$ for the synthetic one. We'll show $\alpha + \beta = \alpha +' \beta$ for all $\alpha + \beta$ by induction on β (with α fixed). We check the cases

1. *zero*: $\alpha + 0 = \alpha = \alpha +' 0$;
2. *successor*: $\alpha + \beta^+ = (\alpha + \beta)^+ = (\alpha +' \beta)^+$, which is the order type of

   $$\underbrace{\quad \alpha \quad}\underbrace{\quad \beta \quad}\underbrace{\ \bullet\ }$$

   which is $\alpha +' \beta^+$.
3. *non-zero limit*: $\alpha + \lambda = \sup\{\alpha + \gamma \mid \gamma < \lambda\} = \sup\{\alpha +' \gamma \mid \gamma < \lambda\} = \alpha +' \lambda$ (since the supremum is a union as sets are nested). □

We can then define multiplication, both inductively and synthetically.

Definition 2.31 (Ordinal Multiplication – Inductive)**.** We define $\alpha \cdot \beta$ by recursion on β (keeping α fixed). We take

$$\begin{aligned} \alpha \cdot 0 &= 0, \\ \alpha \cdot (\beta^+) &= \alpha \cdot \beta + \alpha, \\ \alpha \cdot \lambda &= \sup\{\alpha \cdot \gamma \mid \gamma < \lambda\}, \end{aligned}$$

for non-zero limit λ.

Diagrammatically, our synthetic definition is given by

$$\alpha \cdot \beta = \underbrace{\underbrace{\alpha}\underbrace{\alpha}\cdots\underbrace{\alpha}}_{\beta \text{ times}}.$$

Definition 2.32 (Ordinal Multiplication – Synthetic)**.** $\alpha \cdot \beta$ is the order type of $\alpha \times \beta$, with $(x, y) < (x', y')$ if $y < y'$ or ($y = y'$ and $x < x'$).

We can again check that the inductive and synthetic definition agree, that ordinal multiplication is again not commutative ($\omega \cdot 2 = \omega + \omega$ but $2 \cdot \omega = \omega$) but that it is associative and so on in the exact same way as we did for addition.

We can also define exponentiation. We will give only the inductive definition.

Definition 2.33 (Ordinal Exponentiation)**.** We define α^β by induction on β. We take

$$\begin{aligned} \alpha^0 &= 1, \\ \alpha^{\beta^+} &= \alpha^\beta \cdot \alpha, \\ \alpha^\lambda &= \sup\{\alpha^\gamma \mid \gamma < \lambda\}, \end{aligned}$$

where λ is a non-zero limit ordinal.

3 Posets and Zorn's Lemma

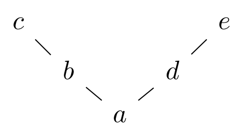
3.1 Posets and Hasse Diagrams

We now look at posets and eventually arrive at a statement and proof of Zorn's lemma.

Definition 3.1 (Poset)**.** A *partially ordered set* or *poset* is a pair $(X, \leq)$, where X is a set and $\leq$ is a relation on X that is reflexive, transitive and antisymmetric (if $x \leq y$ and $y \leq x$ then $x = y$). We write $x < y$ if $x \leq y$ and $x \neq y$. In terms of $<$, a poset is irreflexive and transitive.

Example 3.2 (Examples of Posets)**.**

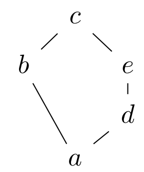
1. Any total order.
2. $(\mathbb{N}^+, \text{'divides'})$
3. For S any set, take $\mathcal{P}(S)$ with $A \leq B$ if $A \subset B$.
4. Take X to be any subset of $\mathcal{P}(S)$ with the same $\leq$.
5. Consider

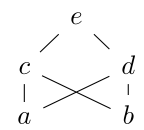
   

   meaning $a \leq b$, $b \leq c$, $a \leq d$, $d \leq e$ and everything following from transitivity. So $a \leq c$ but b and d are unrelated.

Definition 3.3 (Hasse Diagram)**.** A *Hasse diagram* of a poset consists of a drawing of the points in X with an upward line from x to y if y *covers* x (meaning $y > x$ and no z has $y > z > x$).

Example 3.4 (Examples of Hasse Diagrams)**.**

1. 

   (so we have no notion of 'height' or 'rank')
2. 

Definition 3.5 (Chain)**.** A subset S of a poset X is a *chain* if it is a total order. We say S is an *antichain* if no two members of S are related by order.

Definition 3.6 (Supremum)**.** For $S \subset X$, we say $x \in X$ is an *upper bound* for S if $y \leq x$ for all $y \in S$. We say x is a *least upper bound* or *supremum* for S if x is an upper bound for S, and every upper bound y for S has $y \geq x$. We write $\sup S = x$.

Definition 3.7 (Complete)**.** A poset X is *complete* if every set $S \subset X$ has a supremum.

In any complete poset X, there is a greatest element ($\sup X$), and a least element ($\sup \emptyset$).

Definition 3.8 (Order Preserving Map)**.** For a poset X, a function $f : X \to X$ is *order-preserving* if $x \leq y$ implies $f(x) \leq f(y)$.

Theorem 3.9 (Knaster-Tarski Fixed Point Theorem)**.** *Let X be a complete poset, and $f : X \to X$ order-preserving. Then f has a fixed point.*

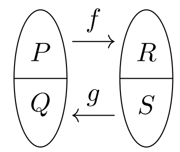
Proof. Let $E = \{x \in X \mid x \leq f(x)\}$, and let $s = \sup E$. We claim $f(s) = s$.

To show $s \leq f(s)$, enough to show that $f(s)$ an upper bound for E (then $s \leq f(s)$ as s is the least upper bound). But $x \in E \Rightarrow x \leq s \Rightarrow f(x) \leq f(s) \Rightarrow x \leq f(x) \leq f(s)$.

