Contents

Paul Russell

Prerequisites none in particular

Books 1. Johnstone } whole course
2. Forster }
3. Hajnal + Hamburger : useful for part I

Details of all books in schedules (+ a couple others)

Chapter 1 Propositional Logic

1.1 Propositional Language

Let P be a set of primitive propositions; that is, P is a set such that

$$\bot, (,), \Rightarrow, \notin, P$$

$\bot$: 'false'

「symbols not in P」

Often $P = \{p_1, p_2, \dots\}$,

or $P = \{p, q, r, \dots\}$.

N.B. Can allow P to be uncountable.

The set of all propositions $L = L(P)$ is defined inductively as follows:

- $P \subset L$
- $\bot \in L$
- if $p, q \in L$ then $(p \Rightarrow q) \in L$

Examples (using $P = \{p, q, r\}$)

$$p,\ (p \Rightarrow q),\ \bot,\ (\bot \Rightarrow \bot),\ (\bot \Rightarrow p),\ (p \Rightarrow \bot),$$
$$((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) \Rightarrow (\bot \Rightarrow (q \Rightarrow \bot)))\quad (*)$$

Remarks 1. A proposition is a finite string of symbols taken from the alphabet $P \cup \{(,), \bot, \Rightarrow\}$

(Not any such string is a proposition : e.g. $\Rightarrow) \bot p \notin L$)

2. When we write down propositions, often we omit brackets or write them in different styles.

e.g. might write $*$ as

$$[(p \Rightarrow q) \Rightarrow (p \Rightarrow r)] \Rightarrow [\bot \Rightarrow (q \Rightarrow \bot)]$$

3. What exactly does 'L defined inductively' mean ?

Define $L_0 = P \cup \{\bot\}$.

Then given L_n, define $L_{n+1} = L_n \cup \{(p \Rightarrow q) \mid p, q \in L_n\}$

Then set $L = \bigcup_{n=0}^{\infty} L_n$.

4. Can define various abbreviations:

- 'not p' $\neg p$ means $(p \Rightarrow \bot)$
- 'p or q' $p \vee q$ means $(\neg p \Rightarrow q)$
- 'p and q' $p \wedge q$ means $\neg(p \Rightarrow (\neg q))$

1.2 Semantic Entailment (validity)

Def A valuation v is a function $v : L \longrightarrow \{0,1\}$ satisfying

- $v(\bot) = 0$; and
- for any $p, q \in L$, $v(p \Rightarrow q) = \begin{cases} 0 \text{ if } v(p)=1, v(q)=0 \\ 1 \text{ otherwise} \end{cases}$

Proposition 1 Let $f : P \to \{0,1\}$ be a function.

Then there is a unique valuation $v : L \longrightarrow \{0,1\}$ such that $v|_P = f$.

Proof Existence We define v inductively on $L_0, L_1, \ldots$

L_0 : set $v(p) = f(p)$ for $p \in P$, and $v(\bot) = 0$

L_n $(n \geq 1)$: suppose v already defined on L_{n-1}

Let $p \in L_n \setminus L_{n-1}$. Then $p = (q \Rightarrow r)$

for some $q, r \in L_{n-1}$.

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

Uniqueness Suppose v' is also a valuation on L with $v'|_P = f$. We show by induction on n, that for all n, v agrees with v' on L_n.

L_0: If $p \in P$ then $v(p) = f(p) = v'(p)$,
and also $v(\bot) = 0 = v'(\bot)$.

L_n $(n \geq 1)$: By ind hyp, know $v|_{L_{n-1}} = v'|_{L_{n-1}}$.

Suppose $p \in L_n$. Then either $p \in L_{n-1}$ so $v(p) = v'(p)$;
or $p \in L_n \setminus L_{n-1}$, i.e.

$p = (q \Rightarrow r)$ for some $q, r \in L_{n-1}$

We know $v(q) = v'(q)$, $v(r) = v'(r)$.

These determine $v(p)$, $v'(p)$ and we are done. □

Definition Suppose $S \subset L$ and that $p \in L$.

We say S semantically entails p if whenever v is a valuation with $v(q) = 1 \ \forall q \in S$, we also have $v(p) = 1$.

We write $S \vDash p$.

If $S = \{q\}$ we write $q \vDash p$.

If $S = \emptyset$ we write $\vDash p$. In this case, we say p is a tautology.

Definition If $v(p) = 1$ (v valuation, $p \in L$), we say p is true in v or v is a model of p.

Remarks 1. $S \vDash p$ says every model of S is a model of p.

2. p is a tautology if every valuation is a model of p.

Example $\{p, (p \Rightarrow q)\} \vdash q$

Indeed, suppose v is a valuation with $v(p) = 1$ and $v(p \Rightarrow q) = 1$. If $v(q) = 0$ and $v(p) = 1$, then by def^n, $v(p \Rightarrow q) = 0$ ✱

Hence $v(q) = 1$. ✓

Examples of tautologies

0. $\vdash (p \Rightarrow p)$ Truth table

p	$p \Rightarrow p$
0	1
1	1

} all 1s so tautology

1. $\vdash p \Rightarrow (q \Rightarrow p)$

p	q	$q \Rightarrow p$	$p \Rightarrow (q \Rightarrow p)$
0	0	1	1
0	1	0	1
1	0	1	1
1	1	1	1

} all 1s so tautology

2. $\vdash (p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$

Suppose not. Then there exists some valuation v with

$$v((p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))) = 0$$

Thus $v(p \Rightarrow (q \Rightarrow r)) = 1$ and

$v((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) = 0$.

Then $v(p \Rightarrow q) = 1$ and $v(p \Rightarrow r) = 0$.

So $v(p) = 1$ and $v(r) = 0$.

So also $v(q) = 1$.

Then $v(q \Rightarrow r) = 0$, so $v(p \Rightarrow (q \Rightarrow r)) = 0$. ✱

3. $\vdash ((p \Rightarrow \bot) \Rightarrow \bot) \Rightarrow p$

p	$p \Rightarrow \bot$	$(p \Rightarrow \bot) \Rightarrow \bot$	$((p \Rightarrow \bot) \Rightarrow \bot) \Rightarrow p$
0	1	0	1
1	0	1	1

[i.e. $\vdash \neg\neg p \Rightarrow p$]

1.3 Syntactic entailment (Proof)

The axioms are as follows:

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

Remark Every axiom is a tautology.

The deduction rule is called modus ponens:

"from p and $p \Rightarrow q$ we can deduce q"

Def^n Suppose $S \subseteq L$ and $p \in L$.

A proof of p from S is a sequence $l_1, l_2, \ldots, l_n \in L$ (with $l_n = p$) such that for all $i = 1, 2, \ldots, n$ we have:

- l_i is an axiom; or
- l_i is a hypothesis, i.e. $l_i \in S$; or
- l_i follows from earlier lines of the proof using modus ponens, i.e. $\exists\, j, k < i$ s.t. $l_k = (l_j \Rightarrow l_i)$

If there exists a proof of p from S we say S syntactically entails p or S proves p, and we write $S \vdash p$. If $S = \{q\}$, we write $q \vdash p$.

If $S = \emptyset$ we say p is a theorem and write $\vdash p$.

Examples 1.] Any axiom is a theorem. If a is an axiom, then a is a proof of a.

2.] $\{p, p \Rightarrow q\} \vdash q$

Write down a proof of q from $\{p, p \Rightarrow q\}$.

1. p (hyp)
2. $p \Rightarrow q$ (hyp)
3. q (modus ponens)

↖ formal proof

3.] $\vdash (p \Rightarrow p)$

ROUGH WORK
nah fam

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

Proposition 2 For all $p \in L$, $\vdash (p \Rightarrow p)$.

Proof Wwrite down the 5-line proof above. □

Mnemonic Theorem $\vdash$; TauTology $\vdash$

Proposition 3 (Deduction Theorem)

Let $S \subseteq L$, $p, q \in L$.

Then $S \vdash (p \Rightarrow q)$ iff $S \cup \{p\} \vdash q$.

Proof "$\Rightarrow$" Suppose we have a proof of $(p \Rightarrow q)$ from S.

We need to write down a proof of q from $S \cup \{p\}$.

Write down the proof of $(p \Rightarrow q)$ from S.

Append: p (hyp), q (modus ponens p, $p \Rightarrow q$)

L2.4

"<=" Suppose $l_1, l_2, \ldots, l_n$ is a proof of q from $S \cup \{p\}$. Need a proof of $(p \Rightarrow q)$ from S.

