Logic and Set Theory

Lectured by I. B. Leader, Lent Term 2005, 2010

Prerequisites. Have met groups, fields, vector spaces (as examples); countability

There are **four** examples sheets.

Books. 1. P. T. Johnstone, 'Notes on Logic & Set Theory', CUP 1987

2. D. Van Dalen, 'Logic and Structure', Springer-Verlag 1980 (good for Chapter 4)

3. A. Hajnal & P. Hamburger, 'Set Theory', CUP 1999 (for cardinals and ordinals)

4. T. Forster, 'Logic, Induction and Sets', CUP 2003 (good bedtime read)

Last updated: Tue 28th Oct, 2014

Please let me know of any corrections: glt1000@cam.ac.uk

Course schedule

LOGIC AND SET THEORY (D) **24 lectures, Lent term**

No specific prerequisites.

Ordinals and cardinals
Well-orderings and order-types. Examples of countable ordinals. Uncountable ordinals and Hartogs' lemma. Induction and recursion for ordinals. Ordinal arithmetic. Cardinals; the hierarchy of alephs. Cardinal arithmetic. [5]

Posets and Zorn's lemma
Partially ordered sets; Hasse diagrams, chains, maximal elements. Lattices and Boolean algebras. Complete and chain-complete posets; fixed-point theorems. The axiom of choice and Zorn's lemma. Applications of Zorn's lemma in mathematics. The well-ordering principle. [5]

Propositional logic
The propositional calculus. Semantic and syntactic entailment. The deduction and completeness theorems. Applications: compactness and decidability. [3]

Predicate logic
The predicate calculus with equality. Examples of first-order languages and theories. Statement of the completeness theorem; *sketch of proof*. The compactness theorem and the Löwenheim-Skolem theorems. Limitations of first-order logic. Model theory. [5]

Set theory
Set theory as a first-order theory; the axioms of ZF set theory. Transitive closures, epsilon-induction and epsilon-recursion. Well-founded relations. Mostowski's collapsing theorem. The rank function and the von Neumann hierarchy. [5]

Consistency
Problems of consistency and independence. [1]

Appropriate books

B.A. Davey and H.A. Priestley *Lattices and Order*. Cambridge University Press 2002 (19.95 paperback).
T. Forster *Logic, Induction and Sets*. Cambridge University Press (50.00 hardback).
A. Hajnal and P. Hamburger *Set Theory*. LMS Student Texts number 48, CUP 1999 (55.00 hardback, 22.99 paperback).
A.G. Hamilton *Logic for Mathematicians*. Cambridge University Press 1988 (25.95 paperback).
P.T. Johnstone *Notes on Logic and Set Theory*. Cambridge University Press 1987 (15.95 paperback).
D. van Dalen *Logic and Structure*. Springer-Verlag 1994 (18.50 paperback).

Chapter 1 : Propositional Logic

Let P be a set of **primitive propositions**. Unless otherwise stated, $P = \{p_1, p_2, p_3, \ldots\}$. The set of **propositions**, written L or $L(P)$, is defined inductively by:

(i) if $p \in P$ then $p \in L$,

(ii) $\bot \in L$ ('$\bot$' is read 'false'),

(iii) if $p, q \in L$ then $(p \Rightarrow q) \in L$.

Examples. $(p_1 \Rightarrow \bot)$, $((p_1 \Rightarrow p_2) \Rightarrow (p_1 \Rightarrow p_3))$, $((p \Rightarrow \bot) \Rightarrow \bot)$.

Notes. 1. A proposition is a finite string of symbols from the alphabet: $\bot \ \Rightarrow \ (\ )\ p_1\ p_2 \ldots$
(Often omit the outer brackets, or use different brackets [ ] for clarity.)

2. 'L defined inductively' means, more precisely, that we set $L_1 = \{\bot\} \cup P$, and for $n \geqslant 1$, $L_{n+1} = L_n \cup \{(p \Rightarrow q) : p, q \in L_n\}$, and then $L = L_1 \cup L_2 \cup \ldots$.
(So L_n = 'things born in time n'.)

3. Every proposition is built up from 1 and 2 using 3 in a *unique* way.
E.g. $[(p_1 \Rightarrow p_2) \Rightarrow (p_1 \Rightarrow p_3)]$ came from $(p_1 \Rightarrow p_2)$ and $(p_1 \Rightarrow p_3)$.

Can now define, for example,

$\neg p$	('not p')	as an abbreviation for	$(p \Rightarrow \bot)$
$p \vee q$	('p or q')	as an abbreviation for	$(\neg p) \Rightarrow q$
$p \wedge q$	('p and q')	as an abbreviation for	$\neg(p \Rightarrow (\neg q))$

Semantic Implication

A **valuation** on L is a function $v : L \to \{0, 1\}$ such that:

(i) $v(\bot) = 0$,

(ii) $v(p \Rightarrow q) = \begin{cases} 0 \text{ if } v(p) = 1,\ v(q) = 0, \\ 1 \text{ otherwise,} \end{cases}$ for all $p, q \in L$.

Remark. On $\{0, 1\}$, can define a constant $\bot$ by $\bot = 0$, and an operation $\Rightarrow$ by

$$(a \Rightarrow b) = \begin{cases} 0 \text{ if } a = 1,\ b = 0, \\ 1 \text{ otherwise.} \end{cases}$$

Then a valuation is precisely a map $v : L \to \{0, 1\}$ that preserves the structure ($\bot$ and $\Rightarrow$), i.e. a homomorphism.

Proposition 1. (i) If v, v' are valuations with $v(p) = v'(p)$ for all $p \in P$, then $v = v'$.

(ii) For any function $w : P \to \{0, 1\}$, there exists a valuation v such that $v(p) = w(p)$ for all $p \in L$.

'A valuation is determined by its values on P, and any values will do.'

Proof. (i) We have $v(p) = v'(p)$ for all $p \in L_1$. But if $v(p) = v'(p)$ and $v(q) = v'(q)$ then $v(p \Rightarrow q) = v'(p \Rightarrow q)$, so $v(p) = v'(p)$ for all $p \in L_2$.
Continuing inductively, we get $v(p) = v'(p)$ for all $p \in L_n$, all n.

(ii) Set $v(p) = w(p)$ for each $p \in P$, and $v(\bot) = 0$; this defines v on L_1. Having defined v on L_n, use

$$v(p \Rightarrow q) = \begin{cases} 0 & \text{if } v(p) = 1,\ v(q) = 0 \\ 1 & \text{otherwise} \end{cases}$$

to define v on L_{n+1}. □

E.g. Let v be the valuation given by: $v(p_1) = v(p_2) = 1$, and $v(p_n) = 0$ for $n \geqslant 3$. Then

$$v(\ (\underbrace{p_1 \Rightarrow p_2}_{1}) \Rightarrow \underbrace{p_3}_{0}\ ) = 0.$$

Say t is a **tautology**, written $\models t$, if $v(t) = 1$ for all v.

Examples.

1. $p \Rightarrow (q \Rightarrow p)$ ('a true statement is implied by anything')

$v(p)$	$v(q)$	$v(q \Rightarrow p)$	$v(p \Rightarrow (q \Rightarrow p))$	
1	1	1	1	
1	0	1	1	⟵ so a tautology, as last column identically 1
0	1	0	1	
0	0	1	1	

2. $(\neg\neg p) \Rightarrow p$, i.e. $((p \Rightarrow \bot) \Rightarrow \bot) \Rightarrow p$ ('law of the excluded middle')

$v(p)$	$v(p \Rightarrow \bot)$	$v((p \Rightarrow \bot) \Rightarrow \bot)$	$v(((p \Rightarrow \bot) \Rightarrow \bot) \Rightarrow p)$	
1	0	1	1	⟵ so a tautology
0	1	0	1	

3. $\big(p \Rightarrow (q \Rightarrow r)\big) \Rightarrow \big((p \Rightarrow q) \Rightarrow (p \Rightarrow r)\big)$

   Suppose not a tautology. Have v with $v\big(p \Rightarrow (q \Rightarrow r)\big) = 1,\ v\big((p \Rightarrow q) \Rightarrow (p \Rightarrow r)\big) = 0$. Then $v(p \Rightarrow q) = 1$ and $v(p \Rightarrow r) = 0$. Thus $v(p) = 1,\ v(r) = 0$, so $v(q) = 1$, so $v(q \Rightarrow r) = 0$, and so $v\big(p \Rightarrow (q \Rightarrow r)\big) = 0$. ⨳

For $S \subset L, t \in L$, say S **entails** t, or **semantically implies** t, written $S \models t$, if

$$v(s) = 1 \text{ for all } s \in S \ \text{ implies } \ v(t) = 1.$$

'Whenever all of S is true, t is true as well.'

E.g. $\{p \Rightarrow q, q \Rightarrow r\} \models (p \Rightarrow r)$.
Need: any valuation v with $v(p \Rightarrow q) = 1,\ v(q \Rightarrow r) = 1$ has $v(p \Rightarrow r) = 1$.
If not, then $v(p \Rightarrow r) = 0$, whence $v(p) = 1,\ v(r) = 0$, so $v(q) = 0$, as $v(q \Rightarrow r) = 1$.
So $v(p \Rightarrow q) = 0$. ⨳

If $v(t) = 1$, say t is **true** in v, or v is a **model** of t. For $S \subset L$, a valuation v is a **model** of S if $v(s) = 1$ for all $s \in S$. Thus $S \models t$ says 'every model of S is a model of t'.

Note. $\models t$ means exactly $\emptyset \models t$.

Syntactic Implication

For a notion of proof, we shall need some axioms and deduction rules. As axioms, we shall take all propositions of the following form.

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

Note. These are all tautologies. Sometimes, they are called 'axiom-schemes', as each is infinitely many axioms.

As our deduction rule, we shall use only **modus ponens**: 'from p and $p \Rightarrow q$, can deduce q'.

For $S \subset L$, $t \in L$, a **proof** of t from S is a finite sequence $t_1, t_2, \ldots, t_n$ of propositions, with $t_n = t$, such that each t_i is either

(i) an axiom,

(ii) a member of S,

(iii) such that there exist $j, k < i$ with $t_j = (t_k \Rightarrow t_i)$.

If there exists a proof of t from S, say S **proves** t, or **syntactically implies** t, written $S \vdash t$. If $\emptyset \vdash t$, say t is a **theorem**, written $\vdash t$. In a proof, S consists of the **hypotheses** or **premises**, and t is the **conclusion**.

Example. $\{p \Rightarrow q,\ q \Rightarrow r\} \vdash (p \Rightarrow r)$ ('go for $(p \Rightarrow q) \Rightarrow (p \Rightarrow r)$')

1.	$[p \Rightarrow (q \Rightarrow r)] \Rightarrow [(p \Rightarrow q) \Rightarrow (p \Rightarrow r)]$	(axiom 2)
2.	$q \Rightarrow r$	(hypothesis)
3.	$(q \Rightarrow r) \Rightarrow [p \Rightarrow (q \Rightarrow r)]$	(axiom 1)
4.	$p \Rightarrow (q \Rightarrow r)$	(modus ponens on 2, 3)
5.	$(p \Rightarrow q) \Rightarrow (p \Rightarrow r)$	(modus ponens on 1, 4)
6.	$p \Rightarrow q$	(hypothesis)
7.	$p \Rightarrow r$	(modus ponens on 5,6)

Example. $\vdash (p \Rightarrow p)$ ('go for $(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)$, using axiom 2')

1.	$\big(p \Rightarrow [(p \Rightarrow p) \Rightarrow p]\big) \Rightarrow \big([p \Rightarrow (p \Rightarrow p)] \Rightarrow (p \Rightarrow p)\big)$	(axiom 2)
2.	$p \Rightarrow [(p \Rightarrow p) \Rightarrow p]$	(axiom 1)
3.	$[p \Rightarrow (p \Rightarrow p)] \Rightarrow (p \Rightarrow p)$	(modus ponens on 1, 2)
4.	$p \Rightarrow (p \Rightarrow p)$	(axiom 1)
5.	$p \Rightarrow p$	(modus ponens on 3, 4)

In showing $S \vdash p$, often helpful to use the following.

Proposition 2 (Deduction Theorem). Let $S \subset L,\ p, q \in L$. Then $S \vdash (p \Rightarrow q)$ if and only if $S \cup \{p\} \vdash q$.

'Provability corresponds to the connective '$\Rightarrow$' in the language.'

Proof. ($\Rightarrow$) Given a proof of $p \Rightarrow q$ from S, write down the lines

$$\begin{array}{ll} p & \text{(hypothesis)} \\ q & \text{(modus ponens)} \end{array}$$

to obtain a proof of q from $S \cup \{p\}$.

($\Leftarrow$) Let $t_1, \ldots, t_n$ be a proof of q from $S \cup \{p\}$. We'll show that $S \vdash (p \Rightarrow t_i)$ for all i.

1. If t_i is an axiom, write down

$$\begin{array}{ll} t_i \Rightarrow (p \Rightarrow t_i) & \text{(axiom 1)} \\ t_i & \text{(axiom)} \\ p \Rightarrow t_i & \text{(modus ponens)} \end{array}$$

2. If $t_i \in S$, write down

$$\begin{array}{ll} t_i \Rightarrow (p \Rightarrow t_i) & \text{(axiom 1)} \\ t_i & \text{(hypothesis)} \\ p \Rightarrow t_i & \text{(modus ponens)} \end{array}$$

   showing $S \vdash (p \Rightarrow t_i)$.

3. If $t_i = p$, then certainly $S \vdash (p \Rightarrow p)$, as $\vdash (p \Rightarrow p)$.
4. If t_i obtained by modus ponens, then we have earlier lines t_j and $t_k = (t_j \Rightarrow t_i)$. By induction, may assume $S \vdash (p \Rightarrow t_j)$ and $S \vdash (p \Rightarrow (t_j \Rightarrow t_i))$, so write down

$$\begin{array}{ll} [p \Rightarrow (t_j \Rightarrow t_i)] \Rightarrow [(p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)] & \text{(axiom 2)} \\ (p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i) & \text{(modus ponens)} \\ p \Rightarrow t_i & \text{(modus ponens)} \end{array}$$

   ('This is the real reason why axiom 2 is the way it is.') □

For example, to show $\{p \Rightarrow q, q \Rightarrow r\} \vdash (p \Rightarrow r)$, enough by the Deduction Theorem to show that $\{p \Rightarrow q, q \Rightarrow r, p\} \vdash r$, which is trival by modus ponens twice.

Question. How are $\vdash$ and $\models$ related?

Aim. Completeness Theorem: $S \vdash t \iff S \models t$.

This is made up of **soundness** (if $S \vdash t$ then $S \models t$) and **adequacy** (if $S \models t$ then $S \vdash t$).

Soundness says 'our axioms are not absurd', and adequacy says 'our set of axioms is strong enough to prove, from S, *every* semantic consequence of S'.

Proposition 3 (Soundness Theorem). Let $S \subset L,\ t \in L$. Then $S \vdash t \Rightarrow S \models t$.

Proof. We have a proof of t from S. We must show that if v is a valuation with $v(s) = 1$ for all $s \in S$ (i.e. v is a model of S) then $v(t) = 1$ (i.e. v is a model of t).

But $v(p) = 1$ 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 $v(p) = 1,\ v(p \Rightarrow q) = 1$ then $v(q) = 1$.

Hence, each line t_i of a proof of t from S has $v(t_i) = 1$. □

Say S is **consistent** if $S \nvdash \bot$.

A special case of adequacy is: $S \models \bot \Rightarrow S \vdash \bot$ (i.e. 'S has no model' $\Rightarrow$ 'S is inconsistent'). Or, in other words: S is consistent $\Rightarrow$ S has a model.

In fact, this would imply adequacy in general. Indeed, given $S \models t$, have that $S \cup \{\neg t\}$ has no model, so we should know $S \cup \{\neg t\} \vdash \bot$, whence:

$$\begin{array}{rll} & S \vdash ((\neg t) \Rightarrow \bot) & \text{(by deduction theorem)} \\ \text{i.e.} & S \vdash (\neg\neg t) & \\ \text{but} & S \vdash ((\neg\neg t) \Rightarrow t) & \text{(axiom 3)} \\ \text{so} & S \vdash t & \text{(modus ponens)} \end{array}$$

So, for adequacy, we must show: S is consistent $\Rightarrow$ S has a model. How might we show this? Given a consistent set S, how can we 'build' a valuation v with $v(s) = 1\ \forall s \in S$? We cannot just set $v(p) = 1$ if $p \in S$ and $v(p) = 0$ if not – since if for example p_3 is not mentioned in S, then we would be setting $v(p_3) = 0,\ v(\neg p_3) = 0$. ⧓

Theorem 4 (Model existence lemma). Let $S \subset L$ be consistent. Then S has a model.

Idea. We want to define $v(p) = 1$ if $p \in S$, and $v(p) = 0$ if not. But this fails if some p has $p \notin S,\ \neg p \notin S$. So we shall try to extend S, keeping it consistent, to 'swallow up' one of p and $\neg p$, for each p.

Proof. First, for any consistent $S \subset L$ and $p \in L$, either $S \cup \{p\}$ or $S \cup \{\neg p\}$ is consistent. For if not, then $S \cup \{p\} \vdash \bot$ and $S \cup \{\neg p\} \vdash \bot$. But then $S \vdash (p \Rightarrow \bot)$ (deduction theorem), i.e. $S \vdash \neg p$, so $S \vdash \bot$. ⧓

Now, L is countable (e.g. as each L_n is countable), so we can list L as $\{t_1, t_2, t_3, \ldots\}$. Set $S_0 = S$. Then set $S_1 = S_0 \cup \{t_1\}$ or $S_0 \cup \{\neg t_1\}$ such that S_1 is consistent, then set $S_2 = S_1 \cup \{t_2\}$ or $S_1 \cup \{\neg t_2\}$ such that S_2 is consistent, and continue inductively.