To show $f(s) \leq s$, enough to show that $f(s) \in E$ (as s is an upper bound for E). But $s \leq f(s)$, so $f(s) \leq f(f(s))$ (as f order-preserving), i.e. $f(s) \in E$. □

Corollary 3.10 (Schröder-Bernstein Theorem)**.** *Let A, B be sets, and let $f : A \to B$ and $g : B \to A$ be injections. Then there exists a bijection from A to B.*

Proof. We try to partition A into P and Q, and B into R and S, such that $f(P) = R$ and $g(S) = Q$. Then we let $h = f$ on R and g^{-1} on Q.



Since $S = B \setminus R$ and $Q = A \setminus P$, so we want

$$P = A \setminus g(B \setminus f(P))$$

Since the function $P \mapsto A \setminus g(B \setminus f(P))$ from $\mathcal{P}(A)$ to $\mathcal{P}(A)$ is order-preserving (and $\mathcal{P}(a)$ is complete), the result follows. □

3.2 Zorn's Lemma

Now we can get to the heart of the matter.

Definition 3.11 (Maximal Element)**.** In a poset X, an element $x \in X$ is *maximal*[3] if no $y \in X$ has $y > x$.

Theorem 3.12 (Zorn's Lemma)**.** *Assuming the Axiom of Choice, let X be a (non-empty) poset in which every chain has an upper bound. Then X has a maximal element.*

Proof. Suppose X has no maximal element. So for each $x \in X$ there is $x' \in X$ with $x' > x$. We know that every chain C has some upper bound $u(C)$.

Let $\gamma = \gamma(X)$ (as guaranteed by Hartog's Lemma). Pick some $x \in X$, and define x_α, $\alpha < \gamma$ recursively by

$$\begin{aligned} x_0 &= x \\ x_{\alpha^+} &= x'_\alpha \\ x_\lambda &= u(\{x_\alpha \mid \alpha < \lambda\}), \end{aligned}$$

for λ a non-zero limit. Note that $\{x_\alpha \mid \alpha < \lambda\}$ is a chain, by induction. Then the x_α, $\alpha < \gamma$, are distinct, so we have injected γ into X, which is a contradiction. □

3.3 Applications of Zorn

We can use Zorn's lemma to prove some powerful results.

Theorem 3.13. *Every vector space V has a basis.*

Proof. Let X be the set of all linearly independent subsets of V, ordered by inclusion. We want to find a maximal element $A \in X$. If one exists, we are done as if it didn't span we could add an element not in the span and remain linearly independent, which would contradict maximality.

We use Zorn's lemma. Given a chain $\{A_i : i \in I\}$, let $A = \cup_{i \in I} A_i$. Then $A \supset A_i$ for all i, so just need to check $A \in X$, that is, A is linearly independent.

Suppose we have a linear dependence in A, say $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$, where $x_1, \ldots, x_n \in A$ and $\lambda_1, \ldots, \lambda_n$ are scalars (not all 0). We must then have $x_i \in A_{k_i}$ for $k_i \in I$, but then some A_{k_i} contains all of these (as the A_i are a chain), contradicting A_{k_i} being linearly independent. □

We can also use Zorn to dispense with our assumption that our set of primitive propositions was countable in our discussion of propositional logic.

Theorem 3.14 (Model Existence Lemma – Uncountable Case)**.** *Let $S \subset L(P)$, for any set of primitive propositions P. Then S consistent implies S has a model.*

Proof. We need a consistent $\bar{S} \subseteq S$ such that $\forall t \in L$, $t \in \bar{S}$ or $\neg t \in \bar{S}$. Then we have a valuation $v(t) = \mathbb{1}[t \in \bar{S}]$ as in our original proof for the countable case.

So we seek a *maximal* consistent $\bar{S} \supseteq S$. If $\bar{S}$ is maximal, then if $t \notin \bar{S}$, then we must have $\bar{S} \cup \{t\}$ inconsistent, i.e. $\bar{S} \cup \{t\} \vdash \bot$. By deduction theorem, this means that $\bar{S} \vdash \neg t$. By maximality, we must have $\neg t \in \bar{S}$. So either t or $\neg t$ is in $\bar{S}$.

Now we show that there is such a maximal $\bar{S}$. Let $X = \{T \subseteq L : T \text{ is consistent }, T \supseteq S\}$. Then $X \neq \emptyset$ since $S \in X$. We show that any non-empty chain has an upper bound. An obvious choice is, again the union.

Let $\{T_i : i \in I\}$ be a non-empty chain. Let $T = \bigcup T_i$. Then $T \supseteq T_i$ for all i. So to show that T is an upper bound, we have to show $T \in X$.

Certainly, $T \supseteq S$, as any T_i contains S (and the chain is non-empty). So we want to show T is consistent. Suppose $T \vdash \bot$. So we have $t_1, \cdots, t_n \in T$ with $\{t_1, \cdots, t_n\} \vdash \bot$, since proofs are finite. Then some T_k contains all t_i since T_i are nested. So T_k is inconsistent. This is a contradiction. Therefore T must be consistent.

Hence by Zorn's lemma, there is a maximal element of X. □

[3] Take care to note the difference between *maximum* and *maximal* elements

3.4 Zorn's Lemma and the Axiom of Choice

We first look at one other consequence of Zorn's lemma.

Theorem 3.15 (Well-Ordering Principle). *Assuming Zorn's Lemma, Every set S can be well-ordered.*

Proof. Let X be the set of pairs (A, R) where $A \subseteq S$ and R is a well ordering of A, and order X by $(A, R) \leq (A', R')$ if the latter extends the former. X is non empty, as say $(\emptyset, \emptyset) \in X$. Now given a chain $\{(A_i, R_i) \mid i \in I\}$, we have an upper bound $\{\cup_{i\in I} A_i, \cup_{i\in I} R_i)$, since our family is nested. So by Zorn's lemma, there exists a maximal element (A, R). We must have $A = S$, as if not we can take $x \in S\backslash A$ and 'take the successor': well order $A \cup \{x\}$ by making $x > y$ for all $y \in A$ which would then contradict the maximality of (A, R). □

Now, we assumed the Axiom of Choice in our proof of Zorn's lemma, and we assumed Zorn's lemma in our proof of the Well-Ordering Principle. We will now see the Well-Ordering Principle implies the Axiom of Choice, and that all three are equivalent.