We show by induction on i that $S \vdash (p \Rightarrow l_i)$ for each $i = 1, 2, \ldots, n$. Then done, as $l_n = q$.

Case (i) l_i is an axiom

Write down: l_i (Axiom)

$l_i \Rightarrow (p \Rightarrow l_i)$ (Axiom 1)

$p \Rightarrow l_i$ (MP)

Case (ii) $l_i \in S$

Exactly same as case (i).

(Justify for l_i as hypothesis)

Case (iii) $l_i = p$

By Prop. 2, $\vdash (p \Rightarrow p)$.

So $S \vdash (p \Rightarrow p)$.

Case (iv) l_i deduced by MP

Then $\exists\, j, k < i$ with $l_k = (l_j \Rightarrow l_i)$.

By ind. hyp., $S \vdash (p \Rightarrow l_j)$ and $S \vdash (p \Rightarrow l_k)$.

Append: $(p \Rightarrow (l_j \Rightarrow l_i)) \Rightarrow ((p \Rightarrow l_j) \Rightarrow (p \Rightarrow l_i))$ (Ax2)

$(p \Rightarrow l_j) \Rightarrow (p \Rightarrow l_i)$ (MP)

$(p \Rightarrow l_i)$ (MP)

□

↑ this is why Axiom 2 is like that

1.4 Completeness

MAIN POINT $\models$ and $\vdash$ turn out to be the same

Easier direction If $S \vdash p$ then $S \models p$.

Proposition 4 (Soundness Theorem)

Let $S \subseteq L$ and $p \in L$ with $S \vdash p$. Then $S \models p$.

Proof Let $l_1, l_2, \ldots, l_n$ be a proof of p from S.

Let v be a model of S. We shall prove $v(p) = 1$.

We know $l_n = p$. So we'll show by induction on i that $v(l_i) = 1$ for $i = 1, 2, \ldots, n$.

Case (i) l_i is an axiom, we know $\models l_i$ so $v(l_i) = 1$.

Case (ii) l_i is a hypothesis, i.e. $l_i \in S$, so $v(l_i) = 1$.

Case (iii) l_i was deduced using modus ponens.

Then $\exists j, k < i$ s.t. $l_k = (l_j \Rightarrow l_i)$.

By ind hyp, $v(l_j) = v(l_k) = 1$.

So $v(l_j \Rightarrow l_i) = 1$, so $v(l_i) = 1$. □

Harder direction Want if $S \models p$ then $S \vdash p$.

Special case: if $S \models \bot$ then $S \vdash \bot$.

(In fact, all of the work is done in proving this special case)

Equivalently: if $S \nvdash \bot$ then $S \not\models \bot$

Definition If $S \nvdash \bot$, we say S is consistent.

(If $S \vdash \bot$, we say S is inconsistent)

What about '$S \not\models \bot$'? Well '$S \models \bot$' means any model of S is a model of $\bot$, i.e. S has no model.

So '$S \not\models \bot$' means 'S has a model'.

So AIM: If S is consistent then S has a model.

Idea Definitely $v(p) = \begin{cases} 1 & \text{if } p \in S, \\ 0 & \text{if } \neg p \in S. \end{cases}$

Nice: S is consistent so never get $p, \neg p \in S$.

Problem: Could sometimes have $p, \neg p \notin S$

Then don't know what to do.

Maybe wait $v(p) = 1$, $v(\neg p) = 0$?

or $v(p) = 0$, $v(\neg p) = 1$?

Would be fine if for all $p \in L$ had $p \in S$ or $\neg p \in S$.

Solution: Try to make S bigger, keeping it consistent, so that this good thing happens.

Lemma 5 Let $S \subset L$ be consistent. Then there exists a consistent $\overline{S} \subset L$ with $S \subset \overline{S}$ s.t. $\forall p \in L$, either $p \in \overline{S}$ or $\neg p \in \overline{S}$.

Proof (N.B. in case P countable only - full proof later)

Each $p \in L$ is a finite string of symbols from the countable alphabet $P \cup \{(,), \Rightarrow, \bot\}$.

Hence L is countable. Write $L = \{p_1, p_2, p_3, \ldots\}$.

We define inductively an increasing sqce $S_0 \subset S_1 \subset \ldots$ of consistent subsets of L with $S \subset S_n$ for all n, and if $n \geq 1$, $p_n \in S_n$ or $\neg p_n \in S_n$.

$n = 0$ Set $S_0 = S$.

$n \geq 1$ If $S_{n-1} \cup \{p_n\}$ is consistent, then set $S_n = S_{n-1} \cup \{p_n\}$.

If $S_{n-1} \cup \{p_n\}$ is not consistent, i.e. $S_{n-1} \cup \{p_n\} \vdash \bot$, then by the deduction theorem $S_{n-1} \vdash (p_n \Rightarrow \bot)$

i.e. $S_{n-1} \vdash (\neg p_n)$

But S_{n-1} is consistent, so $S_{n-1} \cup \{\neg p_n\}$ is consistent.

So take $S_n = S_{n-1} \cup \{\neg p_n\}$.

Now define $\overline{S} = \bigcup_{n=0}^{\infty} S_n$.

First, $S = S_0 \subseteq \overline{S}$. And by construction, for all $p \in L$, either $p \in \overline{S}$ or $\neg p \in \overline{S}$.

Is $\overline{S}$ consistent? If not then there is a proof of $\bot$ from $\overline{S}$. Proofs are finite so this proof only uses finitely many hypotheses, from $S_{n_1}, S_{n_2}, \ldots, S_{n_k}$, say.

Let $n = \max\{n_1, \ldots, n_k\}$.

Then by nestedness, all the hypotheses of the proof lie in S_n, i.e. $S_n \vdash \bot$ ✱ to S_n consistent □

Definition We say S is deductively closed (d.c.) if whenever $p \in L$ and $S \vdash p$ we have $p \in S$

Theorem 6 (Model existence lemma)

Let $S \subseteq L$ be consistent. Then S has a model.

Proof First, define $\overline{S}$ as in Lemma 5.

Observe that $\overline{S}$ is deductively closed.

Indeed, suppose $\overline{S} \vdash p$. Then, as $\overline{S}$ is consistent, $\neg p \notin \overline{S}$. But then by construction of $\overline{S}$, $p \in \overline{S}$ ✓

Define $v : L \to \{0,1\}$ to be the indicator f^n of $\overline{S}$.

Note $\begin{cases} v(p) = 1 \text{ iff } p \in \overline{S} \\ v(p) = 0 \text{ iff } \neg p \in \overline{S} \end{cases}$

As $S \subseteq \overline{S}$, so, as long as v is a valuation, v is a model of S. As $\overline{S}$ is consistent, $\bot \notin \overline{S}$, so $v(\bot) = 0$. ✓

Still need to check if $p, q \in L$ then $v(p \Rightarrow q) = 0$ iff $v(p) = 1, v(q) = 0$.

Case (i) $v(q) = 1$. Then $q \in \overline{S}$. So $\overline{S} \vdash q$.

Hence $\overline{S} \cup \{p\} \vdash q$. By deduction thm, $\overline{S} \vdash (p \Rightarrow q)$.

So $(p \Rightarrow q) \in \overline{S}$ and $v(p \Rightarrow q) = 1$. ✓

Case (ii) $v(p) = 0$. Then $\neg p \in \overline{S}$, i.e. $(p \Rightarrow \bot) \in \overline{S}$ and so $\overline{S} \vdash (p \Rightarrow \bot)$.

Hence by deduction thm, $\overline{S} \cup \{p\} \vdash \bot$.

But $\bot \vdash q$:

1.	$\bot \Rightarrow ((q \Rightarrow \bot) \Rightarrow \bot)$	(Axiom 1)
2.	$\bot$	(hyp)
3.	$(q \Rightarrow \bot) \Rightarrow \bot$	(MP)
4.	$(\neg\neg q) \Rightarrow q$	(Ax 3)
5.	q	(MP)

Thus $\overline{S} \cup \{p\} \vdash q$ and so by deduction thm

$\overline{S} \vdash (p \Rightarrow q)$, so $v(p \Rightarrow q) = 1$. ✓

Case (iii) $v(p) = 1$, $v(q) = 0$

Want $v(p \Rightarrow q) = 0$.

We know $p \in \overline{S}$, $q \notin \overline{S}$. Suppose for a contradiction that $v(p \Rightarrow q) = 1$, i.e. $(p \Rightarrow q) \in \overline{S}$.

