

Logic and Set Theory

Lectured by I. B. Leader, Lent Term 2005

Chapter 1	Propositional Logic	1
Chapter 2	Well-Orderings and Ordinals	7
Chapter 3	Posets and Zorn's Lemma	15
Chapter 4	Predicate Logic	22
Chapter 5	Set Theory	32
Chapter 6	Cardinals	41
Bonus lecture	Incompleteness	44
Examples Sheets		

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)

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, $P = \{p_1, p_2, p_3, \dots\}$. The set of *propositions*, written L or $L(P)$, is defined inductively by:

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

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

- Notes.**
1. A proposition is a finite string of symbols from the alphabet: $\perp \Rightarrow () p_1 p_2 \dots$
(Often omit the outer brackets.)
 2. ' L defined inductively' means, more precisely, that we set $L_1 = \{\perp\} \cup P$, and for $n \geq 1$, $L_{n+1} = L_n \cup \{(p \Rightarrow q) : p, q \in L_n\}$, and then $L = L_1 \cup L_2 \cup \dots$
(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	$(p \Rightarrow \perp)$
$p \vee q$	('or p or q ')	as an abbreviation for	$(\neg p) \Rightarrow q$
$p \wedge q$	('and p and q ')	as an abbreviation for	$\neg(p \Rightarrow (\neg q))$

Semantic Implication

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

1. $v(\perp) = 0$,
2. $v(p \Rightarrow q) = \begin{cases} 0 & \text{if } v(p) = 1, v(q) = 0 \\ 1 & \text{otherwise.} \end{cases}$ ($\forall p, q \in L$),

Remark. On $\{0, 1\}$, can define a constant \perp by $\perp = 0$, 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 $v : L \rightarrow \{0, 1\}$ that preserves the structure (\perp and \Rightarrow), i.e. a homomorphism.

- Proposition 1.**
- (i) If v, v' are valuations with $v(p) = v'(p) \forall p \in P$, then $v = v'$.
 - (ii) For any function $w : P \rightarrow \{0, 1\}$, \exists a valuation v such that $v(p) = w(p) \forall p \in P$.

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

Proof. (i) We have $v(p) = v'(p) \forall p \in L_1$. But if $v(p) = v'(p)$ and $v(q) = v'(q)$ then $v(p \Rightarrow q) = v'(p \Rightarrow q)$, so $v(p) = v'(p) \forall p \in L_2$.

Continuing inductively, we get $v(p) = v'(p) \forall p \in L_n$, all n .

(ii) Set $v(p) = w(p) \forall p \in P$, and $v(\perp) = 0$; 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: $v(p_1) = v(p_2) = 1, v(p_n) = 0 \forall n \geq 3$. Then

$$v(\underbrace{(p_1 \Rightarrow p_2)}_1 \Rightarrow \underbrace{p_3}_0) = 0.$$

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

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	←
0	1	0	1	so a tautology, as
0	0	1	1	last column identically 1

2. $(\neg\neg p) \Rightarrow p$, i.e. $((p \Rightarrow \perp) \Rightarrow \perp) \Rightarrow p$

$v(p)$	$v(p \Rightarrow \perp)$	$v((p \Rightarrow \perp) \Rightarrow \perp)$	$v(((p \Rightarrow \perp) \Rightarrow \perp) \Rightarrow p)$	
1	0	1	1	
0	1	0	1	←
				so a tautology

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

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

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

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

'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 $v(s) = 1 \forall s \in S$. 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 axioms and deduction rules. As axioms, we shall take:

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, \dots, 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 $\exists j, k < i$ with $t_j = (t_k \Rightarrow t_i)$.

If \exists proof of t from S , say S *proves* 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. $[p \Rightarrow ((p \Rightarrow p) \Rightarrow p)] \Rightarrow [(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)]$ (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

$$\begin{array}{l} p \quad (\text{hypothesis}) \\ q \quad (\text{modus ponens}) \end{array}$$

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

(\Leftarrow) Let t_1, \dots, t_n be a proof of q from $S \cup \{p\}$. We shall show that $S \vdash (p \Rightarrow t_i) \forall i$.

1. If t_i is an axiom, write down

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

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

$$\begin{array}{l} t_i \Rightarrow (p \Rightarrow t_i) \quad (\text{axiom}) \\ t_i \quad (\text{hypothesis}) \\ p \Rightarrow t_i \quad (\text{modus ponens}) \end{array}$$

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 $t_j \Rightarrow t_i$. By induction, may assume $S \vdash (p \Rightarrow t_j)$ and $S \vdash (p \Rightarrow (t_j \Rightarrow t_i))$, so write down

$$[p \Rightarrow (t_j \Rightarrow t_i)] \Rightarrow [(p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)] \quad (\text{axiom 2})$$

Now apply modus ponens twice. □

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

Question. How are \vdash and \models related?

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

Made up of: *soundness* — if $S \vdash t$ then $S \models t$,
adequacy — if $S \models t$ then $S \vdash t$.

Soundness is saying: our axioms are not absurd.

Adequacy is saying: 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. Must show that if v is a valuation with $v(s) = 1 \forall s \in S$ (i.e. v is a model of S) then $v(t) = 1$ (i.e. v is a model of t).

But $v(p) = 1 \forall p \in S$ (v is a model of S)

and $v(p) = 1$ for every axiom p (each axiom is a tautology),

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

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

Say S is *consistent* if $S \not\vdash \perp$. One special case of adequacy is

$$S \models \perp \Rightarrow S \vdash \perp \quad (\text{i.e. 'S has no model' implies 'S is inconsistent'})$$

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 $S \cup \{\neg t\} \vdash \perp$, whence

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

So, for adequacy, we need to show: S is consistent $\Rightarrow S$ has a model. How might we show this? Given a consistent set S , how are we going to build up a valuation v with $v(s) = 1 \forall s \in S$?

We cannot just set $v(p) = \begin{cases} 1 & \text{if } p \in S \\ 0 & \text{if not} \end{cases}$, since if for example p_3 not mentioned in S , then we should be setting $v(p_3) = 0$, $v(\neg p_3) = 0$. \times

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

Idea. We want to define $v(p) = \begin{cases} 1 & \text{if } p \in S \\ 0 & \text{if not} \end{cases}$.

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.

Claim. For any consistent $S \subset L$ and $p \in L$, either $S \cup \{p\}$ or $S \cup \{\neg p\}$ is consistent.

Proof of claim. Suppose not, then $S \cup \{p\} \vdash \perp$ and $S \cup \{\neg p\} \vdash \perp$.

Thus $S \vdash (p \Rightarrow \perp)$ (deduction theorem), i.e. $S \vdash \neg p$, so $S \vdash \perp$. \times

Now, L is countable (e.g. as each L_n is countable), so we can list L as $\{t_1, t_2, t_3, \dots\}$. 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 $\overline{S} = S_0 \cup S_1 \cup S_2 \cup \dots$. Then, for each $p \in L$, we have $p \in \overline{S}$ or $\neg p \in \overline{S}$. Also, \overline{S} is consistent: if $\overline{S} \vdash \perp$ then, as proofs are finite, we have $S_n \vdash \perp$ for some n . \times

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

Claim. v is a valuation.

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 $\overline{S} \vdash (p \Rightarrow \perp)$ and $\overline{S} \vdash p$, whence $\overline{S} \vdash \perp$. \times

Proof of claim. $v(\perp) = 0$ as $\perp \notin \overline{S}$ (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). \times
2. If $v(p) = 1$, $v(q) = 1$, we have $p \in \overline{S}$, $q \in \overline{S}$, 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 $(p \Rightarrow q) \in \overline{S}$.
3. If $v(p) = 0$, $v(q) = 1$, the same.
4. If $v(p) = 0$, $v(q) = 0$, we have $\neg p, \neg q \in \overline{S}$, and we want $(p \Rightarrow q) \in \overline{S}$. So it is enough to show that $(p \Rightarrow \perp) \vdash (p \Rightarrow q)$. By the deduction theorem, it is enough to show that $\{p \Rightarrow \perp, p\} \vdash q$. So it is enough to show that $\perp \vdash q$. But $\vdash (\perp \Rightarrow \neg q)$ (axiom 1), and $\vdash (\neg q \Rightarrow q)$ (axiom 3), so $\vdash \perp \Rightarrow q$ \square

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 . This we shall see 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 $S \vdash t$. □

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

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

Remark. Highly non-obvious.