Theorem 3.16 (Axiom of Choice). *Assuming the Well-Ordering Principle, If $\{A_i \mid i \in I\}$ is a family of non-empty sets, there is a* choice function *$f : I \to \cup_{i\in I} A_i$ such that $f(i) \in A_i$.*

Proof. Given our family, well-order $\cup_{i\in I} A_i$. Then we can define $f(i)$ to be the least element of A_i for each $i \in I$. □

4 Predicate Logic

4.1 Language

We now look at a more complicated version of the style of logic that we looked at before.

Definition 4.1 (Language). Let Ω (*function symbols*) and Π (*relation symbols*) be disjoint sets and $\alpha : \Omega \cup \Pi \to \mathbb{N}$ a function (*arity*).

The *Language* $L = L(\Omega, \Pi, \alpha)$ is the set of formulae, defined as follows:

- *Variables.* We have some variables $x_1, x_2, \ldots$ (also written $x, y, z, \ldots$)
- *Terms.* These are defined inductively by
  1. Every variable is a term;
  2. If $f \in \Omega$, $\alpha(f) = n$ and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term.
- *Atomic formulae.* There are three sorts:
  1. $\perp$;
  2. $(s = t)$ for any terms s, t;
  3. $\phi(t_1, \ldots, t_n)$ for any $\phi \in \Pi$ with $\alpha(\phi) = n$ and $t_1, \ldots, t_n$ terms.
- *Formulae.* These are defined inductively by
  1. Atomic formulae are formulae;
  2. $(p \Rightarrow q)$ is a formula for any formulae p, q;
  3. $(\forall x)p$ is a formula for any formula p and variable x.

Example 4.2 (The Language of Groups). We will give a prototypical example, the language of Groups. In this case, we have $\Omega = \{m, i, e\}$, $\Pi = \emptyset$ and $\alpha(m) = 2$, $\alpha(i) = 1$ and $\alpha(e) = 0$. Then $e, x_1, m(x_1, x_2), i(m(x_1, x_1))$ are terms and $m(e, e) = e$, $(\forall x)m(x, i(x)) = e$ are formulae.

Note that a formula is just a string of meaningless symbols. It doesn't make sense to ask if it's true or false. In particular the function and relation symbols are not assigned any meaning aside from its arity.

We have the usual abbreviations as before, and we also have $(\exists x)p$ for $\neg(\forall x)(\neg p)$.

Definition 4.3 (Closed Term). A term is *closed* if it has no variables.

Definition 4.4 (Free and Bound Variables). An occurrence of a variable x in a formula p is *bound* if it is inside brackets of a $(\forall x)$ quantifier. It is *free* otherwise.

Definition 4.5 (Sentence). A *sentence* is a formula with no free variables.

Definition 4.6 (Substitution). For a formula p, a variable x and a term t, the *substitution* $p[t/x]$ is obtained by replacing each free occurrence of x with t.

4.2 Semantic Entailment

In propositional logic we had 'valuations'. We are going to replace these with sets that have operations of the right arity, called a *structure*.

Definition 4.7 (Structure). An *L-struture* is a non-empty set A with a function $f_A : A^n \to A$ for each $f \in \Omega$, with $\alpha(f) = n$ and a relation $\phi_A \subseteq A^n$, for each $\phi \in \Pi, \alpha(\phi) = n$.

Remark. To make life easier, we explicitly forbid A from being empty.

We now want to define[4] 'p holds in A' for a sentence $p \in L$ and an L-structure A.

Definition 4.8 (Interpretation). To define the *interpretation* $p_A \in \{0, 1\}$ for each sentence p and L-structure A, we define inductively:

- *Closed terms.* Define $t_A \in A$ for each closed term t by
$$(f(t_1, \ldots, t_n))_A = f_a A(t_{1_A}, \ldots, t_{n_A})$$
for any $f \in \Omega$, $\alpha(f) = n$ and closed terms $t_1, \ldots, t_n$.
- *Atomic formulae.* We take
$$\perp_A = 0,$$
$$(s = t)_A = \mathbb{1}[s_A = t_A],$$
$$(\phi(t_1, \ldots, t_n))_A = \mathbb{1}[(t_{1_A}, \ldots, t_{n_A}) \in \phi_A]$$
- *Sentences.* We take
$$(p \Rightarrow q)_A = \mathbb{1}[\neg(p_A = 1, q_A = 0)]$$
$$((\forall x)p)_A = \mathbb{1}[p[\overline{a}/x]_{\overline{A}} \text{ for all } a \in A]$$
where for any $a \in A$, we define a new language L' be adding a constant $\overline{a}$ and make A into a L' structure $\overline{A}$ by setting $\overline{a}_{\overline{A}} = a$.

We can now define models and entailment.

Definition 4.9 (Theory). A *theory* is a set of sentences.

Definition 4.10 (Model). If a sentence p has $p_A = 1$, we say that p *holds* in A, or p is *true* in A, or A is a *model* of p. For a theory S, a model of S is a structure that is a model for each $s \in S$.

Definition 4.11 (Semantic Entailment). For a theory S and a sentence t, S *entails* t, written $S \models t$, if every model of S is also a model of t.

Definition 4.12 (Tautology). t is a *tautology*, written $\models t$, if $\emptyset \models t$.

Predicate logic is also called 'first-order logic', where first-order means we are ranging over *elements* of the structure, and not subsets.

4.3 Syntactic Implication

As before, we need axioms and deduction rules.

Definition 4.13 (Axioms of Predicate Logic).

1. $p \Rightarrow (q \Rightarrow p)$ for any formulae p, q;
2. $[p \Rightarrow (q \Rightarrow r)] \Rightarrow [(p \Rightarrow q) \Rightarrow (p \Rightarrow r)]$ for any formulae p, q;
3. $(\neg\neg p) \Rightarrow p$ for any formula p.
4. $(\forall x)(x = x)$ for any variable x.
5. $(\forall x)(\forall y)((x = y) \Rightarrow (p \Rightarrow p[y/x]))$ for any variable x, y and formula p, with y not occurring bound in p.
6. $[(\forall x)p] \Rightarrow p[t/x]$ for any formula p, variable x, term t with no free variable of t occurring bound in$p$.
7. $[(\forall x)(p \Rightarrow q)] \Rightarrow [p \Rightarrow (\forall x)q]$ for any formulae p, q with variable x not occurring free in p.