Now $p \in \overline{S}$ and $(p \Rightarrow q) \in \overline{S}$ so by M.P. $\overline{S} \vdash q$.

So $q \in \overline{S}$ ⨳. □

Corollary 7 (Adequacy) Let $S \subset L$, $p \in L$.

Then if $S \models p$, then $S \vdash p$.

Pf Now $S \cup \{\neg p\} \models \bot$.

By Theorem 6, $S \cup \{\neg p\} \vdash \bot$.

So by deduction theorem, $S \vdash \neg\neg p$.

So by Axiom 3 and MP, $S \vdash p$. □

Theorem 8 (Completeness Theorem) Let $S \subset L$, $p \in L$.

Then $S \vdash p \iff S \models p$.

Proof "$\Rightarrow$" Soundness, "$\Leftarrow$" Adequacy. □

1.5 Applications of Completeness

Recall the completeness theorem: $S \vdash p$ iff $S \models p$

Two major applications:

Corollary 9 (Compactness Theorem)

Let $S \subset L$ and suppose every finite subset of S has a model. Then S has a model.

Proof Trivial if 'has a model' is replaced by 'is consistent'.

Indeed if $S \vdash \bot$ then since proofs are finite, $T \vdash \bot$ for some finite $T \subset S$. □

Corollary 10 (Decidability Theorem)

There exists an algorithm that, given finite $S \subset L$ and $p \in L$, determines in finite time whether or not $S \vdash p$.

Proof Replace $\vdash$ by $\models$.

S finite so do a truth table. □

1.6 Completeness when P is uncountable

Informal, motivation — all done rigorously later.

Our proof of the Completeness Theorem is incomplete.

Only hole is in proof of Lemma 5 where we extend S to $\bar{S}$.

Recall If P ctble, then L ctble. Write $L = \{p_1, p_2, \dots\}$.

Deal with $p_1, p_2, \dots$ in turn to get $S = S_0 \subset S_1 \subset \cdots$ and set $\bar{S} = \bigcup_n S_n$.

Suppose instead we'd written $L = \{p_1, p_2, \dots\} \cup \{q_1, q_2, \dots\}$.

Try the same as before: deal with $p_1, p_2, \dots$ in turn, get $S = S_0 \subset S_1 \subset \cdots$ and set $S_\omega = \bigcup_{n=0}^{\infty} S_n$.

Now repeat the whole process starting from S_ω.

Deal with $q_1, q_2, \dots$ in turn to get $S_\omega \subset S_{\omega+1} \subset \cdots$

Finally set $S = \bigcup_{n=0}^{\infty} S_{\omega+n}$

Idea: could define $\mathbb{N}$ inductively by

- $0 \in \mathbb{N}$
- if $n \in \mathbb{N}$ then $\exists n^+ \in \mathbb{N}$, $n^+ > n$

(And that's it) What does $\mathbb{N}$ look like?

$0, \underset{0^+}{1}, \underset{1^+}{2}, \underset{2^+}{3}, \dots$

Generalize to the Ordinals, defined inductively by

- 0 is an ordinal;
- if n is an ordinal, then n^+ is an ordinal, $n^+ > n$; (n^+: successor ordinals)
- if X is a set of ordinals with no greatest elt, then $\sup X$ is an ordinal with $\sup X > n$ for all $n \in X$ ($\sup X$: non-zero limit ordinals)

(And that's it)

What are the ordinals? First 'few':

$0, \underset{0^+}{1}, \underset{1^+}{2}, \underset{2^+}{3}, \dots, \underset{\sup\{0,1,\dots\}}{\omega}, \underset{\omega^+}{\omega+1}, \underset{(\omega+1)^+}{\omega+2}, \dots, \underset{\sup\{\omega,\omega+1,\dots\}}{\omega 2}, \dots$

$\omega 2+1, \omega 2+2, \dots, \omega 3, \dots, \omega 4, \dots, \dots,$

$\underset{\sup\{\omega,\omega 2,\omega 3,\dots\}}{\omega^2}, \omega^2+1, \dots, \omega^2+\omega, \dots, \omega^2+\omega 2, \dots, \dots, \omega^2 2,$

$\dots \omega^2 3, \dots, \dots, \omega^3, \dots, \omega^4, \dots, \omega^5, \dots,$

$\underset{\sup\{\omega,\omega^2,\omega^3,\dots\}}{\omega^\omega}, \dots, \omega^{\omega^\omega}, \dots, \omega^{\omega^{\omega^\omega}}, \dots, \underset{\sup\{\omega,\omega^\omega,\dots\}}{\omega^{\omega^{\cdots}}}, \dots$ ← ε_0

$\varepsilon_0 + 1, \varepsilon_0 + 2, \dots$

Looks like lots of ordinals, but only countably many appear.

In fact (!) there are uncountably ordinals. Indeed, there are more ordinals than there are things in any set. (*)

What about Lemma 5 in general case?

Try to 'list' L : $p_0, p_1, p_2, \dots, p_\omega, p_{\omega+1}, \dots, p_{\omega 2}, p_{\omega 2+1}, \dots$ indexed by the first ordinals. By (*), we eventually run out of things in L (before we run out of ordinals)

Now try to write a proof as before:

$$S = S_0 \subset S_1 \subset \dots \subset S_\omega \subset \dots \subset S_{\omega 2} \subset \dots \dots$$

Take $\overline{S}$ to be union of all of these things.

This works.

Ch 2: Can we say properly set up definition of ordinals? How do they behave?

Ch 3: Looking at these ideas to make things like the proof above rigorous.

CHAPTER 2 ORDINALS

2.1 Total orders and well-orderings

Let X be a set. A <u>total order</u> of X is a relation $<$ on X satisfying:

(i) $<$ is <u>irreflexive</u>, i.e. for all $x \in X$, $x \not< x$;

(ii) $<$ is <u>transitive</u>, i.e. for all $x, y, z \in X$,
if $x < y$ and $y < z$ then $x < z$;

(iii) $<$ is <u>trichotomous</u>, i.e. for all $x, y \in X$,
$x < y$ or $x = y$ or $y < x$

<u>Remark</u> Given $x, y \in X$, <u>exactly</u> one of the three possibilities in (iii) holds. Indeed, if $x = y$ then $x \not< y$ and $y \not< x$ by (i).

While if $x < y$ and $y < x$ then $x < x$ by (ii) $\not\Rightarrow$ to (i)

Notation Define $x > y$ to mean $y < x$.

Define $x \leq y$ to mean $x < y$ or $x = y$.

Sometimes $<$ is called a strict total order and $\leq$ the corresponding weak total order.

What conditions does $\leq$ satisfy?

(i) for all $x \in X$, $x \leq x$;

(ii) for all $x, y, z \in X$, if $x \leq y$ and $y \leq z$ then $x \leq z$

(iii) if $x, y \in X$ with $x \leq y$ and $y \leq x$ then $x = y$

(iv) if $x, y \in X$ then $x \leq y$ or $y \leq x$

Instead of our original approach, we could have defined a weak total order to be a relation $\leq$ satisfying conditions (i), (ii), (iii), (iv) immediately above.

Then could define $<$ in terms of $\leq$ by $x < y$ if $x \leq y$ and $x \neq y$. And $<$ turns out to be a strict total order.

(Exercise: check all this)

Notation Write $x \geq y$ to mean $y \leq x$.

Examples of total orders The usual order $<$ on $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N}$. **Definition** A total order on a set X is a well-ordering if every non-empty subset of X has a least element.

i.e. if $S \subset X$ with $S \neq \emptyset$, then there is some $m \in S$ such that for all $x \in S$, $m \leq x$.

Examples

- $<$ on $\mathbb{N}$ is a well-ordering
- $<$ on $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ is not a well-ordering:
  $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ respectively has no least element

Last time Usual $<$ on $\mathbb{N}$ is well-ordering.

But usual $<$ on $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ is not, in each case, the whole set has no least element.

What about usual $<$ on $X = \{x \in \mathbb{Q} \mid x \geq 0\}$?

Now X does have a least element, namely zero.

But still not well-ordering, e.g. $\{x \in X \mid x > 0\}$ has no least element.

Examples of well-orderings Usual $<$ on the following $X_i \in \mathbb{R}$.