Let $\overline{S} = S_0 \cup S_1 \cup S_2 \cup \ldots$. Then, for each $p \in L$, we have $p \in \overline{S}$ or $\neg p \in \overline{S}$. Also, $\overline{S}$ is consistent: if $\overline{S} \vdash \bot$ then, as proofs are finite, we have $S_n \vdash \bot$ for some n. ⧓

Note. $\overline{S}$ is **deductively closed**, meaning that if $\overline{S} \vdash p$ then $p \in \overline{S}$. Indeed, if $p \notin \overline{S}$ then $\neg p \in \overline{S}$, so $\overline{S} \vdash (p \Rightarrow \bot)$ and $\overline{S} \vdash p$, whence $\overline{S} \vdash \bot$. ⧓

Define $v : L \to \{0, 1\},\ p \mapsto \begin{cases} 1 & \text{if } p \in \overline{S} \\ 0 & \text{if } p \notin \overline{S} \end{cases}$

Claim. v is a valuation.

Proof of claim. $v(\bot) = 0$ as $\bot \notin \overline{S}$ (since $\overline{S}$ is consistent). For $v(p \Rightarrow q)$:

1. If $v(p) = 1,\ v(q) = 0$, we have $p \in \overline{S},\ q \notin \overline{S}$, and we want $(p \Rightarrow q) \notin \overline{S}$ (i.e. $v(p \Rightarrow q) = 0$). But if $(p \Rightarrow q) \in \overline{S}$, then $\overline{S} \vdash q$ (modus ponens), whence $q \in \overline{S}$ (since $\overline{S}$ is deductively closed). ⧓
2. If $v(q) = 1$, we have $q \in \overline{S}$, and we want $(p \Rightarrow q) \in \overline{S}$. But $\vdash q \Rightarrow (p \Rightarrow q)$ (axiom 1), so $\overline{S} \vdash (p \Rightarrow q)$, whence $(p \Rightarrow q) \in \overline{S}$ (as $\overline{S}$ is deductively closed).
3. If $v(p) = 0$, we have $\neg p \in \overline{S}$, and we want $(p \Rightarrow q) \in \overline{S}$. So it is enough to show that $(p \Rightarrow \bot) \vdash (p \Rightarrow q)$. By the deduction theorem, it is enough to show that $\{p \Rightarrow \bot, p\} \vdash q$. So it is enough to show that $\bot \vdash q$. But $\vdash (\bot \Rightarrow \neg\neg q)$ (axiom 1), and $\vdash (\neg\neg q \Rightarrow q)$ (axiom 3), so $\vdash \bot \Rightarrow q$. □

Remarks. 1. Sometimes Theorem 4 is also called the 'Completeness Theorem'.

2. The proof used that P is countable (to get L countable). In fact, Theorem 4 remains true for any P. We shall see this later – it needs Zorn's Lemma.

By the remark before theorem 4, we now have:

Corollary 5 (Adequacy Theorem). Let $S \subset L,\ t \in L$. Then $S \models t$ implies $S \vdash t$. □

Putting Proposition 3 and Corollary 5 together, we get:

Theorem 6 (Completeness Theorem). Let $S \subset L,\ t \in L$. Then $S \vdash t \Leftrightarrow S \models t$.

Proof. ($\Rightarrow$) Soundness. □
($\Leftarrow$) Adequacy.

Corollary 7 (Compactness Theorem). Let $S \subset L,\ t \in L$. Then if $S \models t$ then some finite $S' \subset S$ has $S' \models t$.

Proof. Trivial if we replace $\models$ by $\vdash$ (as proofs are finite). □

In particular, if $S \models \bot$ (S has no model) then some finite $S' \subset S$ has $S' \models \bot$ (S' has no model). Equivalently, if every finite subset of S has a model, then S has a model. (Useful form of compactness.)

This is actually equivalent to Corollary 7, because $S \models t$ is the same as '$S \cup \{\neg t\}$ has no model', and $S' \models t$ is the same as '$S' \cup \{\neg t\}$ has no model'.

Corollary 7′ (Compactness Theorem, equivalent form). Let $S \subset L$. If every finite subset of S has a model, then so does S.

Another consequence of completeness is:

Corollary 8 (Decidability Theorem). For finite $S \subset L$ and $t \in L$, there is an algorithm to determine in finite time whether or not $S \vdash t$.

Remark. Highly non-obvious.

Proof. Obvious if we replace $\vdash$ by $\models$. To check if $S \models t$, just do a truth table. □

Chapter 2 : Well-Orderings and Ordinals

A **total order** or a **linear order** is a pair $(X, <)$, where X is a set and $<$ is a relation on X that is:

(i) irreflexive : not $x < x \quad (\forall x \in X)$,

(ii) transitive : $x < y,\ y < z \Rightarrow x < z \quad (\forall x, y, z \in X)$,

(iii) trichotomous : $x < y$ or $x = y$ or $y < x \quad (\forall x, y \in X)$.

(Note: in (iii), we cannot have more than one: for if $x < y$, $y < x$ then $x < x$ ⨳.)

Examples. 1. $\mathbb{N}$, usual $<$. (Note: $\mathbb{N} = \{0, 1, 2, \ldots\}$, so $0 \in \mathbb{N}$. Write $\mathbb{N}^+$ for $\mathbb{N} \setminus \{0\}$.)

2. $\mathbb{Q}$, usual $<$.

3. $\mathbb{R}$, usual $<$.

4. $\mathbb{N}^+$, '$a < b$' if $a|b$ (and $a \neq b$) : *not* trichotomous.

5. $\mathbb{P}(S)$, '$a < b$' if $a \subset b$ (and $a \neq b$) : *not* trichotomous (if $|S| > 1$).

Write $y > x$ for $x < y$, and $x \leqslant y$ for '$x < y$ or $x = y$'. In terms of $\leqslant$, a total ordering is:

(i) reflexive : $x \leqslant x \quad (\forall x \in X)$,

(ii) transitive : $x \leqslant y,\ y \leqslant z \Rightarrow x \leqslant z \quad (\forall x, y, z \in X)$,

(iii) antisymmetric : $x \leqslant y,\ y \leqslant x \Rightarrow x = y \quad (\forall x, y \in X)$,

(iv) trichotomous : $x \leqslant y$ or $y \leqslant x \quad (\forall x, y \in X)$.

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

$$\forall S \subset X,\ S \neq \emptyset \ \Rightarrow\ \exists\, x \in S \ \text{ such that } \ y \geqslant x\ \forall y \in S.$$

Examples. 1. $\mathbb{N}$ (usual order).

2. $\mathbb{Z}$: *not* a well-ordering.

3. $\mathbb{Q}$: *not* a well-ordering.

4. $\mathbb{R}$: *not* a well-ordering.

5. $\{x \in \mathbb{Q} : x \geqslant 0\}$: *not* a well-ordering (consider $\{x \in \mathbb{Q} : x > 0\}$).

6. $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\}$.

7. $\{1 - \frac{1}{n} : n = 2, 3, \ldots\} \cup \{1\}$.

8. $\{1 - \frac{1}{n} : n = 2, 3, \ldots\} \cup \{2\}$.

9. $\{1 - \frac{1}{n} : n = 2, 3, \ldots\} \cup \{2 - \frac{1}{n} : n = 2, 3, \ldots\}$.

Remark. A total order X is a well-ordering $\Leftrightarrow$ X has no infinite strictly-decreasing sequence. Indeed, if we have such a sequence $x_1 > x_2 > x_3 > \ldots$, then $\{x_1, x_2, \ldots\}$ has no least element. Conversely, if $S \subset X$ has no least element, then for each $x \in S$ there is an $x' \in S$ with $x' < x$. But then $x > x' > x'' > x''' > \ldots$.

Say total orders x, y are **isomorphic** if there exists a bijection $f : X \to Y$ that is order-preserving ($x < y \Leftrightarrow f(x) < f(y)$).

For example, in the preceding list:

- examples 1 and 6 are isomorphic,
- examples 7 and 8 are isomorphic,
- examples 6 and 7 are *not* isomorphic (only one has a greatest element).

Proposition 1 (Proof by induction). Let X be well-ordered, and let $S \subset X$ such that for all $x \in X$, the following holds: if $y \in S$ for all $y < x$ then $x \in S$. Then $S = X$.

Equivalently: given some property $p(x)$, if $\big(\forall y < x : p(y)\big) \Rightarrow p(x)$ (each x), then $p(x)$ $\forall x$.

Proof. If $S \neq X$, let x be the least member of $X \setminus S$. Then $y \in S$ $\forall y < x$ (by choice of x), whence $x \in S$. ⨳ □

An example of the proof by induction:

Proposition 2. Let X, Y be isomorphic well-orderings. Then there is a *unique* isomorphism from X to Y.

Note. False for general total orders. For example: from $\mathbb{Z} \to \mathbb{Z}$, could take identity or $x \to x - 5$; from $\mathbb{R} \to \mathbb{R}$, could take identity or $x \to x^3$.

Proof. Let $f, g : X \to Y$ be isomorphisms. We shall show that $f(x) = g(x)$ $\forall x \in X$ by induction on X.

So, given $f(y) = g(y)$ $\forall y < x$, we need $f(x) = g(x)$. We must have that $f(x) = a$, the least element of $Y \setminus \{f(y) : y < x\}$ (which $\neq \emptyset$, as $f(x) \in$ it), because if $f(x) > a$ then some $x' > x$ has $f(x') = a$ (f bijective), contradicting f order-preserving.

Similarly, $g(x) = a$, so $f(x) = g(x)$ as required. □

A subset I of a total order X is an **initial segment** if $x \in I, y < x \Rightarrow y \in I$.

For example, for any $x \in X$, the set $I_x = \{y \in X : y < x\}$ is an initial segment. Not every proper initial segment of X need be of this form. For example: in $\mathbb{R}$, could take $\{x : x \leqslant 3\}$; in $\mathbb{Q}$, could take $\{x : x \leqslant 0 \text{ or } x^2 < 2\}$.

Remark. In a well-ordering X, every proper initial segment I *is* of the form I_x, some $x \in X$. Indeed, let x be the least member of $X \setminus I$. Then $y < x$ implies $y \in I$ (by choice of x). Also, $y \in I$ implies $y < x$ (if $y = x$ or $y > x$ we should get $x \in I$).

Aim. Every subset of X (well-ordered) is isomorphic to some initial segment of X.

Note. False for general total orders. For example: could take $\{1, 3, 4\}$ in $\mathbb{Z}$; or could take $\mathbb{Q}$ in $\mathbb{R}$.

For $f : A \to B$ and $C \subset A$, the **restriction** of f to C is $f|_C = \{(x, f(x)) : x \in C\}$.

Theorem 3 (Definition by recursion). Let X be a well-ordered set, and Y any set. Then for any $G : \mathbb{P}(X \times Y) \to Y$, there exists $f : X \to Y$ such that $f(x) = G(f|_{I_x})$ for all $x \in X$. Moreover, f is unique.

Note. In defining $f(x)$, we make use of f on $I_x = \{y : y < x\}$.

Proof. (Existence) Define 'h is an attempt' to mean the following ('the clever bit').

$$h : I \to Y, \text{ some initial segment } I \text{ of } X, \text{ and } x \in I \Rightarrow h(x) = G(h|_{I_x}).$$

Note that if h, h' are attempts, both defined at x, then $h(x) = h'(x)$, by induction: for if $h(y) = h'(y)\ \forall y < x$, then certainly $h(x) = h'(x)$.

Also, for each x, there exists an attempt defined at x, again by induction. Indeed, suppose that for all $y < x$, there is an attempt defined at y. Then there exists a unique attempt h_y defined on $\{z : z \leqslant y\}$. Put $h = \bigcup_{y<x} h_y$: an attempt defined on I_x.

Then $h \cup \{(x, G(h))\}$ is an attempt defined at x (single-valued, by uniqueness).

Now define f by: $f(x) = y$ if there is an attempt h with h defined at x and $h(x) = y$.

(Uniqueness) If f, f' satisfy the conditions then $f(x) = f'(x)\ \forall x$, by induction on x. (As if $f(y) = f'(y)$ for all $y < x$ then $f(x) = f'(x)$.) □

Proposition 4 (Subset collapse). Let X be a well-ordered set, and $Y \subset X$. Then Y is isomorphic to an initial segment of X. Moreover, this initial segment is unique.

Proof. To obtain $f : Y \to X$ that is order-preserving and having image an initial segment, we need $f(x) = \min X \setminus \{f(y) : y \in Y,\ y < x\}$ for all $x \in Y$.

Note that we cannot have $X \setminus \{f(y) : y < x, y \in Y\} = \emptyset$ – for example, because $f(y) \leqslant y$ for all y (by induction), so that $x \notin \{f(y) : y < x,\ y \in Y\}$.

So done (existence and uniqueness) by recursion (Theorem 3) applied to Y. □

In particular, X cannot be isomorphic to a proper initial segment of X – by uniqueness in subset collapse, since X is isomorphic to X.

How do different well-orderings relate to each other?

For well-orderings X, Y, write $X \leqslant Y$ if X is isomorphic to an initial segment of Y.

E.g. If $X = \mathbb{N},\ Y = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \} \cup \{2\}$, then $X \leqslant Y$.

Theorem 5. Let X, Y be well-orderings. Then $X \leqslant Y$ or $Y \leqslant X$.

Proof. Suppose $Y \not\leqslant X$. We need $X \leqslant Y$. For $f : X \to Y$ to be an isomorphism with an initial segment of Y, need: for all $x \in X,\ f(x) = \min Y \setminus \{f(y) : y < x\}$.

Note that we cannot have $Y \setminus \{f(y) : y < x\} = \emptyset$ as then Y is isomorphic to I_x. ⨳.

So done by recursion. □

Proposition 6. Let X, Y be well-orderings with $X \leqslant Y, Y \leqslant X$. Then X, Y are isomorphic. ('The best we could ever hope for.')

Proof. Let $f : X \to Y$ be an isomorphism from X to an initial segment of Y, and $g : Y \to X$ be an isomorphism from Y to an initial segment of X. Then $g \circ f$ is an isomorphism from X to an initial segment of X (as an initial segment of an initial segment is an initial segment), whence $g \circ f : X \to X$ is the identity on X (by uniqueness).

Similarly, $f \circ g$ is the identity of Y. So $f : X \to Y$ is a bijection. □

New well-orderings from old

Say $X < Y$ if $X \leqslant Y$ but X is not isomorphic to Y. Equivalently, $X < Y \Rightarrow X$ isomorphic to a **proper** initial segment of Y.

Find a bigger one. Given a well-ordering X, choose $x \notin X$, and define a well-ordering of $X \cup \{x\}$ by setting $y < x$ for all $y \in X$. This is the **successor** of X, written X^+. Clearly $X < X^+$.

Put some together. Given a set $\{X_i : i \in I\}$ of well-orderings, seek well-ordering X with $X \geqslant X_i$ for all i. Given well-orderings $(X, <_X)$ and $(Y, <_Y)$, say X *extends* Y if $Y \subset X$, and $<_X$ and $<_Y$ agree on Y, and Y is an initial segment of X. Say $\{X_i : i \in I\}$ are *nested* if for all i, j, either X_i extends X_j or X_j extends X_i.

Proposition 7. Let $\{X_i : i \in I\}$ be a nested set of well-orderings. Then there exists a well-ordering X with $X \geqslant X_i$ for all i.

Proof. Let $X = \bigcup_i X_i$. Put $x < y$ if for some i we have $x, y \in X_i$ and $x <_i y$ (where $<_i$ is the well-ordering on X_i). [Equivalently, $< = \bigcup_i <_i$.]

Clearly $<$ is a total order on X, with each X_i an initial segment (by nestedness). Also, given non-empty $S \subset X$, have $S \cap X_i \neq \emptyset$, some i. Then $S \cap X_i$ has a minimal member, x say (as X_i well-ordered). So x is minimal in S (as X_i an initial segment of X). Thus $<$ is a well-ordering on X, and $X \geqslant X_i$ for all i. □

Remark. Proposition 7 also holds when the X_i are not nested.

Ordinals

'Is the collection of all well-orderings itself a well-ordering?'

An **ordinal** is a well-ordered set, with two being regarded as the same if they are isomorphic.

(Just as the rationals consist of all symbols $\frac{m}{n}$ $(m, n \in \mathbb{Z}, n \neq 0)$ with two regarded as the same if $mn' = m'n$. But we cannot formalise ordinals using equivalence classes – see later.)

If X is a well-ordered set, corresponding to an ordinal α, say X has **order-type** α.

Examples. For $k \in \mathbb{N}$, write k for the order-type of the (unique) well-ordering of size k. Write ω for the order-type of $\mathbb{N}$. Then in $\mathbb{R}$, $\{1, 3, 4, 7\}$ has order-type 4, while $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\}$ has order-type ω.

Write $\alpha \leqslant \beta$ if there exists X of order-type α and Y of order-type β with $X \leqslant Y$. (Note: this does not depend on the choice of X, Y.) Similarly for $\alpha < \beta,\ \alpha^+$, etc.

Thus, for all α, β, we have $\alpha \leqslant \beta$ or $\beta \leqslant \alpha$, and $\alpha \leqslant \beta,\ \beta \leqslant \alpha\ \Rightarrow\ \alpha = \beta$.

Theorem 8. Let α be an ordinal. Then the ordinals $< \alpha$ form a well-ordered set, of order-type α.

Proof. Let X be a well-ordering of order-type α. Then the well-orderings $< X$ are precisely (up to isomorphism) the proper initial segments of X. But these are the $I_x,\ x \in X$, and hence biject with X in an order-preserving way ($x \leftrightarrow I_x$). □

Sometimes write I_α for $\{\beta : \beta < \alpha\}$. ('A nice set of order-type α.')

Proposition 9. Let S be a non-empty set of ordinals. Then S has a least element.

Proof. Choose $\alpha \in S$. If α is minimal, done. If α is not minimal, i.e. $S \cap I_\alpha \neq \emptyset$, take a minimal β in $S \cap I_\alpha$ (I_α is well-ordered, by Theorem 8). Then β is minimal in S. □

Theorem 10 (Burali-Forti paradox). The ordinals do not form a set.