Definition 4.14 (Deduction Rules of Predicate Logic). The *deduction rules* are

1. *Modus ponens*: From p and $p \Rightarrow q$, we can deduce q.
2. *Generalization*: From r, we can deduce $(\forall x)r$ provided that no premise used in the proof so far had x as a free variable.

Definition 4.15 (Proof). A *proof* of p from S is a sequence of statements, in which each statement is either an axiom, a statement in S, or obtained via modus ponens or generalization.

Definition 4.16 (Syntactic Implication). If there exists a proof of a formula p from a set of formulae S, we write $S \vdash p$, and say '*S proves t*'.

Definition 4.17 (Theorem). If $S|p$, we say p is a *theorem* of S.

Now we prove the theorems we had for propositional logic.

Proposition 4.18 (Deduction Theorem). *Let $S \subseteq L$ and $p, q \in L$. Then $S \cup \{p\} \vdash q$ if and only if $S \vdash p \Rightarrow q$.*

Proof. The proof is exactly the same as the one for propositional logic, expect in the $\Rightarrow$ case, we have to check Gen. Suppose we have the lines r and $(\forall x)r$ (Gen), and we have a proof of $S \vdash p \Rightarrow r$ (by induction). We want to seek a proof of $p \Rightarrow (\forall x)r$ from S.

We know that no premise used in the proof of r from $S \cup \{p\}$ had x as a free variable, as required by the conditions of the use of Gen. Hence no premise used in the proof of $p \Rightarrow r$ from S has x as a free variable. Hence $S \vdash (\forall x)(p \Rightarrow r)$. If x is not free in p, then we get $S \vdash p \Rightarrow (\forall x)r$ by Axiom 7 (and MP). If x is free in p, then we did not use premise p in our proof r from $S \cup \{p\}$. So $S \vdash r$, and hence $S \vdash (\forall x)r$ by Gen. So $S \vdash p \Rightarrow (\forall x)r$. □

We also have the following whose proofs we omit.

Proposition 4.19 (Soundness Theorem). *Let S be a set of sentences, p a sentence. Then $S \vdash p$ implies $S \models p$.*

Theorem 4.20 (Model Existence Lemma). *Let S be a consistent set of sentences. Then S has a model.*

Corollary 4.21 (Adequacy Theorem). *Let S be a theory, and p a sentence. Then $S \models p$ implies $S \vdash p$.*

Corollary 4.22 (Compactness Theorem). *Let S be a theory such that every finite subset of S has a model. Then so does S.*

Proof. Trivial if we replace 'has a model' with 'is consistent', because proofs are finite. □

We can look at some applications of this: can we axiomatize the theory of finite groups (in the language of groups)?

Corollary 4.23. *The theory of finite groups cannot be axiomatized (in the language of groups).*

Proof. Suppose theory T has models all finite groups and nothing else. Let T' be T together with
$$(\exists x_1)(\exists x_2)(x_1 \neq x)$$
$$(\exists x_1)(\exists x_2)(\exists x_3)(x_1 \neq x_2 \neq x_3)$$
$$\vdots$$
then T' has no model, sine each model has to be simultaneously arbitrarily large and finite, but every finite subset of T' does have a model (say $\mathbb{Z}_n$ for some n), which is a contradiction. □

So 'finiteness is not a first-order property'.

Corollary 4.24. *Let S be a theory with arbitrarily large models. Then S has an infinite model.*

Proof. Same as above. □

Corollary 4.25 (Upward Löwenheim-Skolem Theorem). *Let S be a theory with an infinite model. Then S has an uncountable model.*

Proof. Add constants $\{c_i \mid i \in I\}$ to L for some uncountable I. Let $T = S \cup \{\text{'}c_i \neq c_j\text{'} \mid i, j \in I, i \neq j\}$. Then any finite $T' \subseteq T$ has a model, since it can only mention finitely many of the C_i. So any infinite model of S will do. Hence by compactness, T has a model. □

Theorem 4.26 (Downward Löwenhein-Skolem Theorem). *Let L be a countable language (i.e. Ω and Π are countable). Then if S has a model, then it has a countable model.*

Proof. The model constructed in the proof of model existence theorem is countable. □

5 Set Theory

5.1 Axioms of Set Theory

We now formulate set theory as first-order theory.

Definition 5.1 (Zermelo-Fraenkel Set Theory). *Zermelo-Fraenkel set theory* (ZF) has language $\Omega = \emptyset$, $\Pi = \{\in\}$, with arity 2.

The 'universe of sets' wil mean a model with these axioms, a pair (V, ε), where V is a set and $\in$ is a binary relation on V in which the axioms are true.

Axiom 5.2 (Axiom of Extension). 'Sets with the same members are equal':
$$(\forall x)(\forall y)[(\forall z)(z \in x \Leftrightarrow z \in y) \Rightarrow x = y].$$

Axiom 5.3 (Axiom of Separation). 'For a set x and a property p, we can form $\{z \in x \mid p(z)\}$':
$$(\forall t_1) \cdots (\forall t_n)(\forall x)(\exists y)(\forall z)(z \in y \Leftrightarrow z \in x \wedge p)$$
for each formula p with free variables[5] $t_1, \ldots, t_n, z$.

[4] After it's defined, we can pretty much forget about it.
[5] We do need parameters to say form $\{z \in x \mid z \in t\}$ for some variable t.

Axiom 5.4 (Axiom of Empty Sets)**.** 'There is an empty set':

$$(\exists x)(\forall y)[\neg y \in x].$$

We write $\emptyset$ for the (unique by extension) set guaranteed by this axiom.

Axiom 5.5 (Axiom of Pair Sets)**.** 'We can form $\{x, y\}$':

$$(\forall x)(\forall y)(\exists z)(\forall t)(t \in z \Leftrightarrow t = x \vee t = y).$$

We write $\{x, y\}$ for this z, and $\{x\}$ for $\{x, x\}$.