(i) $X_1 = \mathbb{N}$

(ii) $X_2 = \{1 - \frac{1}{n} \mid n \in \mathbb{N}^+\}$ $(\mathbb{N}^+ = \{1, 2, 3, \dots\})$

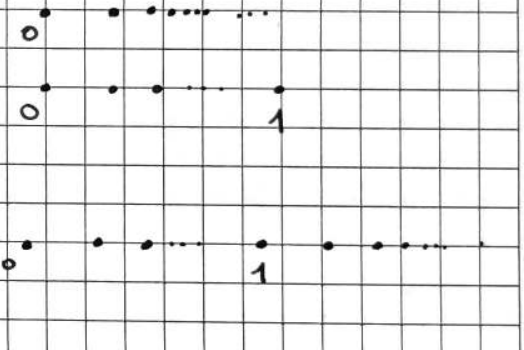
$= \{0, \frac{1}{2}, \frac{3}{4}, \dots\}$

(iii) $X_3 = X_2 \cup \{1\}$

(iv) $X_4 = X_2 \cup \{2\}$

(v) $X_5 = X_2 \cup \{2 - \frac{1}{n} \mid n \in \mathbb{N}^+\}$

(cf $0, 1, 2, \dots, \omega, \omega+1, \omega+2, \dots$)



Definition Let X, Y be sets with total orders $<_X$, $<_Y$ resp. An isomorphism (of totally ordered sets) from X to Y is a bijection $f: X \to Y$ s.t. $\forall x, y \in X$, $f(x) <_Y f(y)$ iff $x <_X y$.

If such exists, say X, Y isomorphic and write $X \cong Y$.

Examples $X_1 \cong X_2$, $X_3 \cong X_4$, no other pair in our examples are isomorphic.

Definition An initial segment of a totally ordered set is a subset $I \subseteq X$ such that whenever $x \in I$ and $y \in X$ with $y < x$ then $y \in I$.

If $x \in X$ then $X_x = \{y \in X \mid y < x\}$ (*) is an initial segment.

Suppose now X is well-ordered and $I \subsetneq X$ is a proper initial segment: then $I = X_x$ for some $x \in X$.

Take $x = \min(X \setminus I)$. So every proper initial segment has the form (*)

Not true for general total orders, e.g.

$I = \{x \in \mathbb{Q} \mid x \leq 0\} \subsetneq \mathbb{Q}$ is a proper initial segment, but $I \neq \mathbb{Q}_x$ for any $x \in \mathbb{Q}$.

Examples • X_2 is an initial segment of X_3, X_4 and X_5

• X_3 is an initial segment of X_5

• X_4 is not an initial segment of X_5, but is (canonically) isomorphic to an initial segment of X_5, namely $(X_5)_{\frac{1}{2}}$

2.2 Well-ordered Induction and Recursion

In $\mathbb{N}$ we have: • proof by induction

• definition by recursion

(e.g. define the Fibonacci numbers recursively by

$F_0 = 1$, $F_1 = 1$, and $\forall n \geq 2$, $F_n = F_{n-1} + F_{n-2}$)

Aim Generalize this to arbitrary well-orderings

Theorem 11 (Well-ordered induction) Let X be a set well-ordered by $<$. Let $Y \subseteq X$ with the following property:

for all $x \in X$, if for all $y \in X$ with $y < x$ we have $y \in Y$ then $x \in X$

Then $Y = X$.

Proof Suppose $Y \neq X$. Then $X \setminus Y \neq \emptyset$, so it has a least element, x say. Then for all $y < x$, $y \in Y$.

But $x \notin Y$. ⚡ □

This is used for proof by induction: if P is some property that an element of a well-ordered set X may or may not have. Suppose we want to prove that for all $x \in X$ we have $P(x)$. Then when proving $P(x)$ for a particular

x we may assume $P(y)$ for all $y < x$.

An example of proof by induction.

Theorem 12 Let X be well-ordered, let $I \subset X$ be an initial segment and suppose $f: X \to I$ is an isomorphism. Then $I = X$ and f is the identity.

Proof We show by induction that for all $x \in X$, $f(x) = x$.

Let $x \in X$. By ind hyp, we know $f(y) = y$ for all $y < x$.

Suppose $f(x) \neq x$.

Then, as f is injective, $f(x) > x$.

But f is surjective onto an initial segment, $x \in I$ is s.t. for some $z \in X$, $x = f(z)$.

But $z > x$, $f(z) < f(x)$. ⚡ □

Theorem 13 (Well-ordered recursion)

Let X, Y be sets with X well-ordered.

Let $G: \mathcal{P}(X \times Y) \to Y$.

Then there exists a unique function $f: X \to Y$ s.t. for all $x \in X$, $f(x) = G(f|_{X_x})$.

Remark Recall that we can think of a function $f: X \to Y$ as the set $\{(x, f(x)) \mid x \in X\}$ of ordered pairs.

So $f \subset X \times Y$. In other words, $f \in \mathcal{P}(X \times Y)$.

Now $f|_{X_x} : X_x \to Y$ so $f|_{X_x} \subset X_x \times Y \subset X \times Y$ so again $f|_{X_x} \in \mathcal{P}(X \times Y)$. So $G(f|_{X_x})$ makes sense.

Proof Define 'h is an attempt' to mean $h: I \to Y$ where I is an initial segment of X s.t. for all $x \in I$, $h(x) = G(h|_{I_x})$.

Claim 1 If h, h' are attempts both defined at $x \in X$ then $h(x) = h'(x)$.

Proof By induction on x.

Recall $\mathrm{dom}(h)$, $\mathrm{dom}(h')$ are initial segments.

So if $y < x$ then h, h' are defined at y.

And so by ind hyp $h(y) = h'(y)$.

Then $h(x) = G(h|_{I_x}) = G(h'|_{I_x}) = h'(x)$. //

Claim 2 For all $x \in X$, there exists an attempt h_x defined at x.

Proof Induction on x. Fix $x \in X$.

By ind hyp, for each $y < x$ there is an attempt h_y defined at y.

Moreover, by Claim 1, these attempts agree wherever they are defined at the same point.

Let $h' = \bigcup_{y \in X_x} h_y$. Then by previous remark, h' is a well-defined function and indeed is an attempt.

If $x \in \mathrm{dom}(h')$ set $h_x = h'$.

If not, $\mathrm{dom}(h') = X_x$; set $h_x = h' \cup \{(x, G(h'))\}$ //

Now set $f = \bigcup_{x \in X} h_x$. By Claim 1, this is a well-defined function, and it satisfies the required properties.

Finally, f itself is an attempt. So Claim 1 tells us it is unique. □

Example of recursion:

Def^n Let X, Y be well-ordered sets. Write $X \leq Y$ to mean X is isomorphic to an initial segment of Y.

Theorem 14 Let X, Y be well-ordered sets. Then $X \leq Y$ or $Y \leq X$. Moreover, if $X \leq Y$ and $Y \leq X$ then $X \cong Y$.

Proof Suppose $Y \not\leq X$. Define $f: X \to Y$ recursively by

$$f(x) = \min\left(Y \setminus \{ f(y) \mid y \in X, y < x \}\right)$$

(noting that this set is non-empty as $Y \not\leq X$)

Then f is an iso from X to an initial segment of Y, so $X \leq Y$.

Now suppose $f: X \to Y$ and $g: Y \to X$ are isomorphisms onto initial segments of Y, X. Then $g \circ f: X \to X$ is an isomorphism of X to an initial segment of itself, so $g \circ f$ is the identity on X.

Similarly $f \circ g$ is the identity on Y. So $X \cong Y$. □

2.3 Ordinals

Def^n An ordinal is a well-ordered set with isomorphic ones considered to be the same.

(c.f. construction of $\mathbb{Q}$ from $\mathbb{Z}$: a rational is an ordered pair (a, b) of integers with $b \neq 0$ where (a, b) is considered the same as (c, d) if $ad = bc$. Can make this more rigorous by setting $(a, b) \sim (c, d)$ iff $ad = bc$. Then $\sim$ is an equivalence relation and can formally define $\mathbb{Q}$ to be the set of equivalence classes.

Doesn't work in this case — no set of all sets.

Formal definition (much) later.)

Any well-ordered set X is isomorphic to a unique ordinal α, the order-type of X.
We write $\alpha = \mathrm{ord}(X)$.
Write $\alpha \in \mathrm{Ord}$ to mean α is an ordinal.
It is easy to check that any set of ordinals is totally ordered by $\leq$ (as defined just before Thm 14)

Theorem 15 Let $\alpha \in \mathrm{Ord}$. Then the collection of all ordinals $\beta < \alpha$ form a well-ordered set I_α.
Moreover $\mathrm{ord}(I_\alpha) = \alpha$.