Proof. Obvious if \vdash replaced 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 : $\text{not } x < x \ (\forall x \in X)$,
- (ii) transitive : $x < y, y < z \Rightarrow x < z \ (\forall x, y, z \in X)$,
- (iii) trichotomous : $x < y$ or $x = y$ or $y < x \ (\forall x, y \in X)$.
(We cannot have more than one of these: if $x < y, y < x$ then $x < x \ \text{\texttimes}$.)

Examples. 1. \mathbb{N} , usual $<$. ($\mathbb{N} = \{0, 1, 2, \dots\}$, 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 $x \leq y$ for ' $x < y$ or $x = y$ '.

In terms of \leq , a total ordering would be

- (i) reflexive : $x \leq x \ \forall x \in X$,
- (ii) transitive : $x \leq y, y \leq z \Rightarrow x \leq z \ \forall x, y, z \in X$,
- (iii) antisymmetric : $x \leq y, y \leq x \Rightarrow x = y \ \forall x, y \in X$,
- (iv) trichotomous : $x \leq y$ or $y \leq x \ \forall x, y \in X$.

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

$$\forall S \subset X, S \neq \emptyset \Rightarrow \exists x \in S \text{ s.t. } y \geq x \ \forall y \in S.$$

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

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

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

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

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

Remark. A total order X is a well-ordering $\Leftrightarrow X$ has no infinite strictly-decreasing sequence $x_1 > x_2 > x_3 > \dots$. Indeed, if we have such a sequence then $\{x_1, x_2, \dots\}$ 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 $x > x' > x'' > \dots$.

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

E.g. – examples 1 and 6 are isomorphic

– examples 7 and 8 are isomorphic

– examples 6 and 7 are **not** isomorphic (one has a greatest element, the other does not)

Proposition 1 (Proof by induction). Let X be well-ordered, and let $S \subset X$ such that $\forall x \in X : \text{if } y \in S \forall y < x \text{ then } x \in S$. Then $S = X$.

[Equivalently, given some property $p(x)$, if $p(y) \forall y < x \Rightarrow p(x)$ (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 $x \in S$. \times □

An example of the use of 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. E.g. – from $\mathbb{Z} \rightarrow \mathbb{Z}$, could take identity or $x \rightarrow x - 5$,
– from $\mathbb{R} \rightarrow \mathbb{R}$, could take identity or $x \rightarrow x^3$.

Proof. Let $f, g : X \rightarrow Y$ 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. E.g. – in \mathbb{R} , take $\{x : x \leq 3\}$,
– in \mathbb{Q} , take $\{x : x \leq 0 \text{ or } x^2 < 2\}$.

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, e.g. in \mathbb{Z} , take $\{1, 3, 4\}$.

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

Theorem 3 (Definition by recursion). Let X be a well-ordered set, and Y any set. Then for any $G : \mathbb{P}(X \times Y) \rightarrow Y$, there exists $f : X \rightarrow Y$ such that $f(x) = G(f|_{I_x}) \forall x \in X$. Moreover, f is unique. (In defining $f(x)$, we make use of f on $I_x = \{y : y < x\}$.)

Proof. (Existence) Define ‘ h is an attempt’ to mean:

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

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

Also, for each x , \exists attempt defined at x , again by induction. Indeed, suppose that $\forall y < x$, \exists attempt defined at y . Then \exists unique attempt h_y defined on $\{z : z \leq y\}$.

Put $h = \bigcup_{y < x} h_y : h$ is an attempt defined on I_x . So $h' = h \cup \{(x, G(h))\}$ is an attempt defined at x .

Now define f by: $f(x) = y$ if \exists 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. \square

Theorem 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 $f : Y \rightarrow X$ that is order-preserving and having image an initial segment, we need $f(x) = \min X \setminus \{f(y) : y < x, y \in Y\} \forall y \in Y$. So done (existence and uniqueness) by recursion (Theorem 3) applied to Y . Note that we cannot have $\{f(y) : y < x, y \in Y\} = X$, for example, because $f(y) \leq y \forall y$ (by induction), so that $x \notin \{f(y) : y < x, y \in Y\}$. \square

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 $X \leq Y$ if X is isomorphic to an initial segment of Y .

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

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

Proof. Suppose: not $Y \leq X$. Need $X \leq Y$. For $f : X \rightarrow Y$ to be an isomorphism with an initial segment of Y , need: $\forall x \in X, f(x) = \min Y \setminus \{f(y) : y < x\}$. So done by recursion. (Cannot have $\{f(y) : y < x\} = Y$ as then Y isomorphic to I_x . \times) \square

Proposition 6. Let X, Y be well-orderings with $X \leq Y, Y \leq X$. Then X, Y isomorphic.

Proof. Let $f : X \rightarrow Y$ be an isomorphism to an initial segment of Y , and $g : Y \rightarrow X$ be an isomorphism to an initial segment of X . Then $g \circ f$ is an isomorphism to an initial segment of X (as an initial segment of an initial segment is an initial segment), whence $g \circ f = \text{identity}$ (by uniqueness).

Similarly for $f \circ g$. So $f : X \rightarrow Y$ is a bijection. \square

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

New well-orderings from old

Find a bigger one. Given a well-ordering X , choose $x \notin X$, and define a well-ordering of $X \cup \{x\}$ by setting $y < x \forall y \in X$. 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 $X \geq X_i \forall i$. Given well-orderings $(X, <_X)$ and $(Y, <_Y)$, say X *extends* Y if $Y \subset X$, $<_X$ and $<_Y$ agree on Y , and Y is an initial segment of X . Say $\{X_i : i \in I\}$ are *nested* if $\forall i, j, X_i$ 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 \exists a well-ordering X with $X \geq X_i$ 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. 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 $X \geq X_i$ for all i . \square

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 regarded as the same if they are isomorphic.

(Just as the rationals consist of all symbols $\frac{m}{n}$ ($m, n \in \mathbb{Z}$) with two regarded as the same if $mn' = m'n$. But we cannot formalise ordinals using equivalence classes - instead, 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 $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\}$ has order-type ω .

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

Thus, $\forall \alpha, \beta$, have $\alpha \leq \beta$ or $\beta \leq \alpha$, and $\alpha \leq \beta$, $\beta \leq \alpha \Rightarrow \alpha = \beta$.

Proposition 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 ($x \leftrightarrow I_x$). \square

Sometimes write I_α for $\{\beta : \beta < \alpha\}$.

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 least element of $S \cap I_\alpha$. (I_α is well-ordered.) \square

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, so has an order-type, α say. So X is isomorphic to I_α , a proper initial segment of X . $\times \square$

Given a set $S = \{\alpha_i : i \in I\}$ of ordinals, it has an upper bound α (i.e. $\alpha \geq \alpha_i \forall i$) 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 $\sup S$.

On the following page are some ordinals. Every ordinal in that picture is countable. Is there an uncountable ordinal? I.e., is there an uncountable well-ordered set?

We can well-order: \mathbb{N} – usual order; \mathbb{Q} – biject with \mathbb{N} . What about \mathbb{R} ? – not obvious.

Amazingly, we can show:

Theorem 11. There is an uncountable ordinal.