We can now pin down what functions are.

Definition 5.6 (Ordered Pair)**.** The *ordered pair* $(x, y) = \{\{x\}, \{x, y\}\}$. Clearly we have $(x, y) = (z, t)$ if and only if $x = z$ and $y = t$. We say x is an ordered pair of $(\exists y)(\exists z)(x = (y, z))$.

Definition 5.7 (Function)**.** We say f is a *function* if

$$(\forall x)(x \in f \Rightarrow x \text{ is an ordered pair})$$
$$\wedge (\forall x)(\forall y)(\forall z)[((x, y) \in f \wedge (x, z) \in f) \Rightarrow y = z].$$

Definition 5.8 (Domain)**.** Call x the *domain* of f, written $x = \operatorname{dom}(f)$ if (f is a function) $\wedge$ $(\forall y)(y \in x \Leftrightarrow (\exists z)((y, z) \in f))$.

Then $f : x \to y$ means (f is a function) $\wedge$ $(x = \operatorname{dom}(f)) \wedge (\forall z)(\forall t)((z, t) \in f \Rightarrow t \in y)$.

Back to axioms.

Axiom 5.9 (Axiom of Union)**.** 'We can form unions':

$$(\forall x)(\exists y)(\forall z)(z \in y \Leftrightarrow (\exists t)(t \in x \wedge z \in t)).$$

Note that in this definition, we think of $A \cup B \cup C$ as $\bigcup\{A, B, C\}$.

Axiom 5.10 (Axiom of Power Sets)**.** 'We can form power sets':

$$(\forall x)(\exists y)(\forall z)(z \in y \Leftrightarrow z \subseteq x),$$

where $z \subseteq x$ means $(\forall t)(t \in z \Leftrightarrow t \in x)$. We write $\mathcal{P}(x)$ for the set generated above.

Axiom 5.11 (Axiom of Infinity)**.** 'There is an infinite set':

$$(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \Rightarrow y^+ \in x)),$$

where $y^+ = y \cup \{y\}$. Any set that satisfies the above axiom is a successor set.

The intersection of successor sets is a successor set, so by taking the intersection of all successor sets we get a least successor set. Call this ω. Then in particular if $x \subseteq \omega$ is a successor set, then $x = \omega$.

We can use this to get induction in V:

$$(\forall x)[(x \subseteq \omega \wedge \emptyset \in x \wedge (\forall y)(y \in x \Rightarrow y^+ \in x)) \Rightarrow x = \omega].$$

Can now define 'x is finite' for $(\exists y)(y \in \omega \wedge x$ bijects with y) and 'x is countable' for (x is finite) $\vee$ (x bijects with ω).

Axiom 5.12 (Axiom of Foundation)**.** 'Every (non-empty) set has an $\in$-minimal element':

$$(\forall x)(x \neq \emptyset \Rightarrow (\exists y)(y \in x \wedge (\forall z)(z \in x \Rightarrow z \notin y))).$$

5.2 Digression on Classes

Now we want to be able to say something like 'for each $x \in I$, we have some A_i. Now take $\{A_i \mid i \in I\}$. We would like our result to be a set, but we don't know that this thing is a function yet. We try to define 'things that look like a function'.

Definition 5.13 (Class)**.** Let $(V, \in)$ be an L-structure. A *class* is a collection C of points in V such that, for some formula p with free variable x (and maybe more parameters), we have $x \in C \Leftrightarrow p$ holds.

Intuitively, everything of the form $\{x \in V \mid p(x)\}$ is a class.

Definition 5.14 (Proper Class)**.** We say C is a *proper class* if C is not a set (in V), that is,

$$\neg(\exists y)(\forall x)(x \in y \Leftrightarrow p).$$

Definition 5.15. A *function-class* F is a collection of ordered pairs such that there is a formula p with free variables x, y (and maybe more) such that

$$(x, y) \in F \Leftrightarrow p \text{ holds},$$
$$\text{and } (x, y) \in F \wedge (x, z) \in F \Rightarrow y = z.$$

5.3 Back to Axioms

We can now give our last axiom.

Axiom 5.16 (Axiom of Replacement)**.** 'The image of a set under a function-class is a set'. This is an axiom scheme, with an instance for each first-order formula p:

$$\underbrace{(\forall t_1) \dots (\forall t_n)}_{\text{parameters}}\underbrace{[(\forall x)(\forall y)(\forall z)(p \wedge p[z/y] \Rightarrow y = z)}_{p \text{ is a function class}}$$
$$\Rightarrow (\forall x)\ \underbrace{(\exists y)(\forall z)(z \in y \Leftrightarrow (\exists t)(t \in x \wedge p[t/x, z/y]))]}_{y \text{ is image of } x}$$

That is all of axioms in ZF. We didn't include the Axiom of Choice though.

Definition 5.17 (ZFC)**.** *ZFC* is the axioms ZF + AC, where AC is the axiom of choice, 'every family of non-empty sets has a choice function':

$$(\forall f)[(\forall x)(x \in \operatorname{dom} f \Rightarrow f(x) \neq \emptyset) \Rightarrow$$
$$(\exists g)(\operatorname{dom} g = \operatorname{dom} f) \wedge (\forall x)(x \in \operatorname{dom} g \Rightarrow g(x) \in f(x))].$$

Here we define a family of sets $\{A_i \mid i \in I\}$ to be a function $f : I \to V$ such that $i \mapsto A_i$.

5.4 Properties of ZF

Now we want to know what V looks like.

Definition 5.18 (Transitive Set)**.** A set x is *transitive* if every member of a member of x is a member of x:

$$(\forall y)((\exists z)(y \in z \wedge z \in x) \Rightarrow y \in x).$$

Lemma 5.19. *Every x is contained in a transitive set.*

Proof. We will form the set

$$x \cup \left(\bigcup x\right) \cup \left(\bigcup\bigcup x\right) \cup \left(\bigcup\bigcup\bigcup x\right) \cup \cdots,$$

as it will be clearly transitive and contains x. By the union axiom, it suffices to obtain the set $\{x, \bigcup x, \bigcup\bigcup x, \dots\}$. We can get this from the axiom of replacement, which we will apply to ω with the function-class $0 \mapsto x$, $1 \mapsto \bigcup x$, ... So we just need to show that this is indeed a function-class.