Proof Each ordinal $\beta < \alpha$ is isomorphic to a proper initial segment of α.
Moreover, by Theorem 12, distinct proper initial segments of α have distinct order types, all $< \alpha$.
Hence $I_\alpha \cong \{ \alpha_x \mid x \in \alpha \} \cong \alpha$. □

Corollary 16 Any non-empty set X of ordinals has a least element.

Proof Let $\alpha \in X$. EITHER α is least elt of X ✓
OR $\emptyset \neq I_\alpha \cap X \subset I_\alpha$. But I_α is well-ordered so $I_\alpha \cap X$ has a least element β. And β must be the least elt of X. □

We seem to have proved that the ordinals are a well-ordered set.
Why do we not just say this?

Theorem 17 (Burali-Forti paradox)
The ordinals do not form a set.

Proof Suppose X is the set of all ordinals.
Then X is well-ordered. Let $\alpha = \mathrm{ord}(X)$.
Now $\alpha \in X$ and $X \cong I_\alpha$, a proper initial segment of X.
↯ □

What do the ordinals look like ?

Definition Let $\alpha \in \mathrm{Ord}$. If α has a greatest elt we say it is a successor ordinal. Otherwise, it is a limit ordinal.

Suppose α is a successor ordinal, with greatest elt x. Let $\beta = \mathrm{ord}(\alpha \setminus \{x\})$. Then $\beta < \alpha$, and α is the least ordinal $> \beta$. Also, $\sup(I\alpha) = \beta$. (*)

Conversely, if β is any ordinal then there is an ordinal $\beta^+ = \mathrm{Ord}(I\beta \cup \{\beta\})$ that is the least ordinal $> \beta$. We call β^+ the successor of β.

'β^+ is β with an extra point added at the end'

If α is a successor ordinal and we take β as in (*) then $\alpha = \beta^+$.

Note that for any ordinal β, β^+ has a greatest element, so must be a successor ordinal rather than a limit ordinal.

Suppose instead α is a limit ordinal. Then $I\alpha$ has no greatest element. So for any $\beta < \alpha$ there is some γ with $\beta < \gamma < \alpha$. So $\sup(I\alpha) = \alpha$.

So we've shown here that $\sup(I\alpha) = \alpha \iff \alpha$ is limit

Proposition 18 Any set X of ordinals has a supremum.

Proof Let $J = \bigcup_{\alpha \in X} I\alpha$ and let $\beta = \mathrm{ord}(J)$.

Each $I\alpha$ is an initial segment of J, so $\alpha \leq \beta$.

Suppose $\forall \alpha \in X$ we have $\alpha \leq \gamma$ where γ is some ordinal.

Then, as the $I\alpha$ are nested, we have $J \leq \gamma$.

So $\beta \leq \gamma$. Thus $\beta = \sup(X)$. $\square$

The least ordinal is $\emptyset$. Define $0 = \emptyset$.

What is the next smallest ordinal? $0^+ = \{*\} = 1$.

" $1^+ = \cdot\,\cdot = 2$

For $n \in \mathbb{N}$, write $n+1 = n^+ \in \mathrm{Ord}$.

So: ⟶ increasing

0 :

1 : $\cdot$

2 : $\cdot\ \cdot$

3 : $\cdot\ \cdot\ \cdot$

4 : $\cdot\ \cdot\ \cdot\ \cdot$

$\vdots$

Let $\omega = \sup(\mathbb{N})$; for $n \in \mathbb{N}$, write $\omega+n+1 = (\omega+n)^+$

So ω : $\cdot\ \cdot\ \cdot \cdots$

$\omega+1$: $\cdot\ \cdot\ \cdot \cdots\ \cdot$

$\omega+2$: $\cdot\ \cdot\ \cdot \cdots\ \cdot\ \cdot$

$\vdots$

Let $\omega 2 = \sup\{\omega+n \mid n \in \mathbb{N}\}$

$\omega 2$: $\cdot\ \cdot \cdots\ \cdot\ \cdot \cdots$

0, ω, $\omega 2$ are limit ordinals

All the rest above are successor ordinals.

Continue: get, via the above construction, the collection of ordinals we found in § 1.6

Examples of well-ordered subsets of $\mathbb{R}$

(i) $\operatorname{ord}(\mathbb{N}) = \omega$ • • • ...

(ii) $\operatorname{ord}(\{\frac{1}{2}, 7, 19\}) = 3$ • • •

(iii) $\operatorname{ord}(\{1 - \frac{1}{n} \mid n \in \mathbb{N}^+\}) = \omega$ • •....

(iv) $\operatorname{ord}(\{1\} \cup \{1 - \frac{1}{n} \mid n \in \mathbb{N}^+\}) = \omega + 1$ •....•

(v) $\operatorname{ord}(\{1 - \frac{1}{n} \mid n \in \mathbb{N}^+\} \cup \{2 - \frac{1}{n} \mid n \in \mathbb{N}^+\}) = \omega 2$ •.... •....

Proposition 19 There exists an uncountable ordinal

(Idea: the ctble ordinals form a set)

Proof Let W be the set of well-orderings of subsets of $\mathbb{N}$.

Let $V = \{\operatorname{ord} R \mid R \in W\}$.

Then V is the set of countable ordinals.

V is a set of ordinals so it is well-ordered by $\leq$.

Let $\alpha = \operatorname{ord}(V) \in \mathrm{Ord}$.

Suppose α is ctble. Then $\alpha \in V$.

But then $V \cong \alpha \cong I\alpha$, a proper initial segment of V.

Hence α is an uncountable ordinal. □ ※

Remark We now know uncountable ordinals exist.

So there must be a least uncountable ordinal.

We denote the least uncountable ordinal by ω_1.

Even more is true:

Theorem 20 (Hartogs' Lemma)

Let X be a set. Then there exists an ordinal α such that there is no injection $\alpha \to X$.

Proof Replicate exactly the proof of Prop 19 but with the set $\mathbb{N}$ replaced by the set X. □

2.4 Ordinal Arithmetic

Addition Let $\alpha, \beta \in Ord$. What is $\alpha + \beta$?

E.g. $\omega = \cdots\cdots$, $3 = \bullet\bullet\bullet$

$\omega + 3 = \cdots\cdots\ \bullet\bullet\bullet$

Def For $\alpha, \beta \in Ord$, let $\alpha + \beta = \leftarrow\alpha\rightarrow\leftarrow\beta\rightarrow$

i.e. the order type of 'α followed by β'

Formally, $\alpha + \beta = Ord(\{0\} \times \alpha, \{1\} \times \beta)$

where we use lexicographic ordering,

$(i, x) < (j, y)$ if either $i < j$

or $i = j$, $x < y$.

Remark This agrees with the usual definition of $+$ when restricted to $\mathbb{N}$.

Proposition 21 $+$ is associative but not commutative

Proof Let $\alpha, \beta, \gamma \in Ord$. Then

$$\begin{aligned}(\alpha+\beta)+\gamma &= \leftarrow\alpha+\beta\rightarrow\leftarrow\gamma\rightarrow \\ &= \leftarrow\alpha\rightarrow\leftarrow\beta\rightarrow\leftarrow\gamma\rightarrow \\ &= \leftarrow\alpha\rightarrow\leftarrow\beta+\gamma\rightarrow \\ &= \alpha+(\beta+\gamma).\end{aligned}$$

However, $1 + \omega = \bullet\ \cdots\cdots = \omega$

$\neq \omega + 1 = \cdots\cdots\ \bullet$ □

Subtraction Suppose $\alpha, \beta \in Ord$ with $\alpha \leq \beta$.

Then let $\gamma = ord(\beta \setminus I)$ where I is the initial segment of β with $\alpha \cong I$.

Then clear by def^n that $\alpha + \gamma = \beta$.

Suppose δ, ε ordinals with $\delta < \varepsilon$. Then δ is (isom to) a proper initial segment of ε. So $\alpha + \delta$ is iso to a proper initial segment of $\alpha + \varepsilon$. So $\alpha + \delta < \alpha + \varepsilon$.

This shows that γ is the unique ordinal with $\alpha + \gamma = \beta$.

Define $\beta - \alpha = \gamma$.

Multiplication Let $\alpha, \beta \in Ord$. We define

$\alpha\beta =$ $\underbrace{\leftarrow \alpha \rightarrow \leftarrow \alpha \rightarrow \cdots}_{\beta}$

i.e. the order type of 'consecutive copies of α indexed by β'

Formally, $\alpha\beta = ord(\alpha \times \beta)$ where we use not quite the lexicographic ordering,

$(x, y) < (z, w)$ if either $y < w$

or $y = w$ and $x < z$.