Idea: Look at $\{\alpha : \alpha \text{ countable}\}$. (Is it a set? cf. 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 members of B are precisely the countable ordinals.

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

Note that ω_1 is the **least** uncountable ordinal (definition of B). Note also that every ordinal in the picture is $< \omega_1$. The same proof shows

Theorem 11' (Hartogs' Lemma). For any set X , \exists ordinal α that does not inject into X .

Proof. Same, with ' X ' in place of ' \mathbb{N} '. \square

The least such α is denoted $\gamma(X)$.

Successors and Limits

Let α be an ordinal. Does it have a greatest element? (E.g. $I_\alpha = \{\beta : \beta < \alpha\}$.)

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

If no: $\forall \beta < \alpha$, $\exists \gamma < \alpha$ s.t. $\gamma > \beta$, so $\sup\{\beta : \beta < \alpha\} = \alpha$. 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)

Some ordinals.

1 2 3 4 \vdots ω $\omega + 1 \leftarrow \text{officially } \omega^+$ $\omega + 2$ $\omega + 3$ \vdots $\omega + \omega = \omega 2 \leftarrow \text{officially } \sup\{\omega, \omega + 1, \omega + 2, \dots\}$ $\omega 2 + 1$ $\omega 2 + 2$ $\omega 2 + 3$ \vdots $\omega 3$ \vdots $\omega 4$ \vdots $\omega 5$ \vdots $\omega \omega = \omega^2 \leftarrow \text{officially } \sup\{\omega, \omega 2, \omega 3, \dots\}$ $\omega^2 + 1$ $\omega^2 + 2$ \vdots $\omega^2 + \omega$ $\omega^2 + \omega + 1$ \vdots $\omega^2 + \omega 2$ \vdots $\omega^2 + \omega^2 = \omega^2 2$ $\omega^2 2 + 1$ \vdots $\omega^2 2 + \omega$ \vdots $\omega^2 3$ \vdots $\omega^2 4$ \vdots ω^3 \vdots $\omega^3 2$ \vdots $\omega^3 3$ \vdots ω^4 \vdots ω^5 \vdots ω^ω \vdots $\omega^{\omega 2}$ \vdots $\omega^{\omega 3}$ \vdots $\omega^\omega \omega = \omega^{\omega+1}$ \vdots $\omega^{\omega+2}$ \vdots	\vdots $\omega^{\omega+3}$ \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 ω^{ω^3} \vdots ω^{ω^4} \vdots ω^{ω^ω} \vdots $\omega^{\omega^{\omega^2}}$ \vdots $\omega^{\omega^{\omega^3}}$ \vdots $\omega^{\omega^{\omega^\omega}}$ \vdots $\omega^{\omega^{\omega^{\omega^{\dots}}}} = \epsilon_0$ $\epsilon_0 + 1$ $\epsilon_0 + 2$ \vdots $\epsilon_0 + \omega$ \vdots $\epsilon_0 2$ \vdots $\epsilon_0 3$ \vdots $\epsilon_0 \omega$ \vdots ϵ_0^2 \vdots ϵ_0^3 \vdots ϵ_0^ω \vdots $\epsilon_0^{\omega^2}$ \vdots $\epsilon_0^{\omega^3}$ \vdots $\epsilon_0^{\omega^\omega}$ \vdots $\epsilon_0^{\omega^{\omega^{\omega^{\dots}}}} = \epsilon_0^{\epsilon_0}$ \vdots $\epsilon_0^{\epsilon_0}$ \vdots $\epsilon_0^{\epsilon_0^{\epsilon_0}}$ \vdots $\epsilon_0^{\epsilon_0^{\epsilon_0^{\dots}}}$ $\epsilon_0^{\epsilon_0^{\epsilon_0^{\dots}}} = \epsilon_1$
---	--

‘ ... and so on ... ’

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. $\omega + 1 = (\omega + 0)^+ = \omega^+$, and $\omega + 2 = (\omega + 1)^+ = \omega^{++}$, but

$$1 + \omega = \sup\{1 + \gamma : \gamma < \omega\} = \omega, \text{ so } + \text{ is not commutative.}$$

Remark. Officially, ‘recursion on the ordinals’ would mean: to define $\alpha + \beta$ (given α, β), define $\alpha + \beta'$, all $\beta' \leq \beta$, by recursion (plus uniqueness). Similarly for proof by induction on the ordinals. (If $p(\alpha)$ false for some α , then p is not everywhere true on $\{\alpha' : \alpha' \leq \alpha\}$.)

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

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

$$\beta < \gamma \Rightarrow \beta^+ \leq \gamma \Rightarrow \alpha + \beta^+ \leq \alpha + \gamma \Rightarrow (\alpha + \beta)^+ \leq \alpha + \gamma \Rightarrow \alpha + \beta < \alpha + \gamma$$

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

Proposition 12. $\forall \alpha, \beta, \gamma : \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.

Proof. By induction on γ (α, β fixed)

$$\mathbf{0.} \alpha + (\beta + 0) = \alpha + \beta = (\alpha + \beta) + 0$$

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

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 no greatest member of $\{\beta + \gamma : \gamma < \lambda\}$.

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

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

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

\geq : For $\delta < \beta + \lambda$, have $\delta \leq \beta + \gamma$, some $\gamma < \lambda$ (definition of $\beta + \lambda$), whence $\alpha + \delta \leq \alpha + (\beta + \gamma)$. Thus RHS \subset 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. we have $\omega + 1 = \xleftarrow{\omega} \bullet = \omega^+$, and $1 + \omega = \bullet \xleftarrow{\omega} = \omega$

Proposition 13. The two definitions of ordinal addition coincide.

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

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

Successors. $\alpha +' \beta^+ = \text{order-type of } \begin{matrix} \leftarrow \alpha \\ \leftarrow \beta \end{matrix} \bullet = (\alpha +' \beta)^+ = (\alpha + \beta)^+ = \alpha + \beta^+$

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

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

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

Ordinal Multiplication

$\alpha\beta$ is defined by recursion on β as follows:

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

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.)

$\omega\omega = \sup\{\omega\gamma : \gamma < \omega\} = \sup\{0, \omega, \omega + \omega, \omega + \omega + \omega, \dots\}$ (as in our 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' $\beta \left\{ \begin{matrix} \vdots \\ \leftarrow \alpha \\ \leftarrow \alpha \end{matrix} \right.$

E.g. $\omega 2 = \text{order-type of } \left\{ \begin{matrix} \leftarrow \omega \\ \leftarrow \omega \end{matrix} \right. = \omega + \omega$

$2\omega = \text{order-type of } \left\{ \begin{matrix} \vdots \\ \leftarrow \bullet\bullet \\ \leftarrow \bullet\bullet \end{matrix} \right. = \omega$.

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 :

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

Note. $2^\omega = \sup\{2^\gamma : \gamma < \omega\} = \omega$, countable.

Chapter 3 : Posets and Zorn's Lemma

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

- (i) reflexive : $x \leq x$ ($\forall x \in X$)
- (ii) transitive : $x \leq y, y \leq z \Rightarrow x \leq z$ ($\forall x, y, z \in X$).
- (iii) antisymmetric : $x \leq y, y \leq x \Rightarrow x = y$ ($\forall x, y \in X$).

Equivalently, writing $x < y$ to mean ' $x \leq y$ and $x \neq y$ ', conditions are

- (i) irreflexivity : not $x < x$ ($\forall x \in X$)
- (ii) transitivity : $x < y, y < z \Rightarrow x < z$ ($\forall x, y, z \in X$).

Examples.

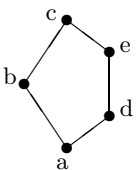
1. any total order.
2. $(\mathbb{N}^+, \text{'divides'})$. Not a total order (3 and 5 are incomparable)
3. For S any set, take $\mathbb{P}(S)$, with $A \leq B$ if $A \subset B$. (Very important.)
4. Take $X =$ any subset of $\mathbb{P}(S)$, same \leq .
E.g. V a vector space, $X =$ all subspaces of V .