Proof. Suppose not: let X be the set of all ordinals. Then X is a well-ordering, say of order-type α. But then X is isomorphic to I_α, a proper initial segment of X. ⨳ □

Given a set $S = \{\alpha_i : i \in I\}$ of ordinals, it has an upper bound α (i.e. $\alpha \geqslant \alpha_i$ for all i) by applying Proposition 7 to the *nested* family $\{I_{\alpha_i} : i \in I\}$. Hence, by Proposition 9, S has a *least* upper bound, written $\sup S$. E.g., $\sup\{2, 4, 6, 8, \ldots\} = \omega$.

On the following page are some ordinals.

Every ordinal in that picture is countable (as a countable union of countable sets is countable). Is there an uncountable ordinal? I.e., is there an uncountable well-ordered set?

We can well-order: $\mathbb{N}$, with the usual order;

$\mathbb{Q}$, by bijecting with $\mathbb{N}$;

$\mathbb{R}$..? Not obvious.

Amazingly, we can show:

Theorem 11. There is an uncountable ordinal.

Idea. Look at $\{\alpha \in ON : \alpha \text{ countable}\}$, where '$\alpha \in ON$' means '$\alpha$ is an ordinal'. (Is this a set? See Burali-Forti.)

Proof. Let $A = \{R \in \mathbb{P}(\mathbb{N} \times \mathbb{N}) : R \text{ is a well-ordering of a subset of } \mathbb{N}\}$, and let $B = \{\text{order-type } (R) : R \in A\}$. So B is precisely the set of all countable ordinals.

Let $\omega_1 = \sup B$. If ω_1 is countable, then it is the greatest countable ordinal (definition of B), contradicting $\omega_1 < \omega_1^+$. So ω_1 is uncountable. □

Some ordinals.

$$
\begin{array}{l}
0 \\
1 \\
2 \\
3 \\
\vdots \\
\omega \\
\omega+1 \quad \leftarrow \text{officially } \omega^{+} \\
\omega+2 \\
\omega+3 \\
\vdots \\
\omega+\omega \;=\; \omega 2 \qquad \leftarrow \text{officially } \sup\{\omega,\ \omega+1,\ \omega+2,\ \ldots\} \\
\omega 2+1 \\
\omega 2+2 \\
\omega 2+3 \\
\vdots \\
\omega 3 \\
\vdots \\
\omega 4 \\
\vdots \\
\omega 5 \\
\vdots \\
\omega\omega \;=\; \omega^2 \qquad \leftarrow \text{officially } \sup\{\omega,\ \omega 2,\ \omega 3,\ \ldots\} \\
\omega^2+1 \\
\omega^2+2 \\
\vdots \\
\omega^2+w \\
\omega^2+w+1 \qquad \text{officially } \sup\{\omega,\ \omega^w,\ \omega^{\omega^\omega},\ \ldots\} \rightarrow \\
\vdots \\
\omega^2+w2 \\
\vdots \\
\omega^2+\omega^2 \;=\; \omega^2 2 \\
\omega^2 2+1 \\
\vdots \\
\omega^2 2+w \\
\vdots \\
\omega^2 3 \\
\vdots \\
\omega^2 4 \\
\vdots \\
\omega^3 \\
\vdots \\
\omega^3 2 \\
\vdots \\
\omega^3 3 \\
\vdots \\
\omega^4 \\
\vdots \\
\omega^5 \\
\vdots \\
\omega^\omega \quad \leftarrow \text{officially } \sup\{\omega,\ \omega^2,\ \omega^3,\ \ldots\} \\
\vdots \\
\omega^\omega 2 \\
\vdots \\
\omega^\omega 3 \\
\vdots \\
\omega^\omega\omega \;=\; \omega^{\omega+1} \\
\vdots \\
\omega^{\omega+2} \\
\vdots \\
\omega^{\omega+3}
\end{array}
\qquad
\begin{array}{l}
\vdots \\
\omega^{\omega 2} \\
\vdots \\
\omega^{\omega 3} \\
\vdots \\
\omega^{\omega 4} \\
\vdots \\
\omega^{\omega^2} \;=\; \omega^{(\omega^2)} \\
\vdots \\
\omega^{\omega^2 2} \\
\vdots \\
\omega^{\omega^2 3} \\
\vdots \\
\omega^{\omega^3} \\
\vdots \\
\omega^{\omega^4} \\
\vdots \\
\omega^{\omega^\omega} \\
\vdots \\
\omega^{\omega^{\omega^2}} \\
\vdots \\
\omega^{\omega^{\omega^3}} \\
\vdots \\
\omega^{\omega^{\omega^\omega}} \\
\vdots \\
\omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}} \;=\; \epsilon_0 \\
\epsilon_0+1 \\
\epsilon_0+2 \\
\vdots \\
\epsilon_0+\omega \\
\vdots \\
\epsilon_0+\epsilon_0=\epsilon_0 2 \\
\vdots \\
\epsilon_0 3 \\
\vdots \\
\epsilon_0\omega \\
\vdots \\
\epsilon_0\epsilon_0=\epsilon_0^2 \\
\vdots \\
\epsilon_0^3 \\
\vdots \\
\epsilon_0^\omega \\
\vdots \\
\epsilon_0^{\omega^2} \\
\vdots \\
\epsilon_0^{\omega^3} \\
\vdots \\
\epsilon_0^{\omega^\omega} \\
\vdots \\
\epsilon_0^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}} \;=\; \epsilon_0^{\epsilon_0} \\
\vdots \\
\epsilon_0^{\epsilon_0} \\
\vdots \\
\epsilon_0^{\epsilon_0^{\epsilon_0}} \\
\vdots \\
\epsilon_0^{\epsilon_0^{\epsilon_0^{\cdot}}} \\
\epsilon_0^{\epsilon_0^{\epsilon_0^{\cdot^{\cdot}}}} \;=\; \epsilon_1 \\
\text{`` ... and so on ... ''}
\end{array}
$$

Note that by construction ω_1 is the *least* uncountable ordinal (definition of B). Note also that every ordinal in the picture is $< \omega_1$. Two surprising properties of the total order ω_1 are:

- it is an *uncountable* ordering in which for each x, the set $\{y : y < x\}$ is *countable*;
- every sequence $\alpha_1, \alpha_2, \ldots$ in ω_1 is bounded above, e.g. by $\sup\{\alpha_1, \alpha_2, \ldots\}$.

Theorem 11′ (Hartogs' Lemma). For any set X, there is an ordinal α that does not inject into X.

Proof. The same as in Theorem 11, with 'X' in place of '$\mathbb{N}$'. □

The least such α is denoted $\gamma(X)$. E.g., $\omega_1 = \gamma(\omega)$.

Successors and Limits

Let α be an ordinal. Does α have a greatest element? (That is, any set of order-type α, such as $I_\alpha = \{\beta : \beta < \alpha\}$.)

If yes: let β be the greatest element, then $\gamma < \alpha \Rightarrow \gamma < \beta$ or $\gamma = \beta$ (and converse trivially), so $\alpha = \beta^+$. Say α is a **successor**.

If no: for all $\beta < \alpha$, there is some $\gamma < \alpha$ such that $\gamma > \beta$, so $\alpha = \sup\{\beta : \beta < \alpha\}$. Say α is a **limit**.

E.g. 5 is a successor ($5 = 4^+$),

ω^+ is a successor,

ω is a limit (no greatest element of $\{\gamma : \gamma < \omega\}$),

(0 is a limit).

Ordinal Arithmetic

Define $\alpha + \beta$ (for ordinals α and β) by recursion on β (α fixed) as follows:

$$\begin{aligned} \alpha + 0 &= \alpha \\ \alpha + \beta^+ &= (\alpha + \beta)^+ \\ \alpha + \lambda &= \sup\{\alpha + \gamma : \gamma < \lambda\} \text{ for } \lambda \text{ a (non-zero) limit} \end{aligned}$$

E.g. $\omega + 1 = \omega + 0^+ = (\omega + 0)^+ = \omega^+$,

$\omega + 2 = (\omega + 1)^+ = \omega^{++}$,

$1 + \omega = \sup\{1 + \gamma : \gamma < \omega\} = \omega$

So $+$ is *not* commutative.

Remark. Officially, since the ordinals are not a set, 'recursion on the ordinals' would mean: to define $\alpha + \beta$ (given α, β), define $\alpha + \gamma$, all $\gamma \leqslant \beta$, by recursion on the set $\{\gamma : \gamma \leqslant \beta\}$ (plus uniqueness). Similarly for proof by induction on the ordinals: if $p(\alpha)$ false for some α, then p is not everywhere true on $\{\gamma : \gamma \leqslant \alpha\}$.

Notes. 1. $\beta \leqslant \gamma \Rightarrow \alpha + \beta \leqslant \alpha + \gamma$ (induction on γ)

2. $\beta < \gamma \Rightarrow \alpha + \beta < \alpha + \gamma$, because:

$$\beta < \gamma \Rightarrow \beta^+ \leqslant \gamma \Rightarrow \alpha + \beta^+ \leqslant \alpha + \gamma \Rightarrow (\alpha + \beta)^+ \leqslant \alpha + \gamma \Rightarrow \alpha + \beta < \alpha + \gamma.$$

3. But $1 < 2$ and yet $1 + \omega = 2 + \omega$ $(= \omega)$.

Proposition 12. For all $\alpha, \beta, \gamma \in ON$, $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.

Proof. By induction on γ (α, β fixed).

0. $\alpha + (\beta + 0) = \alpha + \beta = (\alpha + \beta) + 0$.

Successors. $(\alpha+\beta)+\gamma^+ = ((\alpha+\beta)+\gamma)^+ = (\alpha+(\beta+\gamma))^+ = \alpha+(\beta+\gamma)^+ = \alpha+(\beta+\gamma^+)$.

Limits. For λ a non-zero limit,

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

What about $\alpha + (\beta + \lambda)$?

Claim. $\beta + \lambda$ a limit.

Proof of claim. Have $\beta + \lambda = \sup\{\beta + \gamma : \gamma < \lambda\}$, but $\gamma < \lambda \Rightarrow \exists \gamma' < \lambda$ with $\gamma' > \gamma \Rightarrow \beta + \gamma' > \beta + \gamma$, so there is no greatest member of $\{\beta + \gamma : \gamma < \lambda\}$. So $\sup\{\beta + \gamma : \gamma < \lambda\}$ is a limit.

So $\alpha + (\beta + \lambda) = \sup\{\alpha + \delta : \delta < \beta + \lambda\}$. So our task is:

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

$\leqslant$ **:** For $\gamma < \lambda$, have $\beta + \gamma < \beta + \lambda$, so set on left $\subset$ set on right.

$\geqslant$ **:** For $\delta < \beta + \lambda$, have $\delta \leqslant \beta + \gamma$, some $\gamma < \lambda$ (definition of $\beta + \lambda$), whence $\alpha + \delta \leqslant \alpha + (\beta + \gamma)$. Thus each element of the RHS is $\leqslant$ some element of the LHS. $\square$

Another viewpoint

The definition of ordinal addition given above is called the 'inductive' definition. There is also a 'synthetic' definition: $\alpha + \beta$ is defined to be the order of $\alpha \sqcup \beta$ (disjoint union of α and β, e.g. $(\alpha \times \{0\}) \cup (\beta \times \{1\})$) with all of α coming before all of β.

E.g. $\omega + 1 = \overleftarrow{\omega} \bullet = \omega^+$,

$1 + \omega = \bullet \overleftarrow{\omega} = \omega$.

Armed with the synthetic definition, it is easy to see that $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$: each is the order-type of $\overleftarrow{\alpha}\ \overleftarrow{\beta}\ \overleftarrow{\gamma}$.

We must verify:

Proposition 13. The two definitions of ordinal addition coincide.

Proof. Write + for inductive, $+'$ for synthetic. We shall show $\alpha + \beta = \alpha +' \beta$ for all α, β, by induction on β.

0. $\alpha + 0 = \alpha = \alpha +' 0$

Successors. $\alpha +' \beta^+ =$ order-type of $\overset{\alpha}{\longleftrightarrow}\ \overset{\beta}{\longleftrightarrow}\ \bullet = (\alpha +' \beta)^+ = (\alpha + \beta)^+ = \alpha + \beta^+$

λ a non-zero limit. $\alpha + \lambda = \sup\{\alpha + \gamma : \gamma < \lambda\} = \sup\{\alpha +' \gamma : \gamma < \lambda\} = \alpha +' \lambda$. (*Nested* union, so sup = union.) □

Moral. Synthetic is easier to use, if it is available.

Ordinal Multiplication

Define $\alpha\beta$ by recursion on β as follows:

$$\begin{array}{lcl} \alpha 0 & = & 0 \\ \alpha(\beta^+) & = & \alpha\beta + \alpha \\ \alpha\lambda & = & \sup\{\alpha\gamma : \gamma < \lambda\} \text{ for } \lambda \text{ a (non-zero) limit} \end{array}$$

E.g. $\omega 2 = \omega 1 + \omega = (\omega 0 + \omega) + \omega = \omega + \omega$,

$2\omega = \sup\{2\gamma : \gamma < \omega\} = \omega$ (so multiplication is *not* commutative),

$\omega\omega = \sup\{\omega\gamma : \gamma < \omega\} = \sup\{0, \omega, \omega + \omega, \omega + \omega + \omega, \ldots\}$ (as in our big picture).

Or synthetically: $\alpha\beta$ is order-type of $\alpha \times \beta$, with $(x, y) < (z, t)$ if $y < t$ or $y = t, x < z$.

'β copies of α – go up in rows' $\qquad \beta \left\{ \begin{array}{c} \vdots \\ \overset{\alpha}{\longleftrightarrow} \\ \overset{\alpha}{\longleftrightarrow} \end{array} \right.$

E.g. $\omega 2 =$ order-type of $\left\{ \begin{array}{l} \overset{\omega}{\longleftrightarrow} \\ \overset{\omega}{\longleftrightarrow} \end{array} \right. = \omega + \omega$

$2\omega =$ order-type of $\left\{ \begin{array}{c} \vdots \\ \overset{\bullet\bullet}{\longleftrightarrow} \\ \overset{\bullet\bullet}{\longleftrightarrow} \end{array} \right. = \omega.$

Can check that definitions coincide, and $(\alpha\beta)\gamma = \alpha(\beta\gamma)$, etc.

Could also define exponentiation, towers, and so on. For example, ordinal exponentiation: define α^β by recursion on β:

$$\begin{array}{lcl} \alpha^0 & = & 1 \\ \alpha^{(\beta^+)} & = & \alpha^\beta \alpha \\ \alpha^\lambda & = & \sup\{\alpha^\gamma : \gamma < \lambda\} \text{ for } \lambda \text{ a (non-zero) limit} \end{array}$$

E.g. $\omega^2 = \omega^1\omega = (\omega^0\omega)\omega = (1\omega)\omega = \omega\omega$,

$2^\omega = \sup\{2^\gamma : \gamma < \omega\} = \omega$ – note, this is *countable*.

Chapter 3 : Posets and Zorn's Lemma

A **partially ordered set**, or **poset**, is a pair $(X, \leqslant)$, where $\leqslant$ is a relation on X that is:

(i) reflexive : $x \leqslant x \quad (\forall x \in X)$,

(ii) transitive : $x \leqslant y,\ y \leqslant z \Rightarrow x \leqslant z \quad (\forall x, y, z \in X)$,

(iii) antisymmetric : $x \leqslant y,\ y \leqslant x \Rightarrow x = y \quad (\forall x, y \in X)$.

Equivalently, writing $x < y$ to mean '$x \leqslant y$ and $x \neq y$', the conditions are:

(i) irreflexive : not $x < x \quad (\forall x \in X)$,

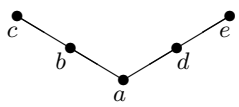
(ii) transitive : $x < y,\ y < z \Rightarrow x < z \quad (\forall x, y, z \in X)$.

Examples. 1. Any total order.

2. $(\mathbb{N}^+, \text{'divides'})$. (Not a total order: 3 and 5 are incomparable.)

3. For S any set, take $\mathbb{P}(S)$, with $A \leqslant B$ if $A \subset B$. (Very important.)

4. Take $X =$ any subset of $\mathbb{P}(S)$, same $\leqslant$.
E.g. V a vector space, $X =$ all subspaces of V.

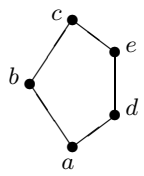
5. 

Meaning: $a \leqslant b$, $b \leqslant c$, $a \leqslant d$, $d \leqslant e$, and everything following by transitivity.
E.g. $a \leqslant c$, but b and d are not related.

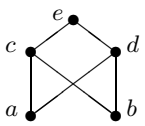
In general, a **Hasse diagram** of a poset consists of a drawing of the points of X with an upward line from x to y if y *covers* x (meaning $y > x$ and no z has $y > z > x$).

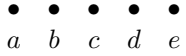
Hasse diagrams can be useful, e.g. $(\mathbb{N}, \leqslant)$, or useless, e.g. $(\mathbb{Q}, \leqslant)$ – no rational covers another!

Examples (continued).

6. 

(so no notion of 'height' or 'rank')

7. 

8. 

A subset S of a poset X is a **chain** if it is a total order ($\forall x, y \in S : x \leqslant y$ or $y \leqslant x$).

E.g. – In a total order, every subset is a chain.

– In example 2, $\{1, 2, 4, 8, 16, \ldots\}$ is a chain.

– In example 5, $\{a, b, c\}$ is a chain, as is $\{a, b\}$.

– Chains can be uncountable, e.g. $(\mathbb{R}, \leqslant)$.

Say S is an **antichain** if no two members of S are related ($\forall x, y \in S$, not $x < y$.)

E.g. – In example 2, $\{p : p \text{ prime}\}$ is an antichain.

– In example 5, $\{b, e\}$ is an antichain.

– In example 7, $\{e\}$ is an antichain, as is $\{a, b\}$.

– In example 8, the whole of X is an antichain.

For $S \subset X$, say $x \in X$ is an **upper bound** for S if $y \leqslant x$ for all $y \in S$, and 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 $y \geqslant x$.