Remark This agrees with the usual multiplication if we restrict to $\mathbb{N}$.

Proposition 22 Ordinal mult^n is associative but not commutative.

Proof Let $\alpha, \beta, \gamma \in Ord$. Then

$(\alpha\beta)\gamma =$ $\underbrace{\leftarrow \alpha\beta \rightarrow \leftarrow \alpha\beta \rightarrow \cdots}_{\gamma}$

$=$ $\underbrace{\underbrace{\leftarrow\alpha\rightarrow\leftarrow\alpha\rightarrow\cdots}_{\beta} \quad \underbrace{\leftarrow\alpha\rightarrow\leftarrow\alpha\rightarrow\cdots}_{\beta} \quad \cdots}_{\gamma}$

$=$ $\underbrace{\leftarrow\alpha\rightarrow\leftarrow\alpha\rightarrow\leftarrow\alpha\rightarrow\cdots}_{\beta\gamma}$

$= \alpha(\beta\gamma)$

However $2\omega =$ $\bullet\bullet \quad \bullet\bullet \quad \bullet\bullet \quad \bullet\bullet\cdots = \omega$

$\neq \omega 2 =$ $\bullet\cdots\bullet \quad \bullet\cdots\bullet$ $\square$

Example What about distributivity? Given $\alpha, \beta, \gamma \in Ord$,

(i) Must $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$? (No)

(ii) Must $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$? (Yes!)

Exercise (This & other things involving ordinal arithmetic will appear on Ex Sheet 2)

2.5 Ordinal Induction and Recursion

If p is a statement about ordinals and we want to prove
$(\forall \alpha \in Ord)\ p(\alpha)$.
Then, by, induction, when proving $p(\alpha)$ we can assume
$(\forall \beta < \alpha)\ p(\beta)$.
(Recall Ord is not a set. But for any $\alpha \in Ord$ then $\alpha \in I_{\alpha^+}$ and I_{α^+} is a well-ordered set of ordinals)

Similarly, we can make recursive definitions over the ordinals: when defining something at α, assume already defined at β for all $\beta < \alpha$.

Typically divide into three cases:
$\alpha = 0$, α is successor, α is non-zero limit (nzl)

Definition For ordinals α, β define $\alpha \oplus \beta$ by recursion over ordinal β:

$\beta = 0$: $\alpha \oplus 0 = \alpha$
$\beta = \delta^+$: $\alpha \oplus \delta^+ = (\alpha \oplus \delta)^+$
β nzl : $\alpha \oplus \beta = \sup\{\alpha \oplus \delta \mid \delta < \beta\}$.

And $\alpha \otimes \beta$ by recursion over ordinal β:

$\beta = 0$: $\alpha \otimes 0 = 0$
$\beta = \delta^+$: $\alpha \otimes \delta^+ = (\alpha \otimes \delta) \oplus \alpha$
β nzl : $\alpha \otimes \beta = \sup\{\alpha \otimes \delta \mid \delta < \beta\}$.

Remark Could combine $\beta = 0$ and β nzl as a single case in def^n of $\otimes$ but not in def^n of $\oplus$.
But generally not worth it – less hassle to keep the three cases separate.

Proposition 23 $\oplus$ is associative

Proof Let $\alpha, \beta, \gamma \in \mathrm{Ord}$. We prove $(\alpha \oplus \beta) \oplus \gamma = \alpha \oplus (\beta \oplus \gamma)$ by induction on γ.

$\gamma = 0$: $(\alpha \oplus \beta) \oplus 0 = \alpha \oplus \beta = \alpha \oplus (\beta \oplus 0)$ ✓

$\gamma = \delta^+$: $(\alpha \oplus \beta) \oplus \delta^+ = ((\alpha \oplus \beta) \oplus \delta)^+$

$= (\alpha \oplus (\beta \oplus \delta))^+$ by ind hyp

$= \alpha \oplus (\beta \oplus \delta)^+$

$= \alpha \oplus (\beta \oplus \delta^+)$ ✓

γ ltd : $(\alpha \oplus \beta) \oplus \gamma = \sup \{ (\alpha \oplus \beta) \oplus \delta \mid \delta < \gamma \}$

$= \sup \{ \alpha \oplus (\beta \oplus \delta) \mid \delta < \gamma \}$ ind hyp

What is $\alpha \oplus (\beta \oplus \gamma)$?

First we need to check $\beta \oplus \gamma$ is a limit.

Indeed, $\beta \oplus \gamma = \sup X$ where $X = \{ \beta \oplus \delta \mid \delta < \gamma \}$.

Suppose $\beta \oplus \gamma$ is not a limit.

Then X has a greatest element, say $\beta \oplus \delta$ where $\delta < \gamma$.

But γ is a limit and $\delta < \gamma$ so $\exists \varepsilon \in \mathrm{Ord}$ with $\delta < \varepsilon < \gamma$.

Then $\beta \oplus \varepsilon \in X$ with $\beta \oplus \varepsilon > \beta \oplus \delta$.

(by induction on ε) ※

Hence $\alpha \oplus (\beta \oplus \gamma) = \sup \{ \alpha \oplus \zeta \mid \zeta < \beta \oplus \gamma \}$.

Let $A = \{ \alpha \oplus (\beta \oplus \delta) \mid \delta < \gamma \}$, $B = \{ \alpha \oplus \zeta \mid \zeta < \beta \oplus \gamma \}$.

We require $\sup A = \sup B$.

Suppose $\delta < \gamma$. Then $\beta \oplus \delta < \beta \oplus \gamma$ so $\alpha \oplus (\beta \oplus \delta) \in B$.

Thus $A \subset B$. Hence $\sup A \leq \sup B$.

On the other hand, suppose $\zeta < \beta \oplus \gamma$.

Then $\beta \oplus \gamma = \sup \{ \beta \oplus \delta \mid \delta < \gamma \}$ so $\zeta < \beta \oplus \delta$ for some $\delta < \gamma$.

Thus $\alpha \oplus \zeta < \alpha \oplus (\beta \oplus \delta) \in A$.

Hence $\sup B \leq \sup A$.

Hence $\sup A = \sup B$, i.e. $(\alpha \oplus \beta) \oplus \gamma = \alpha \oplus (\beta \oplus \gamma)$. □

Proposition 24 $\oplus$ and $+$ coincide ; $\otimes$ and $\times$ coincide

Proof Let $\alpha, \beta \in Ord$. For the first part, we prove by induction on β that $\alpha \oplus \beta = \alpha + \beta$.

$\beta = 0$ $\quad \alpha \oplus 0 = \alpha = \longleftarrow \alpha \longrightarrow \longleftarrow 0 \longrightarrow = \alpha + 0$ ✓

$\beta = \delta^+$ $\quad \alpha \oplus \delta^+ = (\alpha \oplus \delta)^+ = (\alpha + \delta)^+$ (ind hyp)

$= \longleftarrow \alpha \longrightarrow \longleftarrow \delta \longrightarrow \bullet$

$= \longleftarrow \alpha \longrightarrow \longleftarrow \delta^+ \longrightarrow$

$= \alpha + \delta^+$ ✓

β nzl $\quad \alpha \oplus \beta = \sup \{ \alpha \oplus \delta \mid \delta < \beta \}$

$= \sup \{ \alpha + \delta \mid \delta < \beta \}$ (ind hyp)

$= \alpha + \beta$ ✓

Second part (multiplication) — Exercise (ExSheet 2) □

Remarks 1. $+, \times$ are called the 'synthetic' def^{n}s ;
$\oplus, \otimes$ are called the 'inductive' def^{n}s

Generally just use the notations $+, \times$ since P24 says both def^{n}s are the same. (And in fact, as before, usual to just write $\alpha\beta$ not $\alpha \times \beta$).

2. P24 means we often have a choice of methods. Synthetic method is often easier (when available).

3. P24 means we can switch back-and-forth between two viewpoints in the course of a single proof.

Ordinal exponentiation Let $\alpha, \beta \in Ord$. We define α^β by recursion on β:

$\beta = 0$: $\quad \alpha^0 = 1$

$\beta = \delta^+$: $\quad \alpha^{\delta^+} = \alpha^\delta \alpha$

$\beta =$ nzl : $\quad \alpha^\beta = \sup \{ \alpha^\delta \mid \delta < \beta \}$.