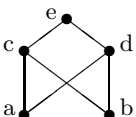
5. $c \bullet \qquad \bullet e$ — meaning: $a \leq b, b \leq c, a \leq d, d \leq e$,
 $\qquad b \bullet \qquad \bullet d$ and everything following by transitivity,
 $\qquad \qquad a \bullet$ e.g. $a \leq c$, not $b \leq d$.

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

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

Examples (continued)

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

7. 

8. $\bullet \bullet \bullet \bullet \bullet$
 $a \quad b \quad c \quad d \quad e$

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

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

- chains can be uncountable, e.g. (\mathbb{R}, \leq) .
- in 2 , $\{1, 2, 4, 8, 16, \dots\}$ is a chain.
- in 5 , $\{a, b, c\}$ is a chain, as is $\{a, b\}$.

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

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

- in 5 , $\{b, e\}$ is an antichain.
- in 8 , the whole of X is an antichain.

For $S \subset X$ and $x \in X$, say x is an *upper bound* for S if $y \leq x \forall y \in S$. Say x is a *least upper bound* or *supremum* for S if x is an upper bound for S , and every upper bound y for S has $y \geq x$.

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 ').]

- in \mathbb{Q} , $S = \{x : x^2 < 2\}$ has 6 as an upper bound, but no supremum.
- in 5 , $\{a, b, c\}$ has only upper bound being c , so supremum = c , but $\{b, d\}$ has no upper bound.
- in 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} , $\sup \in \{x : x \leq 1\}$, but $\sup \notin \{x : x < 1\}$.

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

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

- \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 , $i \in I$, take $\bigcup_{i \in I} A_i$
- $[0, 1]$ is complete, but $(0, 1)$ is not complete – e.g. X itself has no supremum.

Note. X complete $\Rightarrow X$ has a greatest element (an x with $x \geq y \forall y$), namely $\bigvee X$, and also a least element (an x with $x \leq y \forall y$), namely $\bigvee \emptyset$.

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

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 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 $f : X \rightarrow X$ order-preserving. Then f has a fixed point.

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

Claim. $s \in E$.

Proof of claim. Need $s \leq f(s)$, i.e. need $f(s)$ an upper bound for E (then $s \leq f(s)$ as s is the **least** upper bound).

But $x \in E \Rightarrow x \leq s \Rightarrow f(x) \leq f(s) \Rightarrow x \leq f(x) \leq f(s)$.

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

An application of Knaster-Tarski:

Corollary 2 (Schröder-Bernstein Theorem).

Let A, B be sets, and let $f : A \rightarrow B$ and $g : B \rightarrow A$ 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 $h : A \rightarrow B$ by taking $h = f$ on P and g^{-1} on Q .)

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

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

Zorn's Lemma

For X a poset, $x \in X$, say x *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. (Note: 'non-empty' not strictly necessary.)

Proof. Suppose X has no maximal element. So for each $x \in X$, $\exists x' \in X$ 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_α , $\alpha < \gamma$ recursively by:

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

The x_α , $\alpha < \gamma$, are distinct, so we have injected γ into X . \aleph \square

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

A typical application of Zorn: 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 $\{1, x, x^2, x^3, \dots\}$ is a basis

2. $V =$ set of all real sequences. Let $e_n = (0, 0, \dots, 0, 1, 0, 0, \dots)$.

Then the e_n do **not** form a basis – e.g. $(1, 1, 1, \dots)$ is not in the span. In fact, there is no countable basis (easy exercise).

Even more: it is impossible to give an ‘explicit’ basis.

3. Consider \mathbb{R} as a vector space over \mathbb{Q} . Is there a basis?

Theorem 4. Every vector space has a basis. (‘Go for a maximal linearly independent set.’)

Proof. 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. \times)

First, $X \neq \emptyset$, as $\emptyset \in X$.

Given chain $\{A_i : i \in I\}$, let $A = \bigcup_{i \in I} A_i$. Then $A \supset A_i \forall i$, so just need $A \in X$, i.e. A linearly independent.

Suppose we have a linear dependence in A , say $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$, where $x_1, \dots, x_n \in A$ and $\lambda_1, \dots, \lambda_n$ scalars, not all 0. Have $x_1 \in A_{i_1}, \dots, x_n \in A_{i_n}$, some $i_1, \dots, i_n \in I$. But some A_{i_k} has $A_{i_1}, \dots, A_{i_n} \subset A_{i_k}$ (as the A_i are a chain), contradicting A_{i_k} linearly independent. \times

So, by Zorn, \exists maximal $A \in X$. □

Another application of Zorn: completeness theorem for propositional logic, with **no** restrictions on countability of the primitive propositions.

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

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

(Then done, by setting $v(p) = \begin{cases} 1 & \text{if } p \in \bar{S} \\ 0 & \text{if not} \end{cases}$, as in Chapter 1.)

So enough to find a maximal consistent $\bar{S} \supset S$. (If $t \notin \bar{S}$ then $\bar{S} \cup \{t\} \vdash \perp$, so $\bar{S} \vdash (t \Rightarrow \perp)$, so $(t \Rightarrow \perp) \in \bar{S}$, by maximality of \bar{S} .)

So 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 each $T_i \supset S$). If T inconsistent, have $\{t_1, \dots, t_n\} \subset T$ with $\{t_1, \dots, t_n\} \vdash \perp$ (as proofs are finite). Then $t_1 \in T_{i_1}, \dots, t_n \in T_{i_n}$, some $i_1, \dots, i_n \in I$. But some T_{i_k} has $T_{i_1}, \dots, T_{i_n} \subset T_{i_k}$ (as the T_i form a chain), contradicting T_{i_k} consistent. \times

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, $(A, R) \geq (A', R')$ if $A' \subset A$, R and R' agree on A' , and A' is an initial segment of A in the ordering R .)

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

Given a chain $\{(A_i, R_i) : i \in I\}$, the (A_i, R_i) form a nested family, so $(\bigcup_{i \in I} A_i, \bigcup_{i \in I} R_i)$ 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 . \times □

Zorn's Lemma and the Axiom of Choice

In our proof of Zorn's Lemma, we picked, 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, \dots , 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 $f : I \rightarrow \bigcup_{i \in I} A_i$ such that $f(i) \in A_i \forall i$.

This is of a different character from the other rules for building sets (e.g. 'given A and B , can form $A \cup B$ '), 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 $\exists x$ 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 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 $f : J \rightarrow \bigcup_{i \in I} A_i$, some $J \subset I$, such that $f(j) \in A_j \forall j \in J$.

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

$X \neq \emptyset : (\emptyset, \emptyset) \in X$.

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

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

Actually, 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).

So, in fact, ZL \Leftrightarrow AC \Leftrightarrow WO (given the other rules).

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

**

Some notions related to completeness

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

E.g. – 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 $f : X \rightarrow X$ is *inflationary* if $f(x) \geq x \forall x \in X$.

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 $x \leq f(x)$ implies $x = f(x)$.)

Can prove Bourbaki-Witt **without** AC: $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_\omega, x_1 = f(x_0), x_2 = f(x_1), \dots$
In fact, easy to get from B-W to ZL (using AC). So can view B-W as ‘the choice-free part of ZL’.

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 least upper bound of $\{a, b\}$

$a \wedge b$, ‘ a meet b ’, for 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 $\mathbb{P}(S)$ ’ :

1. $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ ($\forall a, b, c, \in X$)
2. $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$ ($\forall a, b, c, \in X$)
3. $\forall a \in X, \exists a' \in X$ 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** $\mathbb{P}(S)$.)