Define f as an *attempt* to mean

$$(f \text{ is a function}) \wedge (\operatorname{dom} f \in \omega) \wedge (\operatorname{dom} f \neq \emptyset)$$
$$\wedge (f(x) = 0) \wedge (\forall n)([(n \in \operatorname{dom} f) \wedge (n \neq 0)]$$
$$\Rightarrow f(n) = \bigcup f(n-1)).$$

We can check by usual ω-induction that

$$(\forall f)(\forall g)(\forall n)([(f \text{ an attempt}) \wedge (g \text{ an attempt})$$
$$\wedge (n \in \operatorname{dom} f) \wedge (n \in \operatorname{dom} g)] \Rightarrow f(n) = g(n)),$$

and also that

$$(\forall n)(n \in \omega \Rightarrow (\exists f)[$$
$$(f \text{ an attempt}) \wedge (n \in \operatorname{dom} f)]),$$

also by ω-induction using the constructions we had before. So we indeed have a function class which we can take as $p(y, z)$ where $p(y, z) = (\exists f)((f$ an attempt $\wedge (y \in \operatorname{dom} f) \wedge (f(y) = z))$. □

Remark. Officially, this says 'let $(V, \in)$ be a model of ZF. Then [statement]. Equivalently, ZF $\vdash$ [statement] (by the completeness theorem).' Also from this proof, we know that in particular x is contained in the *transitive closure* of x, the intersection of all transitive sets containing x, written $\mathrm{TC}(x)$ (as the intersection of transitive sets is transitive).

We want the axiom of foundation to capture that 'sets are built out of simpler sets'. With this, we should want: if $p(x)$ holds whenever $(\forall y \in x)p(y)$, then $p(x)$ holds for all x.

Theorem 5.20 (Principle of $\in$-Induction)**.** *For each formula p with free variables $t_1, \dots, t_n, x$:*

$$(\forall t_1) \dots (\forall t_n)\ [(\forall x)((\forall y)(y \in x \Rightarrow p(y))$$
$$\Rightarrow p(x)) \Rightarrow (\forall x)(p(x))]$$

Proof. Given $t_1, \dots, t_n$, suppose $\neg(\forall x)p(x)$. Then we have $\neg p(x)$ for some x.

Note that we would like to say: choose $\in$-minimal x with $\neg p(x)$, by foundation, and hence we have a contradiction. But $\{x \mid \neg p(x)\}$ need not be a set. Instead, we take $t = \mathrm{TC}(\{x\})$, and $u = \{y \in t \mid \neg p(y)\}$. Then $u \neq \emptyset$, so u has a $\in$-minimal member, say y. Then $\neg p(y)$, but $z \in y \Rightarrow z \in t$ by transitivity, so $z \notin u$, that is, $(\forall z \in y)p(z)$ which is a contradiction. □

Theorem 5.21. *The axiom of foundation and the principle of $\in$-induction are equivalent (in the presence of the other ZF axioms).*

Proof. We already wrote down a proof of the principle of $\in$-induction from foundation, now we prove foundation. Indeed, say that x is *regular* to mean $(\forall y)(x \in y \Rightarrow y$ has a minimal member). So foundation says every x is regular. To prove this by $\in$-induction, it is enough to show that if every $y \in x$ is regular then x is regular.

For z with $x \in z$, if x is minimal in z we are done. Otherwise, we have a $y \in x$ such that $y \in z$. So z has a minimal element (as y is regular). □

Now we also want to do recursion, so we can define $f(x)$ in terms of the $f(y)$, $y \in x$.

Theorem 5.22 ($\in$-Recursion Theorem)**.** *Let G be a function-class $((x, y) \in G \Leftrightarrow p(x, y)$ for some formula p), everywhere defined.*

Then there is a function-class F, $((x, y) \in F \Leftrightarrow q(x, y)$, some formula q), everywhere defined, such that $(\forall x)(F(x) = G(F|_x))$. Moreover, F is unique.

Proof. For existence, we say 'f is an attempt' if

$$(f \text{ is a function}) \wedge (\operatorname{dom} f \text{ is transitive})$$
$$\wedge (\forall x)(x \in \operatorname{dom} f \Rightarrow f(x) = G(f|_x)),$$

where this makes sense as $\operatorname{dom} x$ is transitive. Then $(\forall x)(\forall f)(\forall f')([(f, f'$ attempts$) \wedge (x \in \operatorname{dom} f) \wedge (x \in \operatorname{dom} f')] \Rightarrow f(x) = f'(x))$, by $\in$-induction (as if f and f' agree at all $y \in x$ then they agree at x).

Also, $(\forall x)(\exists f)((f$ is an attempt$) \wedge (x \in \operatorname{dom} f))$, again by $\in$-induction.

If each $y \in x$ has an attempt defined at y, then for each $y \in x$ there is a unique attempt f_y defined on $\mathrm{TC}(\{y\})$. Put $f = \bigcup\{f_y \mid y \in x\}$ and put $f' = f \cup \{(x, G(f|_x))\}$.

So define F by: $q(x, y) =$ '$(\exists f)((f$ an attept$) \wedge (x \in \operatorname{dom} f) \wedge (f(x) = y))'$.

For uniqueness, if we have suitable function-classes F, F' then $(\forall x)(F(x) = F'(x))$, by $\in$-induction. □

Remark. We note that $F|_x = \{(t, F(t)) \mid t \in x\}$ *is* a set, by replacement.

Notice that the proofs of $\in$-induction and $\in$-recursion look similar to induction and recursion on a well-ordered set. This may inspire a thought of what properties of the 'relation' $\in$, that is, the formula $p(x, y) =$ '$x \in y$' have we used?

1. p is *well-founded* – every non-empty set has a p-minimal member.
2. p is *local* – for each y, $\{x \mid p(x, y)\}$ is a set, so that we can build the transitive closure.