Chapter 3 : CARDINALS

3.1 Sizes of sets

$\mathbb{N}$: 'counting numbers' $\longrightarrow$ generalize to ordinals

But also $\mathbb{N}$ = set of sizes of finite sets $\longrightarrow$ generalize to cardinals

We say sets x, y <u>have the same cardinality</u> if there is a bijection $f : x \longrightarrow y$.

The <u>cardinals</u> should have the properties :

(i) any set x has a cardinality, $\text{card}(x)$, which is a cardinal

(ii) $\text{card}(x) = \text{card}(y)$ iff x, y have the same cardinality.

Examples For $n \in \mathbb{N}$, define the cardinal
$$n = \text{card}(\{1, 2, \ldots, n\}).$$
(Special case: $n = 0$, $0 = \text{card}(\emptyset)$.)
Define $\text{card}(\mathbb{N}) = \aleph_0$.
Define $\text{card}(\mathbb{R}) = c$.
Now, $\text{card}(\mathbb{Q}) = \aleph_0$, '$\mathbb{Q}$ is countable'
But $c \neq \aleph_0$, '$\mathbb{R}$ is uncountable'

Arithmetic operations

Def^n Let κ, λ be cardinals. Let K, L be sets with
$$\text{card}(K) = \kappa, \quad \text{card}(L) = \lambda, \quad \text{with } K \cap L = \emptyset.$$
Then $\kappa + \lambda = \text{card}(K \cup L)$,
$\kappa\lambda = \text{card}(K \times L)$,
and $\kappa^\lambda = \text{card}(\{ f \mid f : L \to K \})$.
} also works in these cases if $K \cap L \neq \emptyset$

Remark This reduces to the usual def^{n}s of $+$, $\times$, powers when restricted to $\mathbb{N}$.

Examples 1. For any κ, $2^\kappa \neq \kappa$
Why? This says 'there is no bijection from a set K to the set $\{f \mid f : K \to \{0,1\}\}$';
equivalently 'there is no bijection from K to $\mathcal{P}(K)$'.
Recall (IA) that if such a bijection $g : K \to \mathcal{P}(K)$ exists, then
$\{x \in K \mid x \notin g(x)\} \neq g(y)$ for any $y \in K$ ⁂
2. $2^{\aleph_0} = c$: 'there is a bijection from $\mathcal{P}(\mathbb{N})$ to $\mathbb{R}$'
(Proved in IA).

Basic Properties: Proposition 25 Let κ, λ, μ be cardinals. Then

(i) $(\kappa + \lambda) + \mu = \kappa + (\lambda + \mu)$;
(ii) $\kappa + \lambda = \lambda + \kappa$;
(iii) $\kappa(\lambda\mu) = (\kappa\lambda)\mu$;
(iv) $\kappa\lambda = \lambda\kappa$;
(v) $\kappa(\lambda + \mu) = \kappa\lambda + \kappa\mu$;
(vi) $\kappa^{\lambda+\mu} = \kappa^\lambda \kappa^\mu$;
(vii) $(\kappa\lambda)^\mu = \kappa^\mu \lambda^\mu$;
(viii) $(\kappa^\lambda)^\mu = \kappa^{\lambda\mu}$.

Proof Exercise. E.g. (viii) Let K, L, M be sets with cardinalities κ, λ, μ resp.

Let $X = \{ f \mid f : M \to \{ g \mid g : L \to K \} \}$

and $Y = \{ h \mid h : L \times M \to K \}$.

So $\text{card}(X) = (\kappa^\lambda)^\mu$, $\text{card}(Y) = \kappa^{\lambda\mu}$.

We must find a bijection $\varphi : X \to Y$.

Define $\varphi(f)((\ell, m)) = f(m)(\ell) \quad \forall f \in X, \ell \in L, m \in M$.

This defines $\varphi : X \to Y$. This is a bijection with inverse $\theta : Y \to X$ given by

$$h \mapsto (m \mapsto (\ell \mapsto h((\ell, m)))) \quad \forall h \in Y, \ell \in L, m \in M. \ \square$$

3.2 Comparing cardinals

Definition Let κ, λ be cardinals. Then $\kappa \leq \lambda$ means that if $\text{card}(K) = \kappa$, $\text{card}(L) = \lambda$ then there is an injection $K \to L$.

Define $<$, $\geqslant$ etc in usual way.

Would like $\leq$ to 'be a total order on the cardinals'

(not really — cardinals don't form a set)

But can try to show it has total-order-like properties:

(i) $\forall \kappa, \quad \kappa \leq \kappa$ ✓

(ii) $\forall \kappa, \lambda, \mu, (\kappa \leq \lambda, \lambda \leq \mu) \Rightarrow \kappa \leq \mu$ ✓

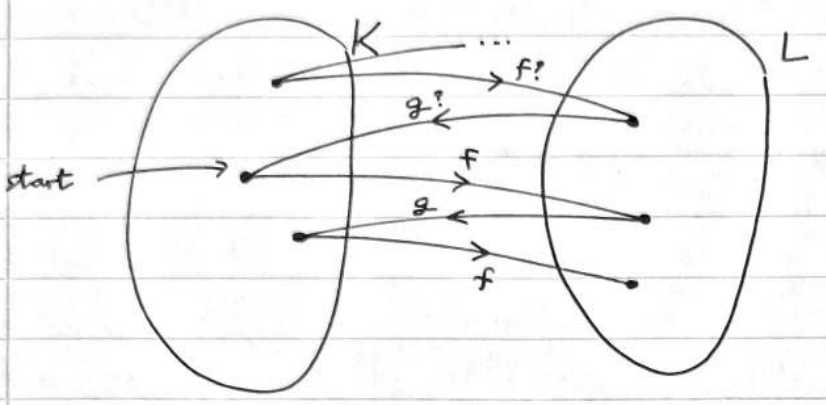
(iii) $\forall \kappa, \lambda, (\kappa \leq \lambda, \lambda \leq \kappa) \Rightarrow \kappa = \lambda$?

(iv) $\forall \kappa, \lambda, \quad \kappa \leq \lambda$ or $\lambda \leq \kappa$?

(iii) is dealt with by:

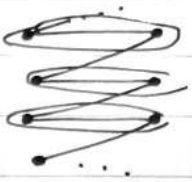
Theorem 26 (Schroder-Bernstein) Let K, L be sets s.t. there are injections $f : K \to L$, $g : L \to K$.

Then there is a bijection $h : K \to L$.

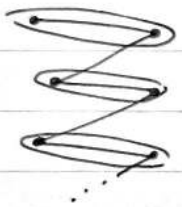


Case (i):

Case (ii):



Case (iii)(a)



(iii)(b)

Proof Assume $K \cap L = \emptyset$. For $x, y \in K \cup L$, say $x \sim y$ if $\exists\ z_0, z_1, \dots, z_n$ with $\{z_0, z_n\} = \{x, y\}$ and $z_i = f(z_{i-1})$ or $z_i = g(z_{i-1})$ for $i = 1, \dots, n$.

Clearly $\sim$ is an equivalence relation.

Let R be an equivalence class; we define a bijection

$h_R : R \cap K \to R \cap L$

Case 1 : $f(R \cap K) = R \cap L$; set $h_R(x) = f(x)$

Case 2 : $f(R \cap K) \neq R \cap L$; then $R = \{y_0, x_0, y_1, x_1, \dots\}$ where the x_i, y_i distinct, $x_i \in K$, $y_i \in L$, $g(y_i) = x_i$, $f(x_i) = y_{i+1}$.

Set $h_R(x_i) = y_i$.

Now let $h = \bigcup_R h_R$. □

(iv) $\forall\ \kappa, \lambda$, $\kappa \leq \lambda$ or $\lambda \leq \kappa$

Follows from:

Theorem 27 (Well-ordering Principle)

Every set can be well-ordered.

Proof Let X be a set. By Hartogs', there is an ordinal α with no injection $\alpha \to X$.

Suppose X cannot be well-ordered.

Now define an injection $\alpha \to X$ recursively by letting, for $x \in \alpha$, $f(x)$ to be an arbitrarily chosen element of $X \setminus \{ f(y) \mid y < x \}$ (noting this set must be non-empty as otherwise we've well-ordered X by bijecting it with an initial segment of α.) ⁂ □

Corollary 28 Let X, Y be sets. Then $\exists$ injection $X \to Y$ or an injection $Y \to X$.