Lindenbaum algebra

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

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

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

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

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

**

Chapter 4 : Predicate Logic

Overview of the set-up

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

$$(\forall x, y, z) \quad (m(x, m(y, z)) = m(m(x, y), z)) \tag{1}$$

$$(\forall x) \quad (m(x, e) = x \wedge m(e, x) = x) \tag{2}$$

$$(\forall x) \quad (m(x, i(x)) = e \wedge m(i(x), x) = e) \tag{3}$$

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

$$(\forall x) \quad (x \leq x) \quad (\text{i.e., } (x, x) \in \leq)$$

$$(\forall x, y, z) \quad ((x \leq y \wedge y \leq z) \Rightarrow x \leq z)$$

$$(\forall x, y) \quad ((x \leq y \wedge y \leq x) \Rightarrow x = y)$$

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

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

Language

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

Variables

Have variables x_1, x_2, x_3, \dots (We may use x, y, \dots)

Terms

Defined inductively by:

1. Every variable is a term
2. If $f \in \Omega$, $\alpha(f) = n$ and t_1, \dots, t_n terms, then so is $ft_1 \dots t_n$.
(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

1. \perp is an atomic formula
2. If s, t terms then $(s = t)$ is an atomic formula
3. If $\phi \in \Pi$, $\alpha(\phi) = n$, and t_1, \dots, t_n terms, then $\phi(t_1 t_2 \dots t_n)$ 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 $\Omega = \emptyset$, $\Pi = \{\leq\}$, $\alpha(\leq) = 2$):
 $x_1 = x_1$, $x_1 \leq x_2$ (officially ‘ $\leq(x_1, x_2)$ ’)

Formulae

Defined inductively by

1. Every atomic formula is a formula
2. If p, q formulae, then so is $(p \Rightarrow q)$
3. If p a formula and x a variable, then $(\forall x)p$ is a formula

E.g. In language of groups:

$$(\forall x)(m(x, e) = x), \quad (\forall x)(m(x, x) = e), \quad (\forall x)(m(x, x) = e \Rightarrow x = e).$$

Remarks. 1. A formula is a string of symbols

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

Closed terms

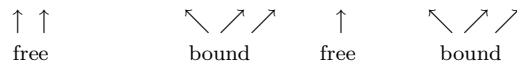
A term is *closed* if it contains no variables.

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



$m(x, x) = e \Rightarrow (\forall x)(\forall y)(m(x, y) = m(y, x))$ ← unhelpful



Sentences

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

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 occurrence 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

1. For each $f \in \Omega$, a function $f_A : A^n \rightarrow A$, where $n = \alpha(f)$
2. For each $\phi \in \Pi$, a set $\phi_A \subset A^n$, where $n = \alpha(\phi)$

E.g. – for $L =$ language of groups, an L -structure is a non-empty set A , with functions $m_A : A^2 \rightarrow A$, $i_A : A \rightarrow A$, $e_A \in A$

– for $L =$ language of posets, an L -structure is a non-empty set A , with a subset $\leq_A \subset A^2$

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.)

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

$$\text{for } f \in \Omega, \alpha(f) = n, \text{ and } t_1, \dots, t_n \text{ closed terms: } (ft_1 \dots t_n)_A = f_A(t_{1A} \dots t_{nA})$$

(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 a sentence p in an L -structure A to be $p_A \in \{0, 1\}$ defined inductively by:

$$\begin{aligned} \perp_A &= 0 \\ (s = t)_A &= \begin{cases} 1 & \text{if } s_A = t_A \\ 0 & \text{if not} \end{cases} && \text{(any closed terms } s, t) \\ \phi(t_1 \dots t_n)_A &= \begin{cases} 1 & \text{if } (t_{1A} \dots t_{nA}) \in \phi_A \\ 0 & \text{if not} \end{cases} && \text{(each } \phi \in \Pi, \alpha(\phi) = n, \\ &&& \text{closed terms } t_1, \dots, t_n) \\ (p \Rightarrow q)_A &= \begin{cases} 0 & \text{if } p_A = 1, q_A = 0 \\ 1 & \text{if not} \end{cases} && \text{(} p, q \text{ sentences)} \\ ((\forall x)p)_A &= \begin{cases} 1 & \text{if } p[\bar{a}/x]_A = 1 \forall a \in A \\ 0 & \text{if not} \end{cases} \end{aligned}$$

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

If $p_A = 1$, say p holds in A , or p 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 $p \forall p \in T$ (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 .

Examples

Theory of groups

L = language of groups: $\Omega = \{m, i, e\}$, $\Pi = \emptyset$
 arities: 2 1 0

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

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 .

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 distributive over $+$
 \times commutative
 $(\forall x)(1x = x)$
 $\neg(0 = 1)$
 $(\forall x)((\neg(x = 0)) \Rightarrow (\exists y)(xy = 1))$

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

So T axiomatises the theory of fields.

Have, e.g., $T \models$ ‘inverses are unique’ :

$$(\forall x)((x \neq 0) \Rightarrow (\forall y)(\forall z)((xy = 1 \wedge xz = 1) \Rightarrow y = z))$$

Theory of posets

$\Omega = \emptyset$, $\Pi = \{\leq\}$ (arity 2)

$$T : \begin{array}{l} (\forall x)(x \leq x) \\ (\forall x)(\forall y)(\forall z)((x \leq y \wedge y \leq z) \Rightarrow x \leq z) \\ (\forall x)(\forall y)((x \leq y \wedge y \leq x) \Rightarrow x = y) \end{array}$$

Theory of graphs

$\Omega = \emptyset$, $\Pi = \{a\}$ (a = ‘is adjacent to’, arity 2)

$$T : \begin{array}{l} (\forall x)(\neg a(x, x)) \\ (\forall x, y)(a(x, y) \Rightarrow a(y, x)) \end{array}$$

If p holds in **all** L -structures, we call p a *tautology*, written $\models p$.
 (Equivalently, $\models p$ means $\emptyset \models p$.)

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

Proofs

Logical axioms. (3 usual, 2 for ‘=’, 2 for ‘ $\forall x$ ’)

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 inference

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 by a rule of deduction.

Say ‘ S proves p ’ or ‘ p is a theorem of 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. $\{x = y, x = z\} \vdash y = z$.

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

Proposition 1 (Deduction Theorem).

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

Proof. (\Rightarrow) As for propositional logic: have $S \vdash (p \Rightarrow q)$, $S \cup \{p\} \vdash p$, so $S \cup \{p\} \vdash q$ by modus ponens.

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

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

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 $S \vdash r$, so $S \vdash (\forall x) r$ (definition of generalisation), so $S \vdash (p \Rightarrow (\forall x) r)$, by axiom 1. \square

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.

**

Proposition 2 (Soundness Theorem).

Let S be a set of sentences, 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. \square

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

Theorem 3 (Model existence lemma, or completeness theorem).

Let L be a language, and S be a consistent set of sentences in 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, closed terms $1 + 0$ and 1 are distinct. So we'd quotient out A by $s \sim t$ if $S \vdash (s = t)$.

3. For 'fields of characteristic 2 or 3': $S =$ field axioms, with ' $1 + 1 = 0 \vee 1 + 1 + 1 = 0$ '. Then $S \not\vdash 1 + 1 = 0$, and $S \not\vdash 1 + 1 + 1 = 0$, 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 c , and add ' $cc = 1 + 1$ ' to S . But this has added to the language, hence it is no longer maximal consistent, so we must loop back to Idea 3.

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 ($\forall p : S_1 \vdash p$ or $S_1 \vdash \neg p$).

For each $(\exists x) p \in S_1$, add a new constant c to the language, and add $p[c/x]$ 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 $p[c/x] \in T_1$, some constant c . 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 $\overline{S} = S_1 \cup S_2 \cup \dots$, in language $\overline{L} = L(\Omega \cup C_1 \cup C_2 \cup \dots, \Pi)$.