So for any $p(x, y)$ that is well-founded and local, we can prove p-induction and p-recursion. If r is a relation on a *set* a, then trivially r is local, so to have these we only need r well-founded. The theorems about well-orderings were a special case of this.

We have almost replicated all of our results about well-orderings, except for subset collapse. We will consider this now. The following definition is motivated by 'can we model a given relation on a set by $\in$'?

Definition 5.23 (Extensional)**.** We say a relation r on a set a is *extensional* if

$$(\forall x, y \in a)([(\forall z \in a)(z\ r\ x \Leftrightarrow z\ r\ y) \Rightarrow x = y]),$$

i.e. it obeys the axiom of extension.

Theorem 5.24 (Mostowski Collapse Theorem)**.** *Let r be a relation on a set a that is well-founded and extensional. Then there exists a transitive set b, and a bijection $f : a \to b$ such that $(\forall x, y \in a)(x\ r\ y \Leftrightarrow f(x) \in f(y))$. Moreover, b and f are unique.*

Proof. Define $f(x) = \{f(y) \mid y\ r\ x\}$, a definition by r-recursion on a. Note that f is a function by replacement (it is an image of a). Let $b = \{f(x) \mid x \in a\}$, which is a set by replacement.

Then b is transitive (by the definition of f), and f surjective (by the definition of b), so we need to just check that f is injective (then we also have $f(x) \in f(y) \Leftrightarrow x\ r\ y$). We shall show that $(\forall y \in a)(f(x) = f(y) \Rightarrow x = y)$ for each $x \in a$, by r-induction on x.

So suppose $f(x) = f(y)$, and that $(\forall t\ r\ x)(\forall y \in a)(f(t) = f(y) \Rightarrow t = u)$. We have $\{f(t) \mid t\ r\ x\} = \{f(u) \mid u\ r\ y\}$ (by the definition of f), so $\{t \mid t\ r\ x\} = \{u \mid u\ r\ y\}$ by our induction hypothesis, so $x = y$ by extensionality.

For uniqueness, if f, f' are suitable then $f(x) = f'(x)$ for all $x \in a$ by r-induction. □

Now we previously defined ordinals as kind of the 'equivalence class' of all well-orderings, but this was a problem since the ordinals wouldn't be sets. We define them formally as follows:

Definition 5.25 (Ordinal)**.** An *ordinal* is a transitive set, totally ordered by $\in$.

This is automatically well-ordered by $\in$, by foundation. Note then that $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}$ are all ordinals, any $n \in \omega$ as $n = \{0, 1, \dots, n-1\}$ as well as ω itself is an ordinal.

So by Mostowski, a well-ordering is order-isomorphic to a unique ordinal, and we call that ordinal the *order-type* of this well-ordering. So well orderings x and y are order-isomorphic if and only if they have teh same order type.

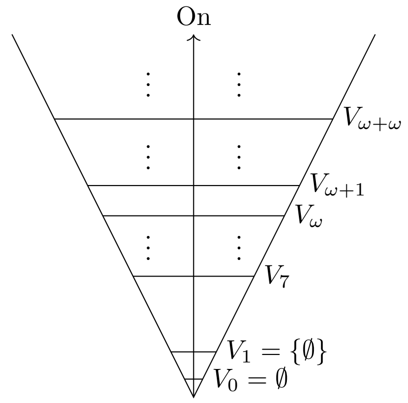
5.5 Picture of the Universe

We now want to see what V looks like.

Definition 5.26 (Von Neumann Hierarchy)**.** Define sets V_α for $\alpha \in \mathrm{On}$ (the class of ordinals) by $\in$-recursion:

$$\begin{aligned} V_0 &= \emptyset, \\ V_{\alpha^+} &= \mathcal{P}(V_\alpha), \\ V_\lambda &= \bigcup\{V_\gamma \mid \gamma < \lambda\}, \end{aligned}$$

for λ a non-zero limit ordinal.



Note that $x \subseteq V_\alpha \Leftrightarrow x \in V_{\alpha+1}$. We would like every x to be in some V_α, and that is indeed true.

Lemma 5.27. *Each V_α is transitive.*

Proof. We induct on α. We have V_0 transitive. If V_α is transitive then $\mathcal{P}(V_\alpha)$ is transitive, as if $x \in y \in \mathcal{P}(V_\alpha)$, then $y \subset V_\alpha$, so $x \in V_\alpha$, so $x \subset V_\alpha$ (as V_α transitive), so $x \in \mathcal{P}(V_\alpha)$. Finally, the union of transitive sets is transitive finishes our last case of limit ordinals. □

Lemma 5.28. *If $\alpha \leq \beta$, then $V_\alpha \subseteq V_\beta$.*

Proof. Fix α and induct on β. If $\beta = \alpha$, we are done. Given $V_\alpha \subset V_\beta$, we have $V_\beta \subset \mathcal{P}(V_\beta)$ (as $x \in V_\beta$ implies $x \subset V_\beta$, as V_β is transitive), so $V_\alpha \subset \mathcal{P}(V_\beta) = V_{\beta^+}$. And for limits this is trivial by definition. □

Theorem 5.29. *Every x belongs to some V_α.*

We first need to note that $x \subset V_\alpha \Leftrightarrow x \in V_{\alpha+1}$, and if $x \subset V_\alpha$ then there is a least such α called the *rank* of x.

Proof. We'll show that $(\forall x)(\exists \alpha)\,(x \in V_\alpha)$ by $\in$-induction on x.

So we are allowed to assume that for each $y \in x$, we have $y \subseteq V_\alpha$ for some α.

So $y \subseteq V_{\mathrm{rank}(y)}$, or $y \in V_{\mathrm{rank}(y)+1}$. Let $\alpha = \sup\{(\mathrm{rank}(y)^+ : y \in x\}$. Then $y \in V_\alpha$ for every $y \in x$. So $x \subseteq V_\alpha$ □

We will take the official definition of rank to be

Definition 5.30 (Rank)**.** The *rank* of a set x is defined recursively by

$$\mathrm{rank}(x) = \sup\{(\mathrm{rank}\, y)^+ \mid y \in x\}.$$

Proposition 5.31. $\mathrm{rank}(x)$ *is the first α such that $x \subseteq V_\alpha$.*

6 Cardinals

6.1 Basic Definitions

We now look at the 'size of sets', working in ZFC.