E.g. – In $\mathbb{R}$, $S = \{x : x^2 < 2\}$ has 7 as an upper bound, and has a least upper bound, $\sqrt{2}$.

Write $\sup S = \sqrt{2}$, or $\bigvee S = \sqrt{2}$ – 'join of S'. ('$\bigvee$ is suggestive of union')

– In $\mathbb{Q}$, $S = \{x : x^2 < 2\}$ has 7 as an upper bound, but no supremum.

– In example 5, $\sup\{a, b, c\}$ is c (the only upper bound), and $\{b, d\}$ has *no* upper bound.

– In example 7, $\{a, b\}$ has upper bounds c, d, e, so *no* least upper bound.

Note. $\bigvee S$ may or may not belong to S. E.g. in $\mathbb{R}$, $\sup \in \{x : x \leqslant 1\}$, but $\sup \notin \{x : x < 1\}$.

A poset X is **complete** if every set $S \subset X$ has a supremum.

E.g. – $(\mathbb{R}, \leqslant)$ is not complete – e.g. $\mathbb{Z}$ has no upper bound. (Note: different from 'completeness' in 'metric space' sense.)

– $[0, 1]$ is complete, but $(0, 1)$ is not complete – e.g. $(0, 1)$ itself has no supremum.

– $\mathbb{Q}$ is not complete – e.g. $\{x : x^2 < 2\}$ or $\mathbb{Q}$ itself.

– $\mathbb{P}(S)$ is *always* complete: given sets A_i for $i \in I$, take $\bigcup_{i \in I} A_i$.

Note. In any complete poset X, there is a greatest element (an x with $x \geqslant y$ for all y), namely $\sup X$, and also a least element (an x with $x \leqslant y$ for all y), namely $\sup \emptyset$.

For a poset X, a function $f : X \to X$ is **order-preserving** if $x \leqslant y \Rightarrow f(x) \leqslant f(y)$.

E.g. – On $\mathbb{N}$, $f(n) = n + 1$.

– On $[0, 1]$, $f(x) = 1 - \frac{1}{2}(1 - x)$ ('halve the distance to 1').

– On $\mathbb{P}(S)$, $f(A) = A \cup \{j\}$, some fixed $j \in S$ – clearly $A \subset B \Rightarrow f(A) \subset f(B)$.

Say $x \in X$ is a **fixed point** of f if $f(x) = x$. Not every order-preserving f has a fixed point – e.g. $f(n) = n + 1$ on $\mathbb{N}$.

Theorem 1 (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 $E = \{x \in X : x \leqslant f(x)\}$, and let $s = \sup E$. We'll show that $f(s) = s$.

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

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

An application of Knaster-Tarski:

Corollary 2 (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. Want to write $A = P \sqcup Q$, $B = R \sqcup S$ such that $f|_P$ bijects P with R, and $g|_S$ bijects S with Q. (Then done: define $h : A \to B$ by taking $h = f$ on P and g^{-1} on Q.)

So we want $P \subset A$ such that $A \setminus g(B \setminus f(P)) = P$.

Let $X = \mathbb{P}(A)$, and define $\theta : X \to X,\ P \mapsto A \setminus g(B \setminus f(P))$. Need a fixed point of θ. But X is complete and θ is order-preserving: $P \subset P' \Rightarrow \theta(P) \subset \theta(P')$, so done by Knaster-Tarski. □

Zorn's Lemma

For X a poset, $x \in X$, say x is **maximal** if no $y \in X$ has $y > x$.

E.g. – In example 5 earlier, c and e are maximal.

– $\mathbb{N}, \mathbb{Q}, \mathbb{R}$ have *no* maximal elements – but then, they have chains without upper bounds.

Theorem 3 (Zorn's Lemma). 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 Hartogs' Lemma).

Pick some $x \in X$, and define $x_\alpha,\ \alpha < \gamma$ recursively by:

$$\begin{aligned} x_0 &= x \\ x_{\alpha^+} &= x'_\alpha \\ x_\lambda &= u(\{x_\alpha : \alpha < \lambda\}) \text{ for } \lambda \text{ a non-zero limit} \\ &\quad \text{(note } \{x_\alpha : \alpha < \lambda\} \text{ is a chain, by induction)} \end{aligned}$$

Then the $x_\alpha,\ \alpha < \gamma$, are distinct, so we have injected γ into X. ⧵⧵ □

Remarks. 1. We could define $x_\lambda = u(\{x_\alpha : \alpha < \lambda\})'$ to avoid thinking about injectivity.

2. Proof was easy given well-orderings, definition by recursion, etc, from chapter 2.

A typical application of Zorn is: does every vector space have a basis?

Recall. A **basis** is a linearly independent ('no finite linear combination = 0') spanning set ('everything is a finite sum from the set').

Examples. 1. V = set of all real polynomials. Then $\{1, x, x^2, x^3, \ldots\}$ is a basis.

2. V = set of all real sequences. Let $e_n = (0, 0, \ldots, 0, 1, 0, 0, \ldots)$. These are linearly independent, but they are not spanning: e.g., $(1, 1, 1, \ldots)$ is not in the span. So they do *not* form a basis. In fact, there is no countable basis (easy exercise). Even more: it is impossible to give an 'explicit' basis.

3. $\mathbb{R}$ as a vector space over $\mathbb{Q}$..? A basis here is called a **Hamel basis**.

Theorem 4. Every vector space V has a basis.

Proof. ('Go for a maximal linearly independent set.')

Let $X = \{A \subset V : A \text{ is linearly independent}\}$, ordered by $\subset$. Seek a maximal element $A \in X$. (Then done: if A does not span then choose x not in the span of A – then $A \cup \{x\}$ is linearly independent. ⨳)

Given chain $\{A_i : i \in I\}$, let $A = \bigcup_{i \in I} A_i$. Then $A \supset A_i$ for all i, so just need $A \in X$, i.e. A 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 $\lambda_1, \ldots, \lambda_n$ scalars, not all 0. Have $x_1 \in A_{i_1}, \ldots, x_n \in A_{i_n}$, some $i_1, \ldots, i_n \in I$. But some A_{i_k} has $A_{i_1}, \ldots, A_{i_n} \subset A_{i_k}$ (as the A_i are a chain), contradicting A_{i_k} linearly independent. ⨳

So, by Zorn, there is a maximal $A \in X$. □

Remarks. 1. The only actual 'maths' (i.e. linear algebra) we did was in the final check. This is very typical of Zorn.

2. '(non-empty)' is not strictly needed in statement of Zorn (as $\emptyset$ has an upper bound, so $X \neq \emptyset$). However, it's often safe to check $X \neq \emptyset$, as our chains should be non-empty.

Another application of Zorn: completeness theorem for propositional logic, without the assumption that the primitive propositions are countable.

Theorem 5. Let $S \subset L(P)$, any set P. Then S consistent $\Rightarrow$ S has a model.

Proof. Seek consistent $\overline{S} \supset S$ such that $\forall t \in L(P)$ have $t \in \overline{S}$ or $(\neg t) \in \overline{S}$.

(Then done, by setting $v(p) = 1$ if $p \in \overline{S}$ and $v(p) = 0$ if not – as in chapter 1.)

We seek a maximal consistent $\overline{S} \supset S$. (If $t \notin \overline{S}$ then $\overline{S} \cup \{t\} \vdash \bot$, so $\overline{S} \vdash (t \Rightarrow \bot)$ (deduction theorem), so $(t \Rightarrow \bot) \in \overline{S}$ by maximality of $\overline{S}$, i.e. $(\neg t) \in \overline{S}$.)

Let $X = \{T \subset L(P) : T \text{ consistent}, T \supset S\}$, ordered by $\subset$.

First, $X \neq \emptyset$, as $S \in X$. Given a non-empty chain $\{T_i : i \in I\}$ in X, let $T = \bigcup_{i \in I} T_i$. Then $T \supset T_i \ \forall i$, so just need $T \in X$.

Have $T \supset S$ (as chain non-empty). If T inconsistent, have $t_1, \ldots, t_n \in T$ with $\{t_1, \ldots, t_n\} \vdash \bot$ (proofs are finite). Then $t_1 \in T_{i_1}, \ldots, t_n \in T_{i_n}$, some $i_1, \ldots, i_n \in I$, whence $t_1, \ldots, t_n \in T_{i_k}$ for some T_{i_k} (as the T_i form a chain), contradicting T_{i_k} consistent. ⨳

So, by Zorn, X has a maximal element. □

One more application of Zorn.

Theorem 6 (Well-ordering Principle). Every set S can be well-ordered.

Remark. Very surprising for $S = \mathbb{R}$, for example.

Proof. Let $X = \{(A, R) : A \subset S,\ R \text{ is a well-ordering of } A\}$, ordered by extension. (That is, $(A, R) \geqslant (A', R')$ if $A' \subset A$, R and R' agree on A', and A' is an initial segment of A in the ordering R.)

First, $X \neq \emptyset$ since $\emptyset$ well-orderable, i.e. $(\emptyset, \emptyset) \in X$.

Given a chain $\{(A_i, R_i) : i \in I\}$, the (A_i, R_i) are a nested family, so $\left(\bigcup_{i \in I} A_i, \bigcup_{i \in I} R_i\right)$ is an upper bound (as in chapter 2).

So, by Zorn, X has a maximal element, say (A, R).

Claim. $A = S$

Proof of claim. Suppose $A \neq S$. Choose $x \in S \setminus A$, and take the successor of A : define a well-ordering on $A \cup \{x\}$ by setting $x > y \ \forall y \in A$.

This contradicts the maximality of A. ⨳ □

Zorn's Lemma and the Axiom of Choice

In our proof of Zorn's Lemma, we chose, for each x, an x' greater than it. This is making infinitely many arbitrary choices. We did the same in Part IA, when proving that a countable union of countable sets is countable: given sets $A_1, A_2, A_3, \ldots$, each listable, we chose a listing of each.

In terms of 'rules for building sets', we are appealing to the **axiom of choice**, which states: 'can choose an element from each of a family $\{A_i : i \in I\}$ of non-empty sets', or more precisely:

> every family $\{A_i : i \in I\}$ of non-empty sets has a *choice function*, meaning an $f : I \to \bigcup_{i \in I} A_i$ such that $f(i) \in A_i$ for all i.

This is of a different character from the other rules for building sets (e.g. 'given A and B, can form $A \cup B$', or 'given A, can form $\mathbb{P}(A)$'), in that the object whose existence is asserted is **not** uniquely specified by its properties (as opposed to, e.g., $A \cup B$).

So it is often of interest to know: does a given proof involve AC or not?

Remark. AC trivial if $|I| = 1$ ($A \neq \emptyset$ **means** that there is x such that $x \in A$)

Similarly, for $|I|$ finite, by induction on $|I|$.

However, for general I it turns out that AC **cannot** be deduced from the other set-building rules.

Does the proof of Zorn's Lemma **need** AC?

Yes, because we can deduce AC from Zorn's Lemma (using only the other set-building rules):

Given $\{A_i : i \in I\}$, each $A_i \neq \emptyset$, a **partial choice function** is a function $f : J \to \bigcup_{i \in I} A_i$, some $J \subset I$, such that $f(j) \in A_j$ for all $j \in J$.

Let $X = \{(J, f) : J \subset I, f \text{ a partial choice function } J \to \bigcup_{i \in I} A_i\}$, ordered by extension: $(J, f) \leqslant (J', f')$ if $J \subset J'$ and $f'|_J = f$.

First, $X \neq \emptyset$ since $(\emptyset, \emptyset) \in X$.

Given a chain $\{(J_q, f_q) : q \in Q\}$ has upper bound $(\bigcup_{q \in Q} J_q, \bigcup_{q \in Q} f_q)$. So, by Zorn, have maximal $(J, f) \in X$.

Want $J = I$. If $J \neq I$, choose $i \in I \setminus J$, choose $x \in A_i$, and consider $(J \cup \{i\}, f \cup \{(i, x)\})$ – contradicts maximality of (J, f). ⨳

Conclusion. ZL $\Leftrightarrow$ AC (given other set-building rules)

Actually, we had well-ordering principle implied by Zorn, and trivially WO $\Rightarrow$ AC (well-order $\bigcup_{i \in I} A_i$, and let $f(i) =$ least element of A_i).

Also, AC $\Rightarrow$ WO (without going via ZL, as in Theorem 6):

Let f be a choice function for $\{A \subset X : A \neq \emptyset\}$. Define x_α, $\alpha < \gamma(X)$, recursively by: having defined x_β, each $\beta < \alpha$, if $\{x_\beta : \beta < \alpha\} = X$ then stop, otherwise set $x_\alpha = f(X - \{x_\beta : \beta < \alpha\})$.

This must stop, else we have injected $\gamma(X)$ into X – contradiction. So we have an injection from X to some well-ordered set (an initial segment of $\gamma(X)$).

Conclusion. ZL $\Leftrightarrow$ AC $\Leftrightarrow$ WO (given the other set-building rules).

Remark. Zorn is hard to prove from first principles because we need ordinals, recursion, Hartogs, etc., and *not* because ZL $\Leftrightarrow$ AC.

**** Non-examinable section ****

Some notions related to completeness

1. Chain-completeness and Bourbaki-Witt

A poset X is **chain-complete** if $X \neq \emptyset$ and every non-empty chain has a least upper bound.

E.g. – Any complete poset.

– Any finite ($\neq \emptyset$) poset X, each chain has a greatest element.

– $X = \{A \subset V : A \text{ is linearly independent}\}$, any vector space V.

Say $f : X \to X$ is **inflationary** if $f(x) \geqslant x \ \forall x \in X$.

Bourbaki-Witt Theorem: every inflationary f on a chain-complete X has a fixed point.

Follows instantly from Zorn: X has a maximal element x, and $x \leqslant f(x)$ implies $x = f(x)$.

Can prove Bourbaki-Witt *without* AC: $x_0 \xrightarrow{f} x_1 \xrightarrow{f} \ldots \xrightarrow{f} x_\omega, \ldots$. (Note: we did not use AC in any ordinal theory, except a remark that ω_1 is not a countable sup.)

In fact, easy to get from B-W to ZL (using AC). So can view B-W as 'choice-free part of ZL'.

2. Lattices and Boolean Algebras

A **lattice** is a poset X in which every finite subset has a least upper bound and a greatest lower bound.

E.g. – $\mathbb{P}(S)$, any set S (as complete).

– $\{A \subset \mathbb{N} : A \text{ finite or } \mathbb{N} \setminus A \text{ finite}\}$.

For $a, b \in X$, write $a \vee b$ ('a join b'), for the least upper bound of $\{a, b\}$; and write $a \wedge b$ ('a meet b'), for the greatest lower bound of $\{a, b\}$

E.g. In $\mathbb{P}(S)$, $A \vee B = A \cup B$, $A \wedge B = A \cap B$.

A lattice X, say with greatest element 1 and least element 0, is a **Boolean algebra** if 'X behaves like $\mathbb{P}(S)$' :

(i) $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ $(\forall a, b, c, \in X)$

(ii) $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$ $(\forall a, b, c, \in X)$

(iii) $\forall a \in X, \ \exists\, a' \in X$ such that $a \vee a' = 1, \ a \wedge a' = 0$.

E.g. $\mathbb{P}(S)$, any S.

Fact. Every *finite* Boolean algebra is isomorphic to $\mathbb{P}(S)$, some finite S.

Not true in general, e.g. $\{A \subset \mathbb{N} : A \text{ finite or } \mathbb{N} \setminus A \text{ finite}\}$. (Countably infinite, so *not* $\mathbb{P}(S)$.)

Example: Lindenbaum algebra

Take propositional language L, with $\sim$ defined by: $p \sim q$ if $\vdash (p \Leftrightarrow q)$.

Have $\leqslant$ on $L/\sim$, defined by $[p] \leqslant [q] \Leftrightarrow \vdash (p \Rightarrow q)$.

Then $L/\sim$ is a Boolean algebra: $[p] \wedge [q] = [p \wedge q]$

$[p] \vee [q] = [p \vee q]$

$[p]' = [\neg p]$

**** End of non-examinable section ****

Chapter 4 : Predicate Logic

Overview of the set-up

Recall that a **group** is a set A, equipped with functions $m : A^2 \to A$ (of 'arity' 2), $i : A^1 \to A$ (arity 1), and a constant $e \in A$ (arity 0, i.e. $e : A^0 \to A$), satisfying:

$$\begin{aligned} (\forall x, y, z) \quad & \big(m(x, m(y, z)) = m(m(x, y), z)\big) \\ (\forall x) \quad & \big(m(x, e) = x \ \wedge \ m(e, x) = x\big) \\ (\forall x) \quad & \big(m(x, i(x)) = e \ \wedge \ m(i(x), x) = e\big) \end{aligned}$$

and a **poset** is a set A, equipped with a predicate $\leqslant \subset A^2$ (arity 2), satisfying:

$$\begin{aligned} (\forall x) \quad & (x \leqslant x) \qquad \text{(i.e., } (x, x) \in \leqslant) \\ (\forall x, y, z) \quad & ((x \leqslant y \ \wedge \ y \leqslant z) \Rightarrow x \leqslant z) \\ (\forall x, y) \quad & ((x \leqslant y \ \wedge \ y \leqslant x) \Rightarrow x = y) \end{aligned}$$

Propositional Logic		**Predicate Logic**
Language	$\longrightarrow$	E.g. language of groups (things like (1), (2), (3) above)
Valuation	$\longrightarrow$	Structure: a set equipped with functions, relations of right arities
A model for S (valuation in which each $s \in S$ holds)	$\longrightarrow$	A model for S (structure in which each $s \in S$ holds)
$S \models t$ (every model for S is a model for t)	$\longrightarrow$	Same (e.g. should have {group axioms} $\models m(e, e) = e$)
$S \vdash t$	$\longrightarrow$	Same (but a bit more complicated)

Let Ω (function symbols) and Π (relation symbols) be disjoint sets, and let α ('arity') be $\alpha : \Omega \cup \Pi \to \mathbb{N}$.