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

Proof of claim.

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

Complete. For sentence $p \in \overline{L}$, have $p \in L_n$ for some n (since p can only mention finitely many constants). So $S_n \vdash p$ or $S_n \vdash (\neg p)$ (as S_n 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 \exists constant c , in language L_{n+1} , with $p[c/x] \in T_n$ (definition of T_n). Then $p[c/x] \in \overline{S}$ (and $c \in \overline{L}$).

On 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.

For $f \in \Omega$, let $f_A([t_1], \dots, [t_n]) = [f(t_1, \dots, t_n)]$.

For $\phi \in \Pi$, let $\phi_A = \{([t_1], \dots, [t_n]) : \overline{S} \vdash \phi(t_1, \dots, t_n)\}$. (Both clearly well-defined.)

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

Proof of claim. An easy induction.

Atomic formulae

$$\begin{aligned} s = t : \overline{S} \vdash (s = t) &\Leftrightarrow [s] = [t] && \text{(definition of } \sim) \\ &\Leftrightarrow s_A = t_A && \text{(definition of operations on } A) \\ &\Leftrightarrow (s = t)_A = 1 && \text{(i.e. } s = t \text{ holds in } A) \end{aligned}$$

$\phi(t_1 \dots t_n) : \text{ same}$

$$\perp : \overline{S} \not\vdash \perp, \text{ and } \perp_A = 0$$

Induction step

$$\begin{aligned} p \Rightarrow q : \overline{S} \vdash (p \Rightarrow q) &\Leftrightarrow \overline{S} \vdash \neg p \text{ or } \overline{S} \vdash q \\ (\Rightarrow: \text{ if } \overline{S} \not\vdash \neg p, \overline{S} \not\vdash q, \text{ then } \overline{S} \vdash p, \overline{S} \vdash \neg q, \text{ by completeness of } \overline{S}. \times) & \end{aligned}$$

$$\Leftrightarrow p_A = 0 \text{ or } q_A = 1 \text{ (induction hypothesis)}$$

$$\Leftrightarrow (p \Rightarrow q)_A = 1$$

$$(\exists x) p : \overline{S} \vdash (\exists x) p \Leftrightarrow \overline{S} \vdash p[t/x], \text{ some closed term } t$$

$$(\Rightarrow: \overline{S} \text{ has witnesses})$$

$$\Leftrightarrow p[t/x]_A = 1, \text{ some closed term } t$$

$$\text{(induction hypothesis)}$$

$$\Leftrightarrow (\exists x) p \text{ holds in } A \text{ (as } A \text{ is the set of}$$

$$\text{(equivalence classes of) closed terms)}$$

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

By remarks before Theorem 3, we have

Corollary 4 (Adequacy Theorem)

Let S be a theory, 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 theory, p a sentence. Then $S \models p \Leftrightarrow S \vdash p$.

Proof. (\Rightarrow) Soundness
 (\Leftarrow) 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).

Corollary 6 (Compactness Theorem)

Let S be a theory. 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 $S \models p$.

A typical application of 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

$$\begin{array}{ll} (\exists x_1)(\exists x_2) (x_1 \neq x_2) & \text{'|G|} \geq 2' \\ (\exists x_1)(\exists x_2)(\exists x_3) (x_1 \neq x_2) \wedge (x_1 \neq x_3) \wedge (x_2 \neq x_3) & \text{'|G|} \geq 3' \end{array}$$

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

Note. We **are** using the full strength of the model existence lemma.

Similarly,

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

Proof. Add sentences as above, and apply compactness as above. □

'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). Done by compactness. \square

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

Corollary 9 (Downward Löwenheim-Skolem Theorem)

Let S be a consistent theory in a countable language (i.e. Ω, Π countable). Then S has a countable model.

Proof. The model constructed in Theorem 3 is countable. \square

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

These axioms are called *Peano arithmetic* (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 should be missing sets like $\{x : x \geq y\}$, y fixed. 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

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

So only **countably many** sets are definable.

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

– set of primes. $p(x) : (\forall y, z) (yz = x \Rightarrow (y = 1 \vee z = 1))$

– $\{x : x \text{ is a power of } 2\}$. $p(x) : (\forall y) (y \text{ prime} \wedge y \text{ divides } x \Rightarrow y = 2)$.

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. \forall sentences p , either $\text{PA} \vdash p$ or $\text{PA} \vdash \neg p$?

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

So there exists p : $\text{PA} \not\vdash p$, $\text{PA} \not\vdash \neg p$. But p true in \mathbb{N} or $\neg p$ true in \mathbb{N} .

Conclusion. \exists sentence p such that p holds in \mathbb{N} but $\text{PA} \not\vdash p$.

(No contradiction with completeness theorem: p holds in **every** model of PA \Rightarrow $\text{PA} \vdash p$.)

Chapter 5 : Set Theory

Try to study the ‘universe of sets’.

We shall view set theory as just another first-order theory.

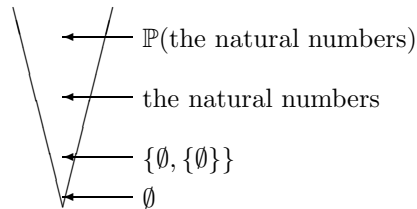
Zermelo-Frankel Set Theory

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.

The ‘universe of sets’ will mean a model (V, \in) of the ZF axioms.

What does (V, \in) look like?



‘Does V really look like this?’

Could view Chapter 5 as a worked example from Chapter 4, but very scary, as (hopefully) every model of ZF 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.’

$$(\forall t_1) \dots (\forall t_n)(\forall x)(\exists y) \underbrace{(\forall z) (z \in y \Leftrightarrow (z \in x \wedge p))},$$

‘the elements of y are those z such that...’

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

Note. Need parameters t_1, \dots, t_n , e.g. to allow $\{z \in x : t \in z\}$, and 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 $\{z \in x : p(x)\}$ 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 of x and y to be $(x, y) = \{\{x\}, \{x, y\}\}$.

Easy to check that $(x, y) = (z, t) \Leftrightarrow (x = z \wedge y = t)$. (It 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

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

and x is the domain of f means $(f \text{ is a function}) \wedge (\forall y)(y \in x \Leftrightarrow (\exists z)((y, z) \in f))$,

and $f : x \rightarrow y$ means $(f \text{ is a function}) \wedge (x \text{ is the domain of } f) \wedge (\forall z)((\exists t)((t, z) \in f) \Rightarrow z \in y)$.

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. $A_1 \cup A_2 \cup A_3 = \bigcup_{x \in \{A_1, A_2, A_3\}} x \iff z \in \text{this} \iff z \in A_1 \text{ or } z \in A_2 \text{ or } z \in A_3$.

We write $\bigcup x$ for this set.

Remark. 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.

6. **Power-set axiom.** ‘Can form power-sets.’

$$(\forall x)(\exists y)(z \in y \Rightarrow z \subset x)$$

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

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

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 $x^+ = x \cup \{x\}$ (the *successor* of x), and then $\emptyset, \emptyset^+, \emptyset^{++}, \emptyset^{+++}, \dots$ are distinct.

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

Write 0 for \emptyset , 1 for \emptyset^+ , 2 for \emptyset^{++} , etc. Then

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

Does V have an infinite set?

From real world: 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: $x \in V$ such that $\emptyset \in x, \emptyset^+ \in x, \emptyset^{++} \in x, \dots$

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

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

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

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

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

and x is *countable* to mean $(x \text{ is finite}) \vee (x \text{ bijects with } \omega)$.

8. Axiom of foundation.

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

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

What we want is: 'sets are built up from simpler sets'.

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

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

9. Axiom of replacement

Often we say: ‘for each $i \in I$, have A_i : take $\{A_i : i \in I\}$ ’.