Remark (Notation)**.** We will write $x \leftrightarrow y$ for $(\exists f)$ (f is a bijection from x to y).

Definition 6.1 (Cardinality)**.** The *cardinality* of a set x, written $\mathrm{card}(x)$, is the least ordinal α such that $x \leftrightarrow \alpha$.

Remark (Scott Trick)**.** If we are just in ZF, we define the *essential rank* of x to be the least rank of all y such that $y \leftrightarrow x$. Then set $\mathrm{card}(x) = \{y \in V_{\mathrm{essrank}(x)^+} \mid y \leftrightarrow x\}$.

Definition 6.2 (Cardinal)**.** We say m is a *cardinal* if $m = \mathrm{card}\, x$, for some x.

6.2 The Alephs

We want to know what the cardinalities of the ordinals are.

Definition 6.3 (Initial Ordinal)**.** We say an ordinal α is *initial* if $(\forall \beta < \alpha)(\neg\beta \leftrightarrow \alpha)$, i.e. it is the smallest ordinal of that cardinality.

Definition 6.4 (Omega Ordinals)**.** We define ω_α for $\alpha \in \mathrm{On}$ by

$$\begin{aligned} \omega_0 &= \omega, \\ \omega_{\alpha+1} &= \gamma(\omega_\alpha), \\ \omega_\lambda &= \sup\{\omega_\alpha \mid \alpha < \lambda\}, \end{aligned}$$

for a non-zero limit ordinal λ.

Each ordinal ω_α is initial (by induction), and *every* initial δ (for $\delta \geq \omega$) is an ω_α. Indeed, the ω_α are unbounded in the ordinals, and taking the least α with $\delta \leq \omega_\alpha$ must have $\delta = \omega_\alpha$ by definition of the ω_α.

Definition 6.5 (Aleph Number)**.** We write $\aleph_\alpha$ for $\mathrm{card}(\omega_\alpha)$.

From the argument above we have

Theorem 6.6. *The $\aleph_\alpha$ are the cardinals of all infinite sets (or, in ZF, the cardinals of all infinite well-orderable sets).*

We will use lower case letters to denote cardinalities and upper case for the sets with that cardinality.

Definition 6.7 (Cardinal Inequality)**.** For cardinals n, m, we write $m \leq n$ if M injects into N, where $\mathrm{card}\, M = m$ and $\mathrm{card}\, N = n$.

So $m \leq n$ and $n \leq m$ implies $n = m$ by Schröder-Bernstein. Write $m < n$ if $m \leq n$ but $m \neq n$.

6.3 Cardinal Arithmetic

We can do arithmetic.

Definition 6.8 (Cardinal Arithmetic)**.** For cardinals m, n, write $m+n$ for $\mathrm{card}(M \sqcup N)$. Write mn for $\mathrm{card}(M \times N)$. Finally, write m^n for $\mathrm{card}(M^N)$, where $M^N = \{f \mid f \text{ is a function } N \to M\}$.

Example 6.9. $\mathbb{R} \leftrightarrow \mathcal{P}(\omega) \leftrightarrow 2^\omega$. So $\mathrm{card}(\mathbb{R}) = \mathrm{card}(\omega) = 2^{\aleph_0}$.

Example 6.10. How many sequences of reals are there? A real sequence is a function from $\omega \to \mathbb{R}$. We have $\mathrm{card}(\mathbb{R}^\omega) = (2^{\aleph_0})^{\aleph_0} = 2^{\alpha_0 \times \aleph_0} = 2^{\alpha_0} = \mathrm{card}(\mathbb{R})$.

We know from countability that $\aleph_0\aleph_0 = \aleph_0$. It turns out generally sums and multiplications are not interesting.

Theorem 6.11. *For every ordinal α, $\aleph_\alpha\aleph_\alpha = \aleph_\alpha$.*

Proof. We'll show $\aleph_\alpha^2 = \aleph_\alpha$ for all α by induction. Define a well-ordering of $\omega_\alpha \times \omega_\alpha$ by 'going up in squares': $(x, y) < (z, w)$ if either $\max(x, y) < \max(z, w)$ or $\max(x, y) = \max(z, w) = \beta$, with $y < \beta, z < \beta$ or $x = z = \beta, y < w$ or $y = w = \beta, x < z$.

For any $\delta \in \omega_\alpha \times \omega_\alpha$, have $\delta \in \beta \times \beta$ for some $\beta < \omega_\alpha$. Hence by induction, have $\beta \times \beta \leftrightarrow \beta$ (or β is finite). So the initial segment I_δ is contained in $\beta \times \beta$, so $\mathrm{card}\,(I_\delta) \leq \mathrm{card}(\beta) < \mathrm{card}\,(\omega_\alpha)$. Hence our well-ordering has order-type at most ω_α. So $\omega_\alpha \times \omega_\alpha \hookrightarrow \omega_\alpha$. Clearly $\omega_\alpha \hookrightarrow \omega_\alpha \times \omega_\alpha$ so $\omega_\alpha \leftrightarrow \omega_\alpha \times \omega_\alpha$. □

Corollary 6.12. *For any ordinals α, β with $\alpha \leq \beta$ we have $\aleph_\alpha + \aleph_\beta = \aleph_\alpha\aleph_\beta = \aleph_\beta$.*

Proof. $\aleph_\beta \leq \aleph_\alpha + \aleph_\beta \leq 2\aleph_\beta \leq \aleph_\alpha\aleph_\beta \leq \aleph_\beta^2 = \aleph_\beta$. □

In general, cardinal *exponentiation* is hard. In ZFC, $2^{\aleph_0} = \aleph_1$ is independent of the axioms (the *Continuum Hypothesis*). ZFC does not even decide if $2^{\aleph_0} < 2^{\aleph_1}$.