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

- **Variables.** Have variables $x_1, x_2, x_3, \ldots$. (We may use $x, y, \ldots$.)
- **Terms.** Defined inductively by:
  - (i) Every variable is a term.
  - (ii) If $f \in \Omega, \ \alpha(f) = n$ and $t_1, \ldots, t_n$ terms, then so is $f t_1 \ldots t_n$. (Can insert brackets and commas if desired.)

  E.g. In language of groups:

  $\Omega = \{m, i, e\}, \ \Pi = \emptyset, \ \alpha(m) = 2, \ \alpha(i) = 1, \ \alpha(e) = 0.$

  Some terms: $x_1, \ m(x_1, x_2), \ i(m(x_1, x_2)), \ e, \ m(x_1, e)$.

- **Atomic formulae.**

  (i) $\bot$ is an atomic formula.

  (ii) If s, t terms then $(s = t)$ is an atomic formula.

  (iii) If $\phi \in \Pi,\ \alpha(\phi) = n$, and $t_1, \ldots, t_n$ terms, then $\phi(t_1 t_2 \ldots t_n)$ is an atomic formula.

  E.g. In language of groups: $x_1 = x_2,\ m(x_1, x_1) = e$.

  In language of posets, (with $\Omega = \emptyset,\ \Pi = \{\leqslant\},\ \alpha(\leqslant) = 2$):
  $x_1 = x_1,\ x_1 \leqslant x_2$ (officially '$\leqslant (x_1, x_2)$').

- **Formulae.** Defined inductively by:

  (i) Every atomic formula is a formula.

  (ii) If p, q formulae, then so is $(p \Rightarrow q)$.

  (iii) If p a formula and x a variable, then $(\forall x)p$ is a formula.

  E.g. In language of groups:

  $$(\forall x)(m(x, e) = x),\ (\forall x)(i(x) = x),\ (\forall x)\big(m(x, x) = e \Rightarrow (\exists y)(m(y, y) = e)\big).$$

Notes. 1. A formula is a string of symbols.

2. Can define '$\neg p$' to mean '$p \Rightarrow \bot$', and similarly $p \vee q,\ p \wedge q$, etc, and '$(\exists\, x)p$' to mean '$\neg(\forall x)(\neg p)$'.

Closed terms. A term is **closed** if it contains no variables.

E.g. In language of groups: $e,\ m(e, e),\ m(e, i(e))$, but not $m(x_1, i(x_1))$.

Free and bound variables. An occurrence of a variable x in a formula p is called **bound** if it is inside the brackets of a '$\forall x$' quantifier. Otherwise, it is **free**.

E.g. $m(x, x) = e \Rightarrow (\exists\, y)\big(m(y, y) = x\big),\quad (\forall x)\big(m(x, x) = e\big)$

$\uparrow\ \uparrow \qquad\quad \nwarrow\ \nearrow\ \nearrow \qquad \uparrow \qquad\quad \nwarrow\ \nearrow\ \nearrow$

free bound free bound

$m(x, x) = e \Rightarrow (\forall x)(\forall y)\big(m(x, y) = m(y, x)\big) \quad \longleftarrow$ unhelpful

$\uparrow\ \uparrow \qquad\quad \uparrow \qquad\quad \uparrow \qquad\quad \uparrow$

free bound bound bound

Sentences. A **sentence** is a formula with no free variables.

E.g. $(\forall x)(m(x, x) = e)$.

Substitution. For p a formula, x a variable, t a term, write $p[t/x]$ for the formula obtained by substituting t for each free occurence of x.

E.g. If p is $(\exists\, y)(m(y, y) = x)$, then $p[e/x]$ is $(\exists\, y)(m(y, y) = e)$.

Semantic Entailment

Let $L = L(\Omega, \Pi, \alpha)$ be a language. An **L-structure** is a non-empty (see later for why) set A, together with

(i) for each $f \in \Omega$, a function $f_A : A^n \to A$, where $n = \alpha(f)$;

(ii) for each $\phi \in \Pi$, a set $A_\phi \subset A^n$, where $n = \alpha(\phi)$.

E.g. For L = language of groups, an L-structure is a set A, with functions $m_A : A^2 \to A$, $i_A : A \to A$, $e_A \in A$. (Note: need not be a group.)

For L = language of posets, an L-structure is a set A, with a relation $\leqslant_A \subset A^2$.

For L-structure A, sentence p, want to define 'p holds in A'.

For example, want '$(\forall x)(m(x,x) = e)$' to hold if and only if each $a \in A$ has $m_A(a,a) = e_A$. So: 'add in $\in A$ and subscript-A and read it aloud'. (Not a definition.)

Formal bit!

Define the **interpretation** of closed term t in L-structure to be $t_A \in A$ inductively by:

for $f \in \Omega,\ \alpha(f) = n,$ and $t_1, \ldots, t_n$ closed terms: $(ft_1 \ldots t_n)_A = f_A\big((t_1)_A \ldots (t_n)_A\big)$

Note. For c a constant, c_A already defined.

E.g. $m(e, m(e,e))_A = m_A(e_A, m_A(e_A, e_A))$.

Define the **interpretation** of sentence p in L-structure A to be $p_A \in \{0,1\}$ inductively by:

Atomic formulae

$$\bot_A = 0$$

$$(s = t)_A = \begin{cases} 1 & \text{if } s_A = t_A \\ 0 & \text{if not} \end{cases} \qquad \text{(any closed terms } s, t)$$

$$\phi(t_1 \ldots t_n)_A = \begin{cases} 1 & \text{if } (t_{1A} \ldots t_{nA}) \in \phi_A \\ 0 & \text{if not} \end{cases} \qquad \begin{array}{l}\text{(each } \phi \in \Pi,\ \alpha(\phi) = n, \\ \text{closed terms } t_1, \ldots, t_n)\end{array}$$

Sentences

$$(p \Rightarrow q)_A = \begin{cases} 0 & \text{if } p_A = 1,\ q_A = 0 \\ 1 & \text{if not} \end{cases} \qquad (p, q \text{ sentences})$$

$$((\forall x)p)_A = \begin{cases} 1 & \text{if } p[\overline{a}/x]_A = 1\ \forall a \in A \\ 0 & \text{if not} \end{cases}$$

(where, for any $a \in A$, we form a new language L' by adding a constant symbol $\overline{a}$, and make A into an L'-structure by setting $\overline{a}_A = a$).

End of formal bit!

Remark. If formula p has free variables, can define $p_A \subset A^{\#\text{free variables}}$.

E.g. If $p = (\exists y)\big(m(y,y) = x\big)$, then $p_A = \{a \in A : \exists b \in A \text{ with } m_A(b,b) = a\}$.

If $p_A = 1$, say p **holds** in A, or p is **true** in A, or A is a **model** of p.

If T is a **theory** (set of sentences), say A is a **model** of T if A is a model of p for all $p \in T$ (i.e. every $p \in T$ holds in A).

For T a theory, p a sentence, say T **entails** p, written $T \models p$, if every model of T is also a model of p.

Say p is a **tautology** if p holds in *all* L-structures, written $\models p$. (Equivalently, $\emptyset \models p$.)

E.g. $(\forall x)\ (x = x)$ is a tautology.

Example: theory of groups

L = language of groups: $\Omega = \{m, i, e\},\ \Pi = \emptyset$

arities: 2 1 0

$$\text{Let } T = \left\{ \begin{array}{c} (\forall x)(\forall y)(\forall z)\big(m(x, m(y, z)) = m(m(x, y), z)\big), \\ (\forall x)\big(m(x, e) = x \wedge m(e, x) = x\big), \\ (\forall x)\big(m(x, i(x)) = e \wedge m(i(x), x) = e\big) \end{array} \right\}.$$

Then an L-structure A is a model for $T \Leftrightarrow A$ is a group. (Two assertions.)

Say T **axiomatises** the class of groups, or 'axiomatises the theory of groups'. Sometimes the elements of T are called the 'axioms' of T.

Example: theory of fields

L = language of fields: $\Omega = \{+, \times, 0, 1, -\},\ \Pi = \emptyset$

arities: 2 2 0 0 1

Let T consist of: abelian group under $(+, -, 0)$
$\times$ commutative, and distributive over $+$
$(\forall x)(1x = x)$
$\neg(0 = 1)$
$(\forall x)\big((\neg(x = 0)) \Rightarrow (\exists\, y)(xy = 1)\big)$

Then an L-structure A is a model of $T \Leftrightarrow A$ is a field.

So T axiomatises the theory of fields.

E.g. $T \models$ 'inverses are unique' : $(\forall x)\big(x \neq 0) \Rightarrow (\forall y)(\forall z)([xy = 1 \wedge xz = 1] \Rightarrow y = z)\big)$.

Example: theory of posets

L = language of posets: $\Omega = \emptyset,\ \Pi = \{\leqslant\}$ (arity 2)

T : $(\forall x)\ (x \leqslant x)$
$(\forall x)(\forall y)(\forall z)\ \big((x \leqslant y \wedge y \leqslant z) \Rightarrow x \leqslant z\big)$
$(\forall x)(\forall y)\ \big((x \leqslant y \wedge y \leqslant x) \Rightarrow x = y\big)$

Example: theory of graphs

L = language of graphs: $\Omega = \emptyset,\ \Pi = \{a\}$ (a = 'is adjacent to', arity 2)

T : $(\forall x)\ (\neg a(x, x))$
$(\forall x, y)\ (a(x, y) \Rightarrow a(y, x))$

Proofs

Logical axioms. (3 usual, 2 for '=', 2 for '$\forall$')

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

Rules of deduction

Modus ponens: from $p, p \Rightarrow q$, can deduce q.

Generalisation: from p, can deduce $(\forall x)\ p$, provided x does not occur free in any premise used in the proof of p.

For $S \subset L,\ p \in L$, a **proof** of p in S is a finite sequence of formulae, ending with p, such that each one is a logical axiom or a hypothesis (member of S) or obtained from earlier lines by a rule of deduction.

Say S **proves** p, or p is a **theorem** of S, if there is a proof of p from S – written $S \vdash p$.

Note. Each logical axiom is a tautology.

Note on $\emptyset$. Suppose we are allowed $\emptyset$ as a structure (for a language L with no constants).

Then $(\forall x)\ \perp$ holds in $\emptyset$, and $\perp$ does not hold in S, so $((\forall x)\perp) \Rightarrow \perp$ does not hold in $\emptyset$

But this is an instance of axiom 6.

Example. $\{x = y,\ x = z\} \vdash y = z$ ('use axiom 5 with p being $x = z$')

1.	$(\forall x)(\forall y)\ ((x = y) \Rightarrow (x = z \Rightarrow y = z))$	(axiom 5)
2.	$\big((\forall x)(\forall y)\ ((x = y) \Rightarrow (x = z \Rightarrow y = z))\big)$ $\Rightarrow \big((\forall y)\ ((x = y) \Rightarrow (x = z \Rightarrow y = z))\big)$	(axiom 6)
3.	$(\forall y)\ ((x = y) \Rightarrow (x = z \Rightarrow y = z))$	(modus ponens on 1, 2)
4.	$\big((\forall y)\ ((x = y) \Rightarrow (x = z \Rightarrow y = z))\big)$ $\Rightarrow ((x = y) \Rightarrow (x = z \Rightarrow y = z))$	(axiom 6)
5.	$(x = y) \Rightarrow (x = z \Rightarrow y = z)$	(modus ponens on 3, 4)
6.	$x = y$	(hypothesis)
7.	$x = z \Rightarrow y = z$	(modus ponens on 5, 6)
8.	$x = z$	(hypothesis)
9.	$y = z$	(modus ponens on 7, 8)

Proposition 1 (Deduction Theorem). Let S be a set of formulae, and p, q formulae. Then $S \vdash (p \Rightarrow q)$ if and only if $S \cup \{p\} \vdash q$

Proof. ($\Rightarrow$) As for propositional logic: have a proof of $p \Rightarrow q$ from S. So add lines 'p' and 'q (modus ponens)' to obtain a proof of q from $S \cup \{p\}$.

($\Leftarrow$) Much as for propositional logic: only new case is generalisation. In other words, in a proof of q from $S \cup \{p\}$, we have written down

$$\frac{r}{(\forall x)\ r} \quad \text{(generalisation)}$$

and we have a proof of $S \vdash (p \Rightarrow r)$ (by induction), and we seek a proof of $S \vdash (p \Rightarrow (\forall x)\ r)$. In deduction of r from $S \cup \{p\}$, no hypothesis had x as a free variable, hence also in deduction of $p \Rightarrow r$ from S, no hypothesis had x as a free variable. Thus $S \vdash (\forall x)\ (p \Rightarrow r)$ by generalisation.

If x not free in p : from $S \vdash (\forall x)\ (p \Rightarrow r)$, get $S \vdash (p \Rightarrow (\forall x)\ r)$ by axiom 7 and modus ponens.

If x free in p : our deduction of r from $S \cup \{p\}$ cannot have used p, so actually $S \vdash r$, so $S \vdash (\forall x)\ r$ (generalisation), so $S \vdash (p \Rightarrow (\forall x)\ r)$, by axiom 1. □

Aim. $S \vdash p \Leftrightarrow S \models p$. For example, if a sentence holds in all groups then it may be deduced from the group theory axioms.

**** Non-examinable section ****

Proposition 2 (Soundness Theorem). Let S be a set of sentences, and p a sentence. Then $S \vdash p \Rightarrow S \models p$.

Proof. Have a proof of p from S, and need to show that every model for S is a model for p. This is an easy induction on the lines of the proof. □

For adequacy: want $S \models p \Rightarrow S \vdash p$, i.e. $S \cup (\neg p) \models \bot \Rightarrow S \cup (\neg p) \vdash \bot$.

Theorem 3 (Model existence lemma, or **completeness theorem).** Let S be a consistent set of sentences in a language L. Then S has a model.

Ideas.

1. Build model out of language itself: take a set of closed terms, with operations, e.g. $(1+1) +_A (1+1) = (1+1)+(1+1)$
2. E.g. for fields, $1+0=1$ for any model, but closed terms $1+0$ and 1 are distinct. So we would quotient out A by $s \sim t$ if $S \vdash (s=t)$ and use equivalence classes.
3. For 'fields of characteristic 2 or 3' : S = field axioms, with '$1{+}1=0 \vee 1{+}1{+}1=0$'. Then $S \nvdash (1+1=0)$, and $S \nvdash (1+1+1=0)$, so $[1+1] \neq [0]$ and $[1+1+1] \neq [0]$. So we do *not* get a field of characteristic 2 or 3.
   So need to extend S to a *maximal* consistent set first.
4. For 'fields with a square root of 2' : S = field axioms, with '$(\exists x)\ (xx = 1+1)$'. Then no closed term t has $[tt] = [1+1]$.
   S lacks 'witnesses' : so add constant symbol c, and add axiom '$cc = 1+1$' to S. But this has added to the language, hence it is no longer maximal consistent, so we must loop back to Idea 3. Problem: doesn't terminate!

Proof. Have consistent S in language $L = L(\Omega, \Pi)$. Extend S to maximal consistent S_1 (by Zorn). Then each sentence $p \in L$ has $p \in S_1$ or $(\neg p) \in S_1$, so S_1 is *complete* (i.e. $\forall p \in L : S_1 \vdash p$ or $S_1 \vdash \neg p$).

For each $(\exists\, x)\ p \in S_1$, add a new constant t to the language, and add $p[t/x]$ to S_1. We obtain T_1, in language $L_1 = L(\Omega \cup C_1, \Pi)$, that **has witnesses** for S_1: for each $(\exists\, x)\ p \in S$, have $p[t/x] \in T_1$, some closed term t. Easy to check T_1 consistent.

Extend T_1 to maximal consistent S_2 (in language L_1), then add witnesses for S_2 to obtain T_2 in language $L_2 = L(\Omega \cup C_1 \cup C_2, \Pi)$. Continue inductively.

Let $\overline{S} = S_1 \cup S_2 \cup \ldots$, in language $\overline{L} = L(\Omega \cup C_1 \cup C_2 \cup \ldots, \Pi)$.

Claim. $\overline{S}$ consistent, complete, and has witnesses (for itself).

Consistent. If $\overline{S} \vdash \bot$, then some $S_n \vdash \bot$ (as proofs are finite). ⨳

Complete. For sentence $p \in \overline{L}$, have $p \in L_n$ for some n (since p can only mention finitely many constants). So $S_{n+1} \vdash p$ or $S_{n+1} \vdash (\neg p)$ (as S_{n+1} complete in language L_n), so $\overline{S} \vdash p$ or $\overline{S} \vdash (\neg p)$.

Witnesses. Given $(\exists\, x)\ p \in \overline{S}$, have $(\exists\, x)\ p \in S_n$, some n. So there exists a closed term t, in language L_{n+1}, with $p[t/x] \in T_n$ (definition of T_n). Then $p[t/x] \in \overline{S}$ (and $t \in \overline{L}$).

On the set of closed terms of $\overline{L}$, define $s \sim t$ to mean $\overline{S} \vdash (s = t)$. Clearly, $\sim$ is an equivalence relation. Let A be the set of equivalence classes, made an $\overline{L}$-structure by:
- for $f \in \Omega$, let $f_A([t_1], \ldots, [t_n]) = [f(t_1, \ldots, t_n)]$.
- for $\phi \in \Pi$, let $\phi_A = \{([t_1], \ldots, [t_n]) : \overline{S} \vdash \phi(t_1, \ldots, t_n)\}$.

Claim. For any sentence $p \in \overline{L}$, have $\overline{S} \vdash p \Leftrightarrow p$ holds in A, i.e. $p_A = 1$.

Then done: A is a model of $\overline{S}$, hence of S.

Proof of claim. An easy induction.