Problem: no reason why $i \mapsto A_i$ should be a function. (I.e., why should there be a set $\{(i, A_i) : i \in I\}$?)

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 $V \times V$.)

Let (V, \in) be a model of ZF. 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 any set t , we have class $\{x : t \in x\}$ – take p to be ‘ $t \in x$ ’.

Say C is a *proper class* if $(\forall y) \neg(\forall x) (x \in y \Leftrightarrow p)$.

E.g. V is a proper class.

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 $p = ‘y = \{x\}’$.

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

$$\underbrace{(\forall t_1) \dots (\forall t_n)}_{\text{parameters}} \left(\underbrace{[(\forall x)(\forall y)(\forall z) ((p \wedge p[z/y]) \Rightarrow z = y)]}_{\text{if } p \text{ a function-class}} \right) \\ \Rightarrow [(\forall x)(\exists y)(\forall z) (z \in y \Leftrightarrow (\exists t)(t \in x \wedge \underbrace{p[t/x, z/y]}_{‘F(t)=z’}))], \\ \underbrace{\hspace{15em}}_{\text{image of } x \text{ under } p \text{ is a set}}$$

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

Note. This is an axiom-scheme.

For example, taking $F(x) = \{x\}$, we have that for any set x we can form $\{\{t\} : t \in x\}$. 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**.

Note. Axiom of choice not included – ZF+AC, denoted ZFC, where

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

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

Goal. ‘What does V look like?’

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

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

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

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

(Officially: ‘let (V, \in) be a model of ZF. Then ... holds in V ’, or ‘ZF \vdash ...’.)

Note. 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 take ‘ $x \cup (\bigcup x) \cup (\bigcup \bigcup x) \cup (\bigcup \bigcup \bigcup x) \cup \dots$ ’.

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

Why is this a function-class?

Say f is an attempt if

$$\begin{aligned} & f \text{ is a function} \wedge \text{dom } f \in \omega \wedge \text{dom } f \neq \emptyset \wedge f(0) = x \\ & \wedge (\forall n) ((n \in \text{dom } f \wedge n \neq 0) \Rightarrow f(n) = \bigcup f(n-1)). \end{aligned}$$

Then it is easy to check that

$$\begin{aligned} & (\forall f)(\forall g)(\forall n) ((f \text{ is an attempt} \wedge g \text{ is an attempt} \\ & \wedge n \in \text{dom } f \wedge n \in \text{dom } g) \Rightarrow f(n) = g(n)) \end{aligned}$$

(by usual ω -induction) and also that

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

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

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

□

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, \dots, t_n, x ,

$$(\forall t_1) \dots (\forall t_n) \left([(\forall x) ((\forall y \in x) \underbrace{p(y)}_{p[y/x]}) \Rightarrow p(x)] \Rightarrow (\forall x) p(x) \right).$$

Proof. Given t_1, \dots, 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 \aleph .)

But $\{x : \neg p(x)\}$ need not be a set!

Let $u = \{y \in \underbrace{TC(\{x\})}_{\text{a set}} : \neg p(y)\}$.

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 $z \in y \Rightarrow z \in TC(\{x\})$ (as $TC(\{x\})$ transitive),

$\Rightarrow z \notin u$ (as u minimal), $\Rightarrow p(z)$. \aleph □

Important remark. 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 $(\forall y) (x \in y \Rightarrow y \text{ has a minimal member})$. So foundation says: every x is regular.

To prove this by \in -induction, enough to show:

Claim. If every $y \in x$ is regular then x is regular.

Proof of claim. 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 $((x, y) \in G \Leftrightarrow p(x, y)$, some formula p), everywhere defined.

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

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

Proof. (Existence) Say ‘ f is an attempt’ if

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

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

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

If each $y \in x$ has an attempt defined at y , then for each $y \in x$ there exists 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: $q(x, y) = ‘(\exists f) (f \text{ an attempt} \wedge x \in \text{dom } f \wedge f(x) = y)’$.

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

\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.) \longleftarrow 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 $a' = \emptyset, b' = \{\emptyset\}, c' = \{\{\emptyset\}\}$: have $\{a', b', c'\}$, transitive, and $x r y \Leftrightarrow x' \in y'$.

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

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

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 \exists transitive set b , and a bijection $f : a \rightarrow b$ 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’ trivially necessary.

Proof. (Existence) Define $f(x) = \{f(y) : y r x\}$ – a definition by r -recursion on the set a . Note that f is a function (not just a function-class), by replacement (it is an image of a). Let $b = \{f(x) : x \in a\}$, 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 $f(y) \in f(x) \Leftrightarrow y \in x$). 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 $\{f(t) : t r x\} = \{f(u) : u r y\}$ (definition of f)
 so $\{t : t r x\} = \{u : u r y\}$ (induction hypothesis)
 so $x = y$ by extensionality.

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

An *ordinal* is a transitive set that is well-ordered by \in . (Equivalently, ‘linearly ordered’, thanks to foundation.)

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

So a well-ordered set x is order-isomorphic to a unique ordinal (by Mostowski) – write *order-type of x* for this ordinal.

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 $\alpha < \beta \Leftrightarrow \alpha \in \beta$.

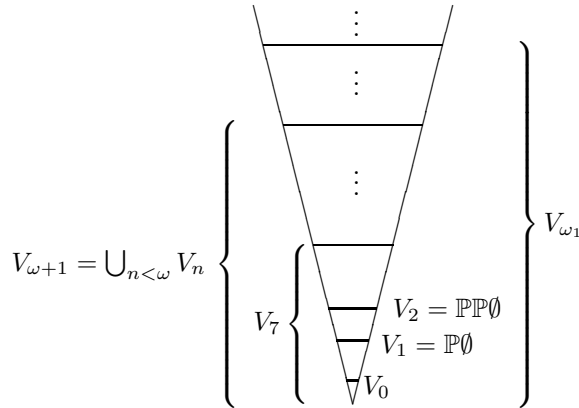
So, for example, $\alpha^+ = \alpha \cup \{\alpha\}$ and $\sup\{\alpha_i : i \in I\} = \bigcup\{\alpha_i : i \in I\}$. (Unhelpful.)

Picture of the Universe

‘Start with \emptyset ; take \mathbb{P} repeatedly.’

Define sets V_α , $\alpha \in \text{ON}$ (ordinals),
by \in -recursion.

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



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

Lemma 5. Each V_α transitive.

Proof. Induction on α .

0. V_0 transitive.

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

Limits. Any union of transitive sets is transitive. □

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

Proof. Induction on β (α fixed).

If $\beta = \alpha$. Done.

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 $V_\alpha \subset \mathbb{P}V_\beta = V_{\beta^+}$.

Unions. Done. □

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

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

2. If $x \subset V_\alpha$ then \exists least such α , called *rank* of x .
(E.g. $\text{rank}(\omega) = \omega$, and indeed $\text{rank}(\alpha) = \alpha \forall \alpha \in \text{ON}$.)

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

So may assume that $(\forall y \in x) (\exists \alpha) (y \subset V_\alpha)$. So for each $y \in x$, have $y \subset V_{\text{rank}(y)}$, 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

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

Want to define ‘card x ’, such that $\text{card } x = \text{card } y \Leftrightarrow (x \leftrightarrow y)$.

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

With AC. Know $x \leftrightarrow \alpha$, some ordinal α , so could define $\text{card } x$ to be the least such α .

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

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. $0, 1, 2, \dots, \omega, \omega_1, \gamma(X)$ for any set X . But **not** ω^2 (as ω^2 countable, so $\omega^2 \leftrightarrow \omega$).

Define ω_α , $\alpha \in \text{ON}$, 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 initial (by induction).

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

Write \aleph_α (‘aleph- α ’) for $\text{card } \omega_\alpha$.