Atomic sentences

$$\begin{array}{rcll}
\bot & : & \overline{S} \nvdash \bot, \text{ and } \bot_A = 0 & \\
s = t & : & \overline{S} \vdash (s = t) \Leftrightarrow [s] = [t] & \text{(definition of } \sim) \\
 & & \Leftrightarrow\ s_A = t_A & \text{(definition of operations on } A) \\
 & & \Leftrightarrow\ (s = t)_A = 1 & \text{(i.e. } s = t \text{ holds in } A) \\
\phi(t_1 \ldots t_n) & : & \text{same} &
\end{array}$$

Induction step

$$\begin{array}{rcl}
p \Rightarrow q\ :\ \overline{S} \vdash (p \Rightarrow q) & \Leftrightarrow & \overline{S} \vdash \neg p \text{ or } \overline{S} \vdash q \\
 & & (\Rightarrow\text{: if } \overline{S} \nvdash \neg p,\ \overline{S} \nvdash q, \text{ then } \overline{S} \vdash p,\ \overline{S} \vdash \neg q, \text{ by completeness of } \overline{S}.\text{⨳}) \\
 & \Leftrightarrow & p_A = 0 \text{ or } q_A = 1 \text{(induction hypothesis)} \\
 & \Leftrightarrow & (p \Rightarrow q)_A = 1 \\
 & & \\
(\exists\, x)\ p\ :\ \overline{S} \vdash (\exists\, x)\ p & \Leftrightarrow & \overline{S} \vdash p[t/x], \text{ some closed term } t\ (\Rightarrow\text{: } \overline{S} \text{ has witnesses}) \\
 & \Leftrightarrow & p[t/x]_A = 1, \text{ some closed term } t \quad \text{(induction)} \\
 & \Leftrightarrow & (\exists\, x)\ p \text{ holds in } A \text{ (as } A \textit{ is} \text{ the set of} \\
 & & \qquad \text{(equivalence classes of) closed terms)}
\end{array}$$

So, in particular, A is a model for S (as $S \subset \overline{S}$). □

Hence, by remarks before Theorem 3, we have:

Corollary 4 (Adequacy Theorem). Let S be a set of sentences, and p a sentence. Then $S \models p \Rightarrow S \vdash p$. □

Theorem 5 (Gödel's Completeness Theorem for First-Order Logic). Let S be a set of sentences, and p a sentence. Then $S \models p \Leftrightarrow S \vdash p$.

Proof. ($\Leftarrow$) Soundness.
($\Rightarrow$) Adequacy. □

Remarks. 1. If L countable (Ω, Π countable) then Zorn not needed.

2. 'First-order' means: our quantifiers/variables ranged over *elements* of the L-structure (*not* subsets).

**** End of non-examinable section ****

Corollary 6 (Compactness Theorem). Let S be a set of sentences. Then if every finite subset of S has a model then S has a model.

Proof. Trivial if we replace 'has a model' with 'is consistent' (as proofs are finite). □

Note. No decidability theorem – not so obvious to check $S \models p$.

Typical applications of completeness (compactness)

Can we axiomatise the theory of finite groups? In other words, is there a set T of sentences (in languages of group theory) such that a group G is a model of $T \Leftrightarrow G$ finite?

Corollary 7. The class of finite groups is *not* axiomatisable (in the language of groups).

Remark. Extraordinary that we can *prove* this, as opposed to just being convinced it is true!

Proof. Suppose T axiomatises finite groups. Form T' by adding to T the sentences

$$\begin{array}{rl} (\exists\, x_1)(\exists\, x_2)\ (x_1 \neq x_2) & \text{'}|G| \geqslant 2\text{'} \\ (\exists\, x_1)(\exists\, x_2)(\exists\, x_3)\ \big((x_1 \neq x_2) \wedge (x_1 \neq x_3) \wedge (x_2 \neq x_3)\big) & \text{'}|G| \geqslant 3\text{'} \end{array}$$

Etc. Then any finite subset of T' has a model ($\mathbb{Z}_m$, some m big enough), and so by compactness, T' itself has a model. ⨳ □

Note. We used compactness, which came from completeness – so we are using the full strength of the model existence lemma.

Corollary 7′ Let S be a theory with arbitrarily large finite models. Then S has an infinite model.

Proof. Add sentences as in Corollary 7, and apply compactness as in Corollary 7. □

'Finiteness is not a first-order property.'

Corollary 8 (Upward Löwenheim-Skolem Theorem). Let S be a theory with an infinite model. Then S has an uncountable model.

Proof. Add to the language an uncountable family $\{c_i : i \in I\}$ of constants, and let $S' = S \cup \{`c_i \neq c_j' : i, j \in I,\ i \neq j\}$. Seek a model for S'.

But any finite subset of S' has a model (as it can mention only finitely many of the c_i – so any infinite model of S will do). So by compactness, S' has a model. □

Remark. Similarly, we can ensure our model does not inject into X, for any fixed X – e.g. add $\gamma(X)$ constants, or $\mathbb{P}(X)$ constants.

For example, there exists an infinite field (e.g. $\mathbb{Q}$), so there is an uncountable field (e.g. $\mathbb{R}$), and so also a field that does not inject into $\mathbb{P}(\mathbb{P}(\mathbb{R}))$, say.

Corollary 9 (Downward Löwenheim-Skolem Theorem). Let S be a theory in a countable language (i.e. Ω, Π countable). If S has a model, then S has a countable model.

Proof. The model constucted in Theorem 3 is countable. □

Peano Arithmetic

We try to make the usual axioms for $\mathbb{N}$ into a first-order theory.

Language: $\Omega = \{0, s, +, \times\},\ \Pi = \emptyset$

	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
arities:	0	1	2	2

Axioms:

1. $(\forall x)(s(x) \neq 0)$
2. $(\forall x)(\forall y)(s(x) = s(y) \Rightarrow x = y)$
3. $(\forall y_1)\ldots(\forall y_n)\big((p[0/x] \wedge (\forall x)(p \Rightarrow p[s(x)/x])) \Rightarrow (\forall x)p\big)$
   (one such for each formula p, free variables $y_1, \ldots, y_n, x$ – an axiom-scheme, meaning an infinite set of axioms)
4. $(\forall x)(x + 0 = x)$
5. $(\forall x)(\forall y)(x + s(y) = s(x + y))$
6. $(\forall x)(x \times 0 = 0)$
7. $(\forall x)(\forall y)(x \times s(y) = (x \times y) + x)$

These axioms are called **Peano arithmetic** (PA), or sometimes *formal number theory.*

Note on axiom 3. For induction, our first guess would be

$$(p[0/x] \wedge (\forall x)(p \Rightarrow p[s(x)/x])) \Rightarrow (\forall x)p$$

But then we would be missing sets such as $\{x : x \geqslant y\}$, for a given y. Hence we add parameters as above.

Then PA has an infinite model ($\mathbb{N}$), and hence, by Upward Löwenheim-Skolem, an uncountable model. Thus PA has a model not isomorphic to $\mathbb{N}$ – doesn't this contradict the fact that $\mathbb{N}$ is uniquely defined by the usual axioms?

Answer. Axiom 3 is only 'first-order induction': it is *not* true induction (over *all* subsets of the structure). E.g., even in $\mathbb{N}$ itself, axiom 3 only refers to countably many subsets.

For $S \subset \mathbb{N}$, say S **definable** (or **definable in PA**) if there exists a formula p (in language of PA) with free variable x such that

$$\forall m \in \mathbb{N} : m \in S \iff (p[m/x] \text{ holds in } \mathbb{N}) \qquad (m = s(s(\ldots s(0)\ldots)))$$

So only *countably many* sets are definable.

E.g. – Set of squares. $p(x) : (\exists\, y)(yy = x)$.

– Set of primes. $p(x) : (\forall y)(y|x \Rightarrow [y = 1 \vee y = x])$, where $1 = s(0)$ and $y|x$ means $(\exists t)(yt = x)$.

– Powers of 2. $p(x) : (\forall y)([y \text{ prime} \wedge y \text{ divides } x] \Rightarrow y = 2)$.

Exercise. $\{x : x \text{ is a power of } 4\}$.

Challenge. $\{x : x \text{ is a power of } 10\}$.

Is PA a complete theory? I.e. for every sentence p, either PA $\vdash p$ or PA $\vdash \neg p$?

Gödel's Incompleteness Theorem says: PA is not complete (and a bit more).

So there exists p: PA $\nvdash p$, PA $\nvdash \neg p$. But one of p and $\neg p$ is true in $\mathbb{N}$.

Conclusion. There exists a sentence p which is true in $\mathbb{N}$, but PA $\nvdash p$.

This does not contradict the completeness theorem, which tells us that if p holds in *every* model of PA, then PA $\vdash p$.

Chapter 5 : Set Theory

Aim. What does the 'universe of sets' look like?

We shall view set theory as 'just' another first-order theory. ('A liberating viewpoint.')

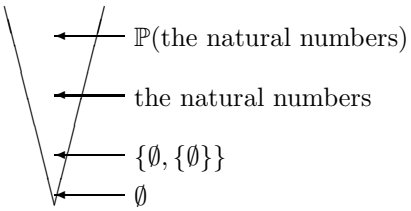
Zermelo-Fraenkel Set Theory

Language of ZF: $\Omega = \emptyset,\ \Pi = \{\in\}$ ($\in$ of arity 2).

Axioms of ZF: 2 to get started, 4 to build things, 3 one might not think of at first.

Then a 'universe of sets' will mean a model $(V, \in)$ of the ZF axioms.

Our question is: what does $(V, \in)$ look like?



('Does V really look like this?')

We could view chapter 5 as a worked example from chapter 4, but very scary, as (hopefully) *every* model of ZF should contain a copy of all of mathematics, and therefore will be incredibly complicated.

Axioms of ZF

1. **Axiom of extension.** 'If two sets have the same members, then they are equal.'

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

   Note. Converse is an instance of a logical axiom.

2. **Axiom of separation.** 'Can form a subset of a set.'

   More precisely, for any x and property p, can form $\{z \in x : p(z)\}$.

$$\underbrace{(\forall t_1)\ldots(\forall t_n)}_{\text{parameters}}(\forall x)(\exists\, y)\ \underbrace{(\forall z)\,[z \in y}_{\text{'the elements of } y \text{ are those } z \text{ such that}\ldots\text{'}} \Leftrightarrow (z \in x \wedge p)]$$

   – one for each formula p, free variables $t_1, \ldots, t_n, z$.

   Note. Need parameters $t_1, \ldots, t_n$, e.g. to allow $\{z \in x : t \in z\}$, for fixed t (parameter).

   Remark. Really, this is an axiom-scheme (an infinite set of axioms).

3. **Empty-set axiom.** 'There is a set with no members.'

$$(\exists\, x)(\forall y)(\neg y \in x)$$

We write $\emptyset$ for this (unique by extension) set. This is an **abbreviation** – so $p(\emptyset)$ is short for $(\exists\, x)((\forall y)(\neg y \in x) \wedge p(x))$.

Similarly, write $\{z \in x : p(z)\}$ for the set guaranteed by separation.

4. **Pair-set axiom.** 'Given x and y, 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 set, and write $\{x\}$ for $\{x, x\}$.

Can now define (as an abbreviation) the ordered pair (x, y) to be $\big\{\{x\}, \{x, y\}\big\}$.

Easy to check that $(x, y) = (z, t) \Leftrightarrow (x = z \wedge y = t)$. (Follows from the axioms so far.)

Say x is an **ordered pair** to mean $(\exists\, y)(\exists\, z)(x = (y, z))$.

Then f is a **function** means

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

and x is the **domain** of f, written $x = \operatorname{dom} f$, means

$$(f \text{ is a function}) \wedge (\forall y)\big(y \in x \Leftrightarrow (\exists\, z)[(y, z) \in f]\big),$$

and $f : x \to y$ means

$$(f \text{ is a function}) \wedge (x = \text{ dom } f) \wedge (\forall z)\big((\exists\, t)[(t, z) \in f] \Rightarrow z \in y\big).$$

5. **Union axiom.** 'Can form unions.'

$$(\forall x)(\exists\, y)(\forall z)[z \in y \Leftrightarrow (\exists\, t)(t \in x \wedge z \in t)]$$

E.g., $A \cup B \cup C = \bigcup\{\underbrace{A, B, C}_{x}\}$ $\longleftarrow$ $z \in$ this $\Longleftrightarrow z \in A$ or $z \in B$ or $z \in C$.

6. **Power-set axiom.** 'Can form power-sets.'

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

where we have written $x \subset y$ for $(\forall z)(z \in x \Rightarrow z \in y)$.

Notes. 1. Write $\bigcup x$ and $\mathbb{P}(x)$ for the sets given by the union and power set axioms. Often write $x \cup y$ for $\bigcup\{x, y\}$, etc.

2. No new axiom needed for intersections: can form $\bigcap x$ $(\neq \emptyset)$ as a subset of y, for any $y \in x$, so $\bigcap x$ obtainable from axiom of separation.

3. Can now form $x \times y$, as a subset of $\mathbb{P}(\mathbb{P}(x \cup y))$ – since if $t \in x$, $u \in y$ then $(t, u) = \big\{\{t\}, \{t, u\}\big\} \in \mathbb{P}(\mathbb{P}(x \cup y))$.

Similarly, can form the set of all functions from x to y, as a subset of $\mathbb{P}(x \times y)$.

7. **Axiom of infinity.**

So far, any model $(V, \in)$ must be infinite. For example, for any set x, let the **successor** of x be $x^+ = x \cup \{x\}$, and then $\emptyset,\ \emptyset^+,\ \emptyset^{++},\ \emptyset^{+++},\ \ldots$ are distinct.

$$\emptyset^+ = \{\emptyset\}, \quad \emptyset^{++} = \{\emptyset, \{\emptyset\}\}, \quad \emptyset^{+++} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \quad \ldots$$

Often write 0 for $\emptyset$, 1 for $\emptyset^+$, 2 for $\emptyset^{++}$, etc. Then

$$0 = \emptyset, \quad 1 = \{0\}, \quad 2 = \{0, 1\}, \quad 3 = \{0, 1, 2\}, \quad \ldots$$

Does V have an infinite set?

From outside (real world of maths): V is an infinite set.
From inside V: V is not a set – meaning $\neg(\exists\, x)(\forall y)(y \in x)$ (Russell's paradox)

We want V to have an infinite set: an $x \in V$ such that $\emptyset \in x,\ \emptyset^+ \in x,\ \emptyset^{++} \in x,\ \ldots$.

Say x is a **successor set** iff $(\emptyset \in x) \wedge (\forall y)(y \in x \Rightarrow y^+ \in x)$.
('A good idea – it is a legal (finite) formula.')

Axiom of infinity: 'There is an infinite set' / 'There is a successor set'

$$(\exists\, x)(x \text{ is a successor set})$$

Note that any intersection of successor sets is a successor set, so there is a *smallest* successor set (namely the intersection of *all* successor sets): call it ω. (This will be our version, in V, of the natural numbers.)

Thus $(\forall x)\big(x \in \omega \Leftrightarrow (\forall y)(y \text{ is a successor set} \Rightarrow x \in y)\big)$, so any subset of ω that is a successor set must equal ω (by definition of ω):

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

(Genuine induction, applying to *all* $x \subset \omega$ – not like PA in chapter 4.)

Can check $(\forall x)(x \in \omega \Rightarrow x^+ \neq \emptyset)$, and $(\forall x)(\forall y)\big([(x \in \omega) \wedge (y \in \omega) \wedge (x^+ = y^+)] \Rightarrow x = y\big)$, so ω satisfies (in V) our usual rules for the natural numbers.

Can now define x is **finite** to mean $(\exists\, y)\big((y \text{ bijects with } x) \wedge (y \in \omega)\big)$, and x is **countable** to mean (x is finite) $\vee$ (x bijects with ω).

8. **Axiom of foundation.** 'Sets are built up from simpler sets'.

Want to disallow things like $x \in x$ (i.e. '$\{x\}$ has no $\in$-minimal member'), and $x \in y \in x$ (i.e. '$\{x, y\}$ has no $\in$-minimal member'), etc.

Similarly, want to disallow $x_1 \in x_0,\ x_2 \in x_1,\ x_3 \in x_2, \ldots$ (i.e. '$\{x_0, x_1, x_2, \ldots\}$ has no $\in$-minimal member').

Axiom of foundation: 'Every (non-empty) set has an $\in$-minimal member.'

$$(\forall x)\big(x \neq \emptyset \Rightarrow (\exists\, y)[y \in x \wedge (\forall z)(z \in x \Rightarrow z \notin y)]\big)$$

9. **Axiom of replacement**

Often say: 'for each $i \in I$, have A_i: take $\{A_i : i \in I\}$'. But there is a problem: why should $i \mapsto A_i$ be a function? (I.e., why should $\{(i, A_i) : i \in I\}$ be a set?)

Want 'the image of a set, under something looking like a function, is a set'. This is going from I out into the universe.

Digression on classes

Idea. $x \mapsto \{x\}$ looks like a function, but it is not, as it is not even a set. (Every function f has a domain, defined for example as a suitable subset of $\bigcup\bigcup f$, and this f would have domain V ⨠.)

Let $(V, \in)$ be an L-structure. A **class** is a collection C of members of V such that for some formula p, free variables x (and maybe more), $x \in C \Leftrightarrow p$ holds.

E.g. V is a class – take p to be '$x = x$'.

For $t \in V$, $\{x \in V : t \in x\}$ is a class – take p to be '$t \in x$'. (t is a parameter.)

Clearly every set is a class – take p to be '$x \in y$'.

If a class C is not a set, i.e. $\neg(\exists y)(\forall x)\,(x \in y \Leftrightarrow p(x))$, say C is a **proper class**.

E.g. V is a proper class, as is $\{x \in V : x \text{ infinite}\}$.

Similarly, a **function-class** is a collection F of ordered pairs such that for some formula p, free variables x, y (and maybe more), we have: (x, y) belongs to $F \Leftrightarrow p$ holds, and if $(x, y) \in F,\ (x, z) \in F$, then $y = z$.

E.g. '$x \mapsto \{x\}$' is a function-class: take $p =$ '$y = \{x\}$'.

End of digression

Axiom of replacement: 'The image of a set under a function-class is a set.'

$$\underbrace{(\forall t_1)\ldots(\forall t_n)}_{\text{parameters}}\Big(\underbrace{\big[(\forall x)(\forall y)(\forall z)\big((p \wedge p[z/y]) \Rightarrow z = y\big)\big]}_{p \text{ is a function-class}}$$
$$\underbrace{\Rightarrow \big[(\forall x)(\exists\, y)(\forall z)\big(z \in y \Leftrightarrow (\exists\, t)(t \in x \wedge \underbrace{p[t/x, z/y]}_{\text{`}F(t)=z\text{'}}))\big]\Big)}_{\text{image of } x \text{ under } p \text{ is a set}},$$

– one for each formula p, free variables $x, y, t_1, \ldots, t_n$.

Note. This is an axiom-scheme.

For example, taking $F(x) = \{x\}$, we have that for any set x we can form $\{\{t\} : t \in x\}$, by replacement with p being '$y = \{x\}$'. This is a *bad* example, as we can deduce it from earlier axioms (like power-set) instead. See later for a *good* example.

The above axioms are called the **ZF axioms**.

Remarks. 1. Sometimes 'foundation' is called **regularity**. Sometimes 'separation' is called **comprehension**.

2. Axiom of choice is *not* included – can have ZF+AC, denoted ZFC, where:

Axiom of choice: 'Every family of non-empty sets has a choice function.'

$$(\forall f)\Big(\big[(f \text{ is a function}) \wedge (\forall x)(x \in \operatorname{dom} f \Rightarrow f(x) \neq \emptyset)\big] \Rightarrow$$
$$(\exists g)\big[(g \text{ a function}) \wedge (\operatorname{dom} g = \operatorname{dom} f) \wedge (\forall x)(x \in \operatorname{dom} g \Rightarrow g(x) \in f(x))\big]\Big)$$

Goal. 'What does V look like?'

Say x is **transitive** if every member of a member of x is a member of x:

$$(\forall y)\big([(\exists z)(y \in z \wedge z \in x)] \Rightarrow y \in x\big), \text{ or equivalently } \bigcup x \subset x.$$

E.g. ω is transitive, as $n = \{0, 1, 2, \ldots, n-1\}$ for each $n \in \omega$.

Lemma 1. Every set x is contained in a transitive set.

Remarks. 1. Officially: 'let $(V, \in)$ be a model of ZF. Then ... holds in V', or 'ZF $\vdash \ldots$'

2. Any intersection of transitive sets is transitive, so once we have proved lemma 1, we shall know that every x is contained in a *smallest* transitive set, the **transitive closure** of x, written $TC(x)$.

Proof. Want to form '$x \cup (\bigcup x) \cup (\bigcup\bigcup x) \cup (\bigcup\bigcup\bigcup x) \cup \ldots$'.

(Clearly transitive, and contains x.)

This will be by the union axiom, applied to the set $\{x, \bigcup x, \bigcup\bigcup x, \ldots\}$, which will be a set, by axiom of replacement ('a good example') applied to ω and function-class $0 \mapsto x,\ 1 \mapsto \bigcup x,\ 2 \mapsto \bigcup\bigcup x,\ \ldots$.

But why is this a function-class?

Define f is an **attempt** ('the clever bit') to mean

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

Then it is easy to check that

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

(by usual ω-induction) and also that

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

(also by ω-induction), so take function-class to be $p(y, z)$, where

$$p(y, z) = (\exists f)\ \big((f \text{ is an attempt}) \wedge (y \in \operatorname{dom} f) \wedge (f(y) = z)\big). \qquad \square$$

Want foundation to be telling us 'sets are built up from simpler sets'. If this is correct, we should want: if $p(x)$ holds whenever $(\forall y \in x)\ p(y)$, then $p(x)$ holds for all x.

Theorem 2 (Principle of $\in$-induction). For each formula p, free variables $t_1, \ldots, t_n, x$,

$$(\forall t_1)\ldots(\forall t_n)\big(\big[(\forall x)(\forall y)(y \in x \Rightarrow \underbrace{p(y)}_{\text{officially } p[y/x]}) \Rightarrow p(x)\big] \Rightarrow (\forall x)p(x)\big).$$

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

(Want to say: choose $\in$-minimal x with $\neg p(x)$, by foundation, and hence ⨳.
But $\{x : \neg p(x)\}$ need not be a set! E.g., p could have been '$x \neq x$'.)

Let $t = TC(\{x\})$, and $u = \{y \in t : \neg p(y)\}$. Then $u \neq \emptyset$ (as $x \in u$), so u has an $\in$-minimal member, say y. Then $\neg p(y)$ (as $y \in u$). But $z \in y \Rightarrow z \in t$ (as t transitive), so $z \notin u$ (as y minimal). I.e., $(\forall z \in y)\, p(z)$. ⨳ □

Remarks. 1. We are using the existence of transitive closures.

2. Foundation is actually *equivalent* to $\in$-induction, as we can deduce foundation from $\in$-induction (in presence of the other ZF axioms). Indeed, say '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, enough to show: if every $y \in x$ is regular then x is regular.

Proof. For z with $x \in z$, if x minimal in z, done. If x not minimal in z, have $y \in z$, some $y \in x$, so z has a minimal member (as y regular).

Note. Definition of 'regular' was a clever idea.

What about recursion? Want to be able to define $f(x)$ in terms of the $f(y),\ y \in x$.

Theorem 3 ($\in$-recursion theorem). Let G be a function-class $\big((x,y) \in G \Leftrightarrow p(x,y)$, some formula $p\big)$, everywhere defined.

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

Note. $F|_x = \{(t, F(t)) : t \in x\}$ *is* a set, by replacement.

Proof. (Existence) Say 'f is an attempt' if

$$(f \text{ is a function}) \wedge (\text{dom } f \text{ is transitive}) \wedge (\forall x)\big(x \in \text{dom}\, f \Rightarrow f(x) = G(\underbrace{f|_x}_{\substack{\text{makes sense, as} \\ \text{dom } f \text{ is transitive}}})\big).$$

Then $(\forall x)(\forall f)(\forall f')\big([(f, f' \text{ attempts}) \wedge (x \in \text{dom}\, f) \wedge (x \in \text{dom}\, f')] \Rightarrow f(x) = f'(x)\big)$, by $\in$-induction (as if f and f' agree at all $y \in x$ then they agree at x).

Also, $(\forall x)(\exists f)\big((f \text{ is an attempt}) \wedge (x \in \text{dom}\, f)\big)$, 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 $TC(\{y\})$. Put $f = \bigcup\{f_y : y \in x\}$ and then put $f' = f \cup \{(x, G(f|_x))\}$.

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

(Uniqueness) If we have suitable F, F' then $(\forall x)(F(x) = F'(x))$, by $\in$-induction. □

Note. Proofs of $\in$-induction and $\in$-recursion look similar to induction and recursion on a well-ordered set (from chapter 2).

Which properties of the 'relation' $\in$ (i.e. 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, $\{x : p(x, y)\}$ is a set.) $\leftarrow$ to build transitive closure.

So, for any $p(x, y)$ that is well-founded and local, can prove p-induction and p-recursion. If r is a relation on a *set* a, then trivially r is local, so to have r-induction and r-recursion, we just need r well-founded. (So theorems in chapter 2 were a special case of this.)

'Can we model a relation by $\in$?'

Example. On $\{a, b, c\}$ consider relation $r = \{(a, b), (b, c)\}$.

Take $a' = \emptyset$, $b' = \{\emptyset\}$, $c' = \big\{\{\emptyset\}\big\}$. Then $\{a', b', c'\}$ is transitive, and $x \ r \ y \Leftrightarrow x' \in y'$.

Say a relation r on a set a is **extensional** if

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

E.g. The relation above.

The analogue of 'subset collapse' is:

Theorem 4 (Mostowski's Collapsing 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.

Remark. 'Well-founded' and 'extensional' are trivially necessary.

Proof. (Existence) Define $f(x) = \{f(y) : y \ r \ x\}$ – a definition by r-recursion on the set a ('the only sensible choice'). Note that f is a function (not just a function-class), by replacement (it is an image of a). Let $b = \{f(x) : x \in a\}$, which is a set by replacement.

Then b transitive (definition of f), and f surjective (definition of b), and so just need to check that f is injective (then 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 $(\forall \ t \ r \ x)(\forall u \in a)(f(t) = f(u) \Rightarrow t = u)$.

Have $\{f(t) : t \ r \ x\} = \{f(u) : u \ r \ y\}$ (by definition of f), so $\{t : t \ r \ x\} = \{u : u \ r \ y\}$ (by induction hypothesis), so $x = y$ by extensionality.

(Uniqueness) If f, f' suitable then $f(x) = f'(x) \ \forall x \in a$ (by r-induction) $\square$

An **ordinal** is a transitive set that is totally ordered by $\in$. (Equivalently, 'well-ordered', thanks to foundation.)

E.g. $\emptyset, \{\emptyset\}$, any $n \in \omega$ (as $n = \{0, 1, \ldots, n - 1\}$), ω itself.

So by Mostowski, a well-ordering is order-isomorphic to a unique ordinal – that ordinal is the **order-type** of this well-ordering. (This was owed from chapter 2.)

So well-orderings x and y are order-isomorphic $\Leftrightarrow$ they have the same order type.

Amusing remark. For a well-ordering x, Mostowski sends each initial segment I_y to its order-type, so x is sent to $\{\text{order-type of } I_y : y \in x\}$. So, for each ordinal α, we have $\alpha = \{\beta : \beta < \alpha\}$. Thus $\alpha < \beta \Leftrightarrow \alpha \in \beta$.

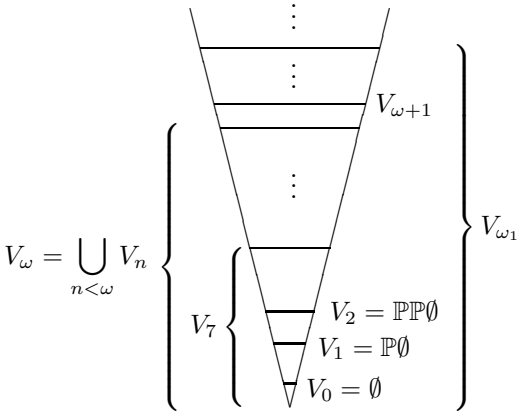
So, for example, $\alpha^+ = \alpha \cup \{\alpha\}$ and $\sup\{\alpha_i : i \in I\} = \bigcup\{\alpha_i : i \in I\}$ $\longleftarrow$ unhelpful.

Picture of the Universe

'Start with $\emptyset$; take $\mathbb{P}$ repeatedly (lots!).'

Define sets V_α, $\alpha \in ON$, by $\in$-recursion:

$V_0 = \emptyset$
$V_{\alpha^+} = \mathbb{P}V_\alpha$
$V_\lambda = \bigcup_{\gamma<\lambda} V_\gamma$, for λ a non-zero limit



Want: each $x \in V$ belongs to some V_α.

Lemma 5. Each V_α is transitive.

Proof. Use induction on α.

0. V_0 transitive.

Successors. V_α transitive $\Rightarrow$ $\mathbb{P}V_\alpha$ transitive, as if $x \in y \in \mathbb{P}V_\alpha$, then $y \subset V_\alpha$, so $x \in V_\alpha$, so $x \subset V_\alpha$ (as V_α transitive), so $x \in \mathbb{P}V_\alpha$

Limits. Any union of transitive sets is transitive. □

Lemma 6. $V_\alpha \subset V_\beta$ for all $\alpha \leqslant \beta$.

Proof. Induction on β (α fixed).

If $\beta = \alpha$. Done, as $V_\alpha \subset V_\alpha$.

Successors. Given $V_\alpha \subset V_\beta$, have $V_\beta \subset \mathbb{P}V_\beta$ (as $x \in V_\beta \Rightarrow x \subset V_\beta$, as V_β transitive), so $V_\alpha \subset \mathbb{P}V_\beta = V_{\beta^+}$.

Limits. Done. □

Theorem 7. Every x belongs to some V_α. '$V = \bigcup_{\alpha \in ON} V_\alpha$'.

Notes. 1. $x \subset V_\alpha \Leftrightarrow x \in V_{\alpha+1}$.

2. If $x \subset V_\alpha$ then there exists a least such α, called **rank** of x.

E.g. $\text{rank}(\emptyset) = 0$, $\text{rank}(\{\emptyset\}) = 1$, $\text{rank}(\omega) = \omega$. In general, $\text{rank}(\alpha) = \alpha$ for all $\alpha \in ON$.

Proof. We shall show $(\forall x)(\exists\, \alpha)(x \subset V_\alpha)$ by $\in$-induction.

So may assume that given x, each $y \in x$ has $y \subset V_{\text{rank}(y)}$, so $y \in V_{\text{rank}(y)^+}$.

Let $\alpha = \sup\{\text{rank}(y)^+ : y \in x\}$. Then $y \in V_\alpha$, all $y \in x$, so $x \subset V_\alpha$. □

Remarks. 1. The V_α are the **von Neumann hierarchy**.

2. Proof shows $\text{rank}(x) = \sup\{\text{rank}(y)^+ : y \in x\}$. ('Best way to work out ranks.')

Chapter 6 : Cardinals

Looking at 'sizes of sets'. Work in ZFC.

Say $x \leftrightarrow y$ if $(\exists\, f)$ (f a bijection from x to y).

Want to define 'card x', such that card x = card $y \Leftrightarrow (x \leftrightarrow y)$.

Cannot define card $x = \{y : y \leftrightarrow x\}$, as this might not be a set.

We know that $x \leftrightarrow \alpha$, some ordinal α, so could define card x to be the least such α.

[Just in ZF, use the 'Scott trick': let $\alpha(x)$ = least rank(y), over all $y \leftrightarrow x$ (sometimes called the 'essential rank' of x), and put card $x = \{y \subset V_{\alpha(x)} : y \leftrightarrow x\}$.]

Say m is a **cardinal** or **cardinality** if m = card x, some x.

The Alephs

What are the cardinalities of well-orderable sets / ordinals ?

Say α is **initial** if $(\forall \beta < \alpha)$ $(\neg \beta \leftrightarrow \alpha)$.

E.g. 0, 1, 2, ..., ω, ω_1, and indeed $\gamma(X)$ for any set X. But *not* ω^2 (countable, so $\omega^2 \leftrightarrow \omega$).

Define ω_α, $\alpha \in ON$, recursively by:

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

Then each ω_α is initial (by induction). Also, *every* infinite initial δ is an ω_α. Indeed, the ω_α are unbounded in the ordinals – e.g. we have $\omega_\alpha \geqslant \alpha$ $\forall \alpha$, by induction. Taking the least α with $\delta \leqslant \omega_\alpha$, must have $\delta = \omega_\alpha$, by definition of the ω_α.

Write $\aleph_\alpha$ ('aleph-α') for card ω_α.

So the alephs are the cardinalities of the infinite well-orderable sets.

E.g. card $\omega = \aleph_0$, card $\omega_1 = \aleph_1$.

For cardinals m, n, say $m \leqslant n$ if there is an injection $M \to N$, where M, N sets with card $M = m$, card $N = n$. (This does not depend on choice of M, N).

Similarly, $m < n$ means ($m \leqslant n$ and $m \neq n$).

E.g. card $\omega <$ card $\mathbb{P}(\omega)$.

Notes. If $m \leqslant n,\ n \leqslant m$, then $n = m$ (Schröder-Bernstein), so $\leqslant$ is a partial order.

> In ZFC, it is a total order. In fact, just in ZF it need not be a total order – the $\aleph_\alpha$ are the cardinalities of the well-orderable sets.

Cardinal Arithmetic

For cardinals m, n, define:

$$\begin{aligned} m+n &\quad \text{to be card}\,(M \sqcup N) \\ mn &\quad \text{to be card}\,(M \times N) \\ m^n &\quad \text{to be card}\,(M^N), \text{ where } M^N = \{f : f \text{ a function from } N \text{ to } M\} \end{aligned}$$

with M, N any sets with card $M = m$, card $N = n$. (Does not depend on choice of M, N.)

Similarly, $\sum_{i\in I} m_i = \text{card}\,(\sqcup_{i\in I} M_i)$. (Note: AC used here.) Etc.

E.g. 1. $\mathbb{R} \leftrightarrow \mathbb{P}\omega \leftrightarrow 2^\omega$, so card $\mathbb{R}$ = card $(\mathbb{P}\omega)$ = card $(2^\omega) = 2^{\aleph_0}$

2. How many sequences of reals are there?
Want card $(\mathbb{R}^\omega) = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0\aleph_0} = 2^{\aleph_0}$. (Using (D) and (A) below.)

Note. Used obvious facts like:

(A) $\aleph_0\aleph_0 = \aleph_0$ (as $\omega \times \omega \leftrightarrow \omega$)

(B) $m + n = n + m$ (as $M \sqcup N \leftrightarrow N \sqcup M$)

(C) $mn = nm$ (as $M \times N \leftrightarrow N \times M$)

(D) $(m^n)^p = m^{np}$ (as $(M^N)^P \leftrightarrow M^{N\times P}$)

Warning. Cardinal exponentiation is *not* the same as ordinal exponentiation.

E.g. Ordinals: ω^ω is countable (as $\omega^\omega = \sup\{\omega^\beta : \beta < \omega\}$)
Cardinals: $\aleph_0^{\aleph_0} \geqslant 2^{\aleph_0} > \aleph_0$, so uncountable.

We know $\aleph_0\aleph_0 = \aleph_0$. What about $\aleph_1\aleph_1$? All + and $\cdot$ are easy on the alephs, thanks to:

Theorem 1. For all $\alpha \in ON$, we have $\aleph_\alpha\aleph_\alpha = \aleph_\alpha$.

Proof. Induction on α.