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

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

For cardinals m, n (a *cardinal* is $\text{card } x$, some x), say $m \leq n$ if \exists injection $M \rightarrow N$, where M, N sets with $\text{card } M = m$, $\text{card } N = n$ (does not depend on choice of M, N).

Similarly, $m < n$ means $(m \leq n \text{ and } m \neq n)$.

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

Note. – If $m \leq n$, $n \leq m$, then $n = m$ (Schröder-Bernstein), so \leq is a partial order.

- In ZFC, it is a total order.
- In fact, in ZF it need not be a total order.

Cardinal Arithmetic

Define $m + n$ to be $\text{card}(M \sqcup N)$. (Clearly well-defined)

and mn to be $\text{card}(M \times N)$

and m^n to be $\text{card}(M^N)$, where $M^N = \{f : f \text{ a function from } N \text{ to } M\}$

with M, N any sets with $\text{card } M = m$, $\text{card } N = n$.

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

2. How many sequences of reals are there?

Want $\text{card}(\mathbb{R}^\omega) = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \aleph_0} = 2^{\aleph_0}$. (Using (A) and (D) below.)

Note. Used obvious facts like

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

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

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

(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: $\aleph_0^{\aleph_0} \geq 2^{\aleph_0} > \aleph_0$, so uncountable.

We know $\aleph_0 \aleph_0 = \aleph_0$. What about $\aleph_1 \aleph_1$?

+ and \cdot are easy on the alephs, thanks to:

Theorem 1. $\aleph_\alpha \aleph_\alpha = \aleph_\alpha$.

Proof. Induction on α .

Define a well-ordering of $\omega_\alpha \times \omega_\alpha$ by: $(x, y) < (z, t)$ if:

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

‘Go up in squares.’

Consider an initial segment I_δ . Then $I_\delta \subset \beta \times \beta$, some $\beta < \omega_\alpha$.

But $\text{card } \beta < \text{card } \omega_\alpha$ (as ω_α initial), so $\beta \times \beta \leftrightarrow \beta$ (induction hypothesis),

so $\text{card } I_\delta \leq \text{card}(\beta \times \beta) = \text{card } \beta < \text{card } \omega_\alpha$.

Thus **every** proper initial segment has order-type $< \omega_\alpha$, whence order-type $\leq \omega_\alpha$.

Thus $\omega_\alpha \times \omega_\alpha$ injects into ω_α , so $\aleph_\alpha \aleph_\alpha \leq \aleph_\alpha$.

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

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

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

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

Remark. 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 $2^{\aleph_0} = \aleph_1$. (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: \exists sentence p such that $PA \not\vdash p$ and $PA \not\vdash \neg p$.

Equivalently: \exists sentence p , **true in \mathbb{N}** , such that $PA \not\vdash p$.

‘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 such that p true $\Leftrightarrow p$ not provable.

[Then done: if false then $PA \vdash p$, whence p holds in every model of PA, and in particular p holds in \mathbb{N} \times . 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 ‘ p not provable’ **has** to be longer than p !

Recall. $S \subset \mathbb{N}$ is *definable* or *definable in language of PA* if \exists formula $p(x)$ (in language of PA, x a free variable (and nothing else)) such that $\forall m \in \mathbb{N} : m \in S \Leftrightarrow p[m/x]$ holds in \mathbb{N} .

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: $p(x) : (\forall y) (\underbrace{(y|x)}_{(\exists z)(yz=x)} \wedge y \text{ prime}) \Rightarrow y = 2$.

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

E.g. $f(n) = 2^n$ definable, meaning:

\exists formula $p(x, y)$ such that $\forall m, n \in \mathbb{N} : m = 2^n \Leftrightarrow p[n/x, m/y]$ holds.

Coding

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

Assign to each a distinct value:

$$v(0) = 1, v(s) = 2, v(+)= 3, \dots, v(') = 11, v(=) = 12$$

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 $2^9 5^7$.

‘ m codes a formula’ is definable (as \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, \dots, p_n of formulae by:

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

Observe that ‘ m codes an axiom’ is definable. (Has to be checked, but easy.)

Similarly, ‘ l, m, n are such that 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 $\theta(m) =$ ‘ m codes a provable statement’ (i.e. ‘ S_m is provable’) is definable, as $\theta(m) \Leftrightarrow (\exists n) \theta(m, n)$.

Clever bit:

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

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

Thus the sentence $\phi(N)$ will do!

We have shown:

Theorem 1. PA incomplete.

Could we make PA complete by adding a sentence p (true in \mathbb{N}) to it?

Answer: no. If $PA' = PA \cup \{p\}$, run same proof.

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

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

Theorem 2. T is not definable.

‘ 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 p from PA?

Answer. We used existence of a model of PA (\mathbb{N}), i.e. used $\text{con}(\text{PA}) =$ ‘PA is consistent’ = ‘ $\neg \exists$ proof of ‘ \perp ’.’

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

Theorem 3. $\text{PA} \not\vdash \text{con}(\text{PA})$.

Does $\text{ZF} \vdash \text{con}(\text{PA})$? $((\forall x \in \omega)(x \text{ does not code a proof of } \perp).)$

Yes: as $\text{ZF} \vdash$ ‘PA has a model’ (namely ω).

However, running Theorem 1 of ZF language, get

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

As for Theorem 1 \Rightarrow Theorem 3, get

Theorem 5. If ZF consistent, then $\text{ZF} \not\vdash \text{con}(\text{ZF})$.

Notes typeset in L^AT_EX by Gareth Taylor

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

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 $\perp \Rightarrow q$. 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, \dots 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 $S - \{s\} \not\vdash s$. 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?

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 $\omega^{\alpha_1}n_1 + \dots + \omega^{\alpha_k}n_k$, where $\alpha_1 > \dots > \alpha_k$ are ordinals (and k and n_1, \dots, 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 $\alpha_1 < \alpha_2 < \alpha_3 < \dots$ 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. Define a topological space $[0, \omega_1)$ on the set of countable ordinals by taking as a base of open sets the intervals $(\alpha, \beta) = \{\gamma : \alpha < \gamma < \beta\}$. Prove that $[0, \omega_1)$ is sequentially compact (i.e. every sequence has a convergent subsequence), but not compact. Is there a metric giving this topology?
14. 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?
- +15. 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 $f : X \rightarrow X$ 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. Write down axioms (in the language of partial orders) for the theory of total orders that are dense (between any two elements is a third) and have no greatest or least element. Show that every countable model of this theory is isomorphic to \mathbb{Q} .
11. Is every countable model of Peano Arithmetic isomorphic to \mathbb{N} ?
12. 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.
13. 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)
- +14. Let \mathbb{A} denote the subfield of \mathbb{C} consisting of the algebraic numbers. If a sentence (in the language of fields) holds in \mathbb{C} , must it also hold in \mathbb{A} ?

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. 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 .
5. 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 $\in |_{HF}$)?
6. Which of the axioms of ZF are satisfied in the structure $V_{\omega+\omega}$?
7. What is the cardinality of the set of all continuous functions from \mathbb{R} to \mathbb{R} ?
8. Is there an ordinal α such that $\omega_\alpha = \alpha$?
9. 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_ω to $\aleph_\omega^{\aleph_0}$, and deduce that $2^{\aleph_0} \neq \aleph_\omega$.
10. Suppose that no uncountable subset of \mathbb{R} can be well-ordered. Prove that there exist sets x and y such that there is an injection from x to y and a surjection from x to y but no bijection from x to y .
11. Prove that a countable union of countable sets cannot have cardinality \aleph_2 .
12. 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?
13. Assume that ZF is consistent. We extend the language of ZF by adding uncountably many new constants, and extend the axioms of ZF by adding the assertions that these constants are distinct and all belong to ω . Explain why this theory has a model. In this model of ZF, ω is uncountable – doesn't this contradict the fact that ω is countable?
- +14. 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 $n = 2^m$.