Define a well-ordering of $\omega_\alpha \times \omega_\alpha$ by 'going up in squares':

$$\begin{aligned} (x,y) < (z,t) \text{ if either } & \max(x,y) < \max(z,t), \\ \text{or } & \max(x,y) = \max(z,t) = \beta, \text{ say, and } y < \beta,\ t = \beta, \\ & \qquad \text{or } y = t = \beta,\ x < z, \\ & \qquad \text{or } x = z = \beta,\ y < t. \end{aligned}$$

For any $\delta \in \omega_\alpha \times \omega_\alpha$, consider an initial segment I_δ. Then $I_\delta \subset \beta \times \beta$, some $\beta < \omega_\alpha$. But card $\beta <$ card ω_α (as ω_α initial), so $\beta \times \beta \leftrightarrow \beta$ (by the induction hypothesis) or β is finite, so card $I_\delta \leqslant$ card $(\beta \times \beta) =$ card $\beta <$ card ω_α.

Thus *every* proper initial segment has order-type $< \omega_\alpha$, whence our well-ordering has order-type $\leqslant \omega_\alpha$. Thus $\omega_\alpha \times \omega_\alpha$ injects into ω_α, so $\aleph_\alpha\aleph_\alpha \leqslant \aleph_\alpha$.

Trivially $\aleph_\alpha \leqslant \aleph_\alpha\aleph_\alpha$, so $\aleph_\alpha\aleph_\alpha = \aleph_\alpha$. □

Corollary 2. Let $\alpha \leqslant \beta$. Then $\aleph_\alpha + \aleph_\beta = \aleph_\alpha \aleph_\beta = \aleph_\beta$.

Proof. $\aleph_\beta \leqslant \aleph_\alpha + \aleph_\beta \leqslant 2\aleph_\beta \leqslant \aleph_\alpha \aleph_\beta \leqslant \aleph_\beta \aleph_\beta = \aleph_\beta$. □

Example. In ZFC, an infinite set X has $X \leftrightarrow X \sqcup X$.

Remarks. However, exponentiation is *hard*. For example $2^{\aleph_0}$ might not even be an aleph (if no AC).

Even in ZFC, cannot prove or disprove $2^{\aleph_0} = \aleph_1$. (The **continuum hypothesis**.)

Even today, not all implications about values of cardinal exponentiation ($\aleph_\alpha^{\aleph_\beta}$) are known.

** Bonus lecture : Incompleteness **

(PTJ chapters 4 and 9 for everything.)

Peano Arithmetic (PA): axioms (7 of them), in language 0, s, $+$, $\cdot$.

Aim. PA incomplete: i.e., there exists a sentence p such that PA $\nvdash p$ and PA $\nvdash \neg p$.

Equivalently: there exists a sentence p, *true in* $\mathbb{N}$, such that PA $\nvdash p$.

Here, 'true' = 'true in $\mathbb{N}$', 'provable' = 'PA proves it'. So we want p, true, but not provable.

Idea. Find p saying 'I am not provable', or, more precisely, p with p true $\Leftrightarrow$ p not provable.

[Then done: if p false then PA $\vdash p$, whence p holds in every model of PA, and in particular p holds in $\mathbb{N}$ ⩛. So p true, hence p not provable.]

We shall 'code up' formulae, proofs, etc. inside PA (i.e. as natural numbers). But it still looks as if, in any format, 'p not provable' *has* to be longer than p.

Recall. $S \subset \mathbb{N}$ is **definable** or **definable in the language of PA** if there exists a formula $p(x)$ (in the language of PA, x a free variable) such that for all $m \in \mathbb{N}$, $m \in S \Leftrightarrow p(m)$ holds in $\mathbb{N}$ (where $p(m)$ means $p[\underbrace{ss\ldots s}_{m}0/x]$).

E.g. Set of primes: $p(x) : (\forall y)(\forall z)([yz = x] \Rightarrow [y = 1 \vee z = 1]) \wedge (x \neq 1)$.
Say 'm is prime' is definable.

Similarly, 'm is a power of 2' definable: $p(x) : (\forall y)([\underbrace{y|x}_{(\exists z)(yz=x)} \wedge y \text{ prime}] \Rightarrow y = 2)$.

Similarly, a function $f : \mathbb{N} \to \mathbb{N}$ is **definable** if there is a formula $p(x, y)$ such that $\forall\, m, n \in \mathbb{N}$ $f(m) = n \Leftrightarrow p(m, n)$ true.

Fact. Any function $f : \mathbb{N} \to \mathbb{N}$ given by some algorithm is definable.

E.g. $f(n) = 2^n$ definable, meaning: there is a formula $p(x, y)$ such that for all $m, n \in \mathbb{N}$, $m = 2^n \Leftrightarrow p[n/x, m/y]$ holds.

Coding

Language symbols are: $0\ \ s\ \ +\ \ \cdot\ \ \perp\ \ \Rightarrow\ \ (\ \ )\ \ \forall\ x\ \ '\ \ =$

Assign to each a distinct value: $v(0) = 1,\ v(s) = 2,\ v(+) = 3,\ \ldots,\ v(') = 11,\ v(=) = 12$.

Now code a formula p by raising successive primes to the powers of successive symbols.

E.g. If p is '$(\forall x)(x = 0)$', have code $c(p) = 2^7\ 3^9\ 5^{10}\ 7^8\ 11^7\ 13^{10}\ 17^{12}\ 19^1\ 23^8$.

Not every number codes a formula, e.g. $2^7\ 3^7$ or $2^9\ 5^7$.

For any $m \in \mathbb{N}$, 'm codes a formula' is definable (as there exists an algorithm).

Write S_m for the formula coded by m (and set $S_m =$ '$\perp$' if m does not code a formula).

Now code a finite sequence $p_1, \ldots, p_n$ of formulae by:

$$S(p_1, \ldots, p_n) = 2^{c(p_1)} 3^{c(p_2)} \ldots (n^{\text{th}} \text{ prime})^{c(p_n)}.$$

Observe that 'm codes an axiom' is definable (where 'axiom' means logical or PA), as there is an algorithm. (Easy check.)

Similarly, 'l, m, n code formulae, with S_n obtained from S_l, S_m by modus ponens' is definable, and same for generalisation.

So $\theta(m, n) =$ 'n codes a proof of S_m' is definable.

So $\phi(m) =$ 'm codes a provable statement' (i.e. 'S_m is provable') is definable, as $\phi(m) \Leftrightarrow (\exists\, n)\, \theta(m, n)$.

Clever bit:

Consider $\chi(m) =$ 'm codes a formula, with one free variable, and $S_m(m)$ is a non-provable statement'. Clearly definable, say by formula $p(x)$. (I.e. $p(m)$ holds in $\mathbb{N} \Leftrightarrow \chi(m)$ true.)

Let the code for p be N (i.e. $p(x) = S_N$). So $\chi(N)$ asserts: 'N codes a formula, with one free variable, and $S_N(N)$ unprovable'. (I.e. '$\chi(N)$ not provable'.)

Thus the sentence $p(N)$ will do!

We have shown:

Theorem 1 (Gödel's Incompleteness Theorem.) PA is incomplete. □

Could we make PA complete by adding some clever sentence p (true in $\mathbb{N}$) to it?
Answer: no. If $\text{PA}' = \text{PA} \cup \{p\}$, run the same proof.

However, we can certainly enlarge PA to a complete theory: e.g. set $T = \{p : p \text{ true in } \mathbb{N}\}$.

Why does the above proof not still work? (I.e. with PA replaced by T throughout.) It can only be because:

Theorem 2. T is not definable. □

So 'm codes a *true* statement' not definable. 'Truth is not definable.'

Why does the proof of Theorem 1 not formalise, in PA, into a proof of our true-but-unprovable p from PA?

Answer. We used existence of a model of PA (namely $\mathbb{N}$), i.e. we used con(PA) = 'PA is consistent' = '$(\forall n)(n$ does not code a proof of $\perp)$'.

Formalising Theorem 1 gives: $\text{PA} \cup \text{con(PA)} \vdash p$.

Hence:

Theorem 3. PA $\nvdash$ con(PA). □

Does ZF $\vdash$ con(PA)? (So, $(\forall x \in \omega)(x$ does not code a proof of $\bot)$.)

Yes: as ZF $\vdash$ 'PA has a model' (namely ω).

However, running Theorem 1 of ZF language, we get:

Theorem 4. If ZF consistent, then ZF incomplete. (Get p as before, if ZF consistent.) □

And as for Theorem 1 $\Rightarrow$ Theorem 3, get:

Theorem 5. If ZF consistent, then ZF $\nvdash$ con(ZF). □

Lent 2010 **LOGIC AND SET THEORY – EXAMPLES 1** IBL

1. Which of the following propositions are tautologies?
(i) $(p_1 \Rightarrow (p_2 \Rightarrow p_3)) \Rightarrow (p_2 \Rightarrow (p_1 \Rightarrow p_3))$
(ii) $((p_1 \vee p_2) \wedge (p_1 \vee p_3)) \Rightarrow (p_2 \vee p_3)$
(iii) $(p_1 \Rightarrow (\neg p_2)) \Rightarrow (p_2 \Rightarrow (\neg p_1))$

2. Write down a proof of $\bot \Rightarrow q$. Use this to write down a proof of $p \Rightarrow q$ from $\neg p$.

3. Use the Deduction Theorem to show that $p \vdash \neg\neg p$.

4. Show that $\{p, q\} \vdash p \wedge q$ in three different ways: by writing down a proof, by using the Deduction Theorem, and by using the Completeness Theorem.

5. Give propositions p and q for which $(p \Rightarrow q) \Rightarrow \neg(q \Rightarrow p)$ is a tautology.

6. Explain carefully why the set of all propositions is countable.

7. Three people each have a set of beliefs: a consistent deductively closed set. Show that the set of propositions that they all believe is also consistent and deductively closed. Must the set of propositions that a majority believe be consistent? Must it be deductively closed?

8. Can the third axiom be deduced from the first two? In other words, is there a proof of $(\neg\neg p) \Rightarrow p$ that uses only the first two axioms and modus ponens?

9. Let $t_1, t_2, \ldots$ be propositions such that, for every valuation v, there exists n with $v(t_n) = 1$. Use the Compactness Theorem to show that in fact we may bound the values of n: there must be an N such that, for every valuation v, there exists $n \leq N$ with $v(t_n) = 1$.

10. Two sets S, T of propositions are *equivalent* if $S \vdash t$ for every $t \in T$ and $T \vdash s$ for every $s \in S$. A set S of propositions is *independent* if for every $s \in S$ we have $S - \{s\} \nvdash s$. Show that every finite set of propositions has an independent subset equivalent to it. Give an infinite set of propositions that has no independent subset equivalent to it. Show, however, that for every set of propositions there exists an independent set equivalent to it.

11. Give a direct proof of the Compactness Theorem (not making use of the notion of syntactic implication).

12. Give an explicit function f from natural numbers to natural numbers such that every tautology of length n has a proof that is at most $f(n)$ lines long.

13. A set S of propositions is a *chain* if for any distinct $p, q \in S$ we have $p \vdash q$ or $q \vdash p$ but not both. Write down an infinite chain. If the set of primitive propositions is allowed to be uncountable, can there exist an uncountable chain?

$^{+}$14. Suppose that the set of primitive propositions is allowed to be uncountable. Is it true that for every set of propositions there exists an independent set equivalent to it?

Lent 2010 **LOGIC AND SET THEORY – EXAMPLES 2** IBL

1. Write down subsets of the reals that have order-types $\omega + \omega$, ω^2 and ω^3.

2. Let α and β be non-zero ordinals. Must we have $\alpha + \beta > \alpha$? Must we have $\alpha + \beta > \beta$?

3. Is there a non-zero ordinal α with $\alpha\omega = \alpha$? What about $\omega\alpha = \alpha$?

4. Show that the inductive and the synthetic definitions of ordinal multiplication coincide.

5. Let α, β, γ be ordinals. Prove that $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.

6. Let α, β, γ be ordinals. Must we have $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$? Must we have $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$?

7. Let α and β be ordinals with $\alpha \geq \beta$. Show that there is a unique ordinal γ such that $\beta + \gamma = \alpha$. Must there exist an ordinal γ with $\gamma + \beta = \alpha$?

8. An ordinal written as $\omega^{\alpha_1} n_1 + \ldots + \omega^{\alpha_k} n_k$, where $\alpha_1 > \ldots > \alpha_k$ are ordinals (and k and $n_1, \ldots, n_k$ are non-zero natural numbers), is said to be in *Cantor Normal Form*. Show that every non-zero ordinal has a unique Cantor Normal Form. What is the Cantor Normal Form for the ordinal ϵ_0?

9. Is ω_1 a successor or a limit?

10. Let α be a countable (non-zero) limit ordinal. Prove that there exists an increasing sequence $\alpha_1 < \alpha_2 < \alpha_3 < \ldots$ with supremum equal to α. Is this result true for $\alpha = \omega_1$?

11. Show that, for every countable ordinal α, there is a subset of $\mathbb{Q}$ of order-type α. Why is there no subset of $\mathbb{R}$ of order-type ω_1?

12. Let X be a totally ordered set such that every subset of X is isomorphic to some initial segment of X. Prove that the ordering must be a well-ordering.

13. Is it possible to select for each countable (non-zero) limit ordinal α an ordinal $x_\alpha < \alpha$ in such a way that the x_α are distinct?

$^+$14. Let X be a totally ordered set such that the only order-preserving injection from X to itself is the identity. Must X be finite?

1. How many different partial orders (up to isomorphism) are there on a set of 4 elements? How many of these are complete?

2. Which of the following posets (ordered by inclusion) are complete?
(i) The set of all subsets of $\mathbb{N}$ that are finite or have finite complement
(ii) The set of all independent subsets of a vector space V
(iii) The set of all subspaces of a vector space V

3. Let X be a complete poset, and let $f : X \to X$ be order-reversing (meaning that $x \leq y$ implies $f(x) \geq f(y)$). Give an example to show that f need not have a fixed point. Show, however, that there must exist either a fixed point of f or two distinct points x and y with $f(x) = y$ and $f(y) = x$.

4. Use Zorn's Lemma to show that every partial order on a set may be extended to a total order.

5. Give a direct proof of Zorn's Lemma (not using ordinals and not using the Axiom of Choice) for countable posets.

6. Show that the statement 'for any sets X and Y, either X injects into Y or Y injects into X' is equivalent to the Axiom of Choice (in the presence of the other rules for building sets). [Hint for one direction: Hartogs' Lemma.]

7. What is yellow and equivalent to the Axiom of Choice?

8. Formulate sets of axioms in suitable languages (to be specified) for the following theories.
(i) The theory of fields of characteristic 2
(ii) The theory of posets having no maximal element
(iii) The theory of bipartite graphs
(iv) The theory of algebraically closed fields
(v) The theory of groups of order 60
(vi) The theory of simple groups of order 60
(vii) The theory of real vector spaces

9. Write down axioms (in the language of groups) for the theory of groups that are either infinite or have order a multiple of 100.

10. Show that the theory of fields of positive characteristic is not axiomatisable (in the language of fields), and that the theory of fields of characteristic zero is axiomatisable but not finitely axiomatisable.

11. Is every countable model of Peano Arithmetic isomorphic to $\mathbb{N}$?

12. Write down axioms, in a suitable language, for the theory of groups that have an element of infinite order. Can this be done in the language of groups?

13. Let L be the language consisting of a single function symbol f, of arity 1. Write down a theory T that asserts that f is a bijection with no finite orbits, and describe the countable models of T. Prove that T is a complete theory.

14. Show that the following theories are not axiomatisable.
(i) The theory of connected graphs (in the language of graphs)
(ii) The theory of simple groups (in the language of groups)
$^+$(iii) The theory of non-abelian simple groups (in the language of groups)

Lent 2010 **LOGIC AND SET THEORY – EXAMPLES 4** IBL

1. Show that the Empty-Set Axiom is deducible from the Axioms of Infinity and Separation (or, if you prefer, just from the Axiom of Infinity), and that the Axiom of Separation is deducible from the Axiom of Replacement.

2. Show that the Pair-Set Axiom is deducible from the Axioms of Empty-Set, Power-Set and Replacement.

3. Write down sentences (in the language of ZF) to express the assertions that, for any two sets A and B, the product $A \times B$ and the set of all functions from A to B exist. Indicate how to deduce these sentences from the axioms of ZF.

4. Is it true that if x is a transitive set then the relation $\in$ on x is a transitive relation? Does the converse hold?

5. What is the rank of $\{2, 3, 6\}$? What is the rank of $\{\{2, 3\}, \{6\}\}$? Work out the ranks of $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$, using your favourite constructions of these objects from ω.

6. A set x is called *hereditarily finite* if each member of $TC(\{x\})$ is finite. Prove that the class HF of hereditarily finite sets coincides with V_ω. Which of the axioms of ZF are satisfied in the structure HF (i.e. the set HF, with the relation $\in |HF$)?

7. Which of the axioms of ZF are satisfied in the structure $V_{\omega+\omega}$?

8. What is the cardinality of the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$?

9. Is there an ordinal α such that $\omega_\alpha = \alpha$?

10. Explain why, for each $n \in \omega$, there is no surjection from $\aleph_n$ to $\aleph_{n+1}$. Use this fact to show that there is no surjection from $\aleph_\omega$ to ${\aleph_\omega}^{\aleph_0}$, and deduce that $2^{\aleph_0} \neq \aleph_\omega$.

11. If ZF is consistent then, by Downward Löwenheim-Skolem, it has a countable model. Doesn't this contradict the fact that, for example, the power-set of ω is uncountable?

12. Assume that ZF is consistent. We extend the language of ZF by adding new constants $\alpha_1, \alpha_2, \ldots$, and extend the axioms of ZF by adding (for each n) the assertions that α_n is an ordinal and that $\alpha_{n+1} < \alpha_n$. Explain why this theory has a model. In this model of ZF, haven't we contradicted the fact that the ordinals are well-ordered?

13. Prove (in ZF) that a countable union of countable sets cannot have cardinality $\aleph_2$.

14. The function-classes $x + y = x \triangle y$ and $xy = x \cap y$ 'make V into a ring', in the sense that all of the axioms for a ring hold in this structure. Is it possible to make V into a ring with 1?

$^+$15. Show that the function $f(n) = 2^n$ is definable in the language of PA – in other words, find a formula $p(x, y)$ in the language of PA such that, in the natural numbers, $p(m, n)$ holds if and only if $n = 2^m$.