Proof True for well-ordered sets so use Theorem 27 to well-order X, Y. □

Remark When proving WOP, we had to make infinitely many arbitrary choices. c.f. 'A ctble union of ctble sets is ctble'

Proof List the sets $A_1, A_2, A_3, \ldots$

List each set $A_i = \{ a_{i1}, a_{i2}, a_{i3}, \ldots \}$. □

For each of the infinitely many sets, we've had to pick a listing. This uses Alexandria Ocasio-Cortez. This is different from our other standard assumptions about sets (e.g. if A, B sets then so is $A \cup B$), as the things it produces are not uniquely specified. So might ask questions like:

1. Can we prove it from other axioms? (No)
2. In a given case, do we actually need AoC?
3. What can we prove without using AoC?

3.3 Definition of Cardinals

WOP says every cardinal is $\text{card}(\alpha)$ for some $\alpha \in \text{Ord}$.

Def An ordinal is an initial ordinal if it is the least ordinal of given cardinality.

Examples $0, 1, 2, \ldots$ are initial ordinals

ω is an initial ordinal

But if $\omega < \alpha < \omega_1$, is not an initial ordinal

But then ω_1 is an initial ordinal

For $\alpha \in \text{Ord}$, define ω_α recursively by:

$\alpha = 0$ $\quad \omega_0 = \omega$

$\alpha = \delta^+$ $\quad \omega_{\delta^+}$ is the least initial ordinal $> \omega_\delta$ $\quad (*)$

α nzl $\quad \omega_\alpha = \sup \{ \omega_\delta \mid \delta < \alpha \}$

(*) This exists by Hartogs' Lemma.

The infinite initial ordinals are precisely the ω_α for $\alpha \in \text{Ord}$.

(Why? If $\alpha \neq 0$ then ω_α is the least initial ordinal which is greater than ω_β for all $\beta < \alpha$. Additionally, by induction on α, $\forall \alpha$, $\omega_\alpha \geq \alpha$)

By WOP, the cardinalities of the initial ordinals are all the cardinals.

Notation Write $\aleph_\alpha = \text{card}(\omega_\alpha)$

So the cardinals are just $\aleph_\alpha$ ($\alpha \in \text{Ord}$) together with $0, 1, 2, 3, \ldots$

Formal definition A cardinal is an initial ordinal

Remarks 1. With this definition, $\aleph_\alpha = \omega_\alpha$

2. Not a helpful way to think of things — gives us a way to make a formal def^n but don't generally want to think of cardinals like this.

(Think '$\aleph_\alpha = \text{card}(\omega_\alpha)$' not '$\aleph_\alpha = \omega_\alpha$')

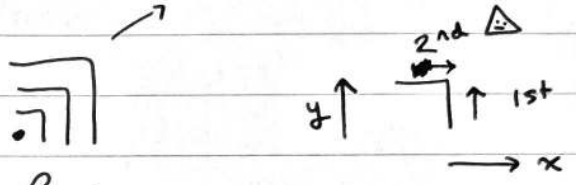
3. Ordinal exponentiation and cardinal exponentiation are not s.

Knowing that these are all the cardinals makes arithmetic much simpler:

Theorem 29 Let $\alpha \in Ord$. Then $\aleph_\alpha \aleph_\alpha = \aleph_\alpha$.

Proof By induction on α. Clearly $\aleph_\alpha \leq \aleph_\alpha \aleph_\alpha$.
So by Schroeder-Bernstein ETS, $\aleph_\alpha \aleph_\alpha \leq \aleph_\alpha$,
i.e. we need to show there is an injection $\omega_\alpha \times \omega_\alpha \longrightarrow \omega_\alpha$.

Well-order $\omega_\alpha \times \omega_\alpha$ by
'going up in squares'



Define $(x,y) < (z,t)$ if

EITHER $\max(x,y) < \max(z,t)$

OR $\max(x,y) = \max(z,t) = \beta$ say,

AND $y < \beta = t$

or $x = z = \beta$, $y < t$

or $y = t = \beta$, $x < z$

Let I be a proper initial segment of $\omega_\alpha \times \omega_\alpha$ in this well-ordering. Then $I = (\omega_\alpha \times \omega_\alpha)_\delta$ for some $\delta \in \omega_\alpha \times \omega_\alpha$.
ω_α is a limit ordinal, so $\delta \in \beta \times \beta$ for some $\beta < \omega_\alpha$.
Thus $I \subset \beta \times \beta$. But ω_α is initial so $card(\beta) < \aleph_\alpha$
That is, $card(\beta) = \aleph_\gamma$ for some ordinal $\gamma < \alpha$.
Then $card(I) \leq card(\beta \times \beta) = card(\beta)\, card(\beta)$

$$= \aleph_\gamma \aleph_\gamma = \aleph_\gamma \quad \text{(by ind hyp)}$$
$$< \aleph_\alpha$$

Hence $ord(I) < \omega_\alpha$. So every proper initial segment of our well-ordering of $\omega_\alpha \times \omega_\alpha$ has order-type $< \omega_\alpha$.
So $ord(\omega_\alpha \times \omega_\alpha) \leq \omega_\alpha$.
Hence $card(\omega_\alpha \times \omega_\alpha) \leq card(\omega_\alpha)$.
i.e. $\aleph_\alpha \aleph_\alpha \leq \aleph_\alpha$. □

Corollary 30 Let κ, λ be infinite cardinals. Then
$$\kappa + \lambda = \kappa\lambda = \max(\kappa, \lambda)$$

Proof Wlog $\kappa \leq \lambda = \aleph_\alpha$, say. Then
$$\aleph_\alpha = \lambda \leq \kappa + \lambda \leq \lambda + \lambda = 2\lambda \leq \kappa\lambda \leq \lambda\lambda = \aleph_\alpha \aleph_\alpha = \aleph_\alpha . \quad \square$$

This tells us that $+$, $\times$ are very straightforward for cardinals. However, exponentiation is not at all straightforward.

E.g. Is $2^{\aleph_0} = \aleph_1$?

($2^{\aleph_0} = \text{card}(\mathcal{P}(\mathbb{N})) = \text{card}(\mathbb{R})$; $\aleph_1$: smallest uncountable cardinal)

I.e. 'Is there an infinite set bigger than $\mathbb{N}$ but smaller than $\mathbb{R}$?'

Don't know. (Actually impossible to prove or disprove from standard assumptions about sets, even including AoC)

3.4 Cardinals without AoC

The theory of cardinals we have developed relies heavily on the Axiom of Choice (AoC) via the well-ordering principle.

Did we need AoC to prove WOP? Yes — any proof of WOP must use AoC somewhere.

Let's show this by assuming WoP and proving AoC.

Axiom of Choice Let $\mathcal{A} = \{A_i : i \in I\}$ be a set of non-empty sets. Then there exists a function $f : I \to \bigcup_{i \in I} A_i$ such that $\forall i \in I$, $f(i) \in A_i$.

(f is called a *choice function* for $\mathcal{A}$.)

Proof of AoC from WOP

Well-order $\bigcup_{i \in I} A_i$ and let $f(i) = \min(A_i)$ for each $i \in I$. $\square$

What does our theory of cardinals look like if we decide not to assume AoC?

§ 3.1 (Basic properties): No use made of AoC

§ 3.2 (Ordering): We proved 4 properties of $\leq$:

(i) $\forall\ \kappa,\ \kappa \leq \kappa$ ✓

(ii) $\forall\ \kappa, \lambda, \mu,\ (\kappa \leq \lambda,\ \lambda \leq \mu) \Rightarrow \kappa \leq \mu$ ✓

(iii) $\forall\ \kappa, \lambda,\ (\kappa \leq \lambda,\ \lambda \leq \kappa) \Rightarrow \kappa = \lambda$ ✓

(iv) $\forall\ \kappa, \lambda,\ \kappa \leq \lambda$ or $\lambda \leq \kappa$

(i), (ii), (iii) still survive without AoC

BUT our proof of (iv) used WOP and hence AoC

In fact, AoC is essential here: any proof that given any two sets, one injects into the other, must use AoC (Ex Sheet 2)

So if we don't assume AoC, we could have 'incomparable' cardinals, we could get cardinals κ, λ s.t. $\kappa \not\leq \lambda$ and $\lambda \not\leq \kappa$.

§ 3.3 (definition of cardinals in terms of initial ordinals)

Again, highly dependent on AoC via WOP.

Without AoC? Everything still valid as 'theory of cardinalities of well-ordered sets'.

Still can define our $\aleph$s and behave in the same way, e.g. $\aleph_\alpha + \aleph_\beta = \aleph_\alpha \aleph_\beta = \aleph_{\max(\alpha,\beta)}$

BUT there could be other infinite cardinals that are not $\aleph$s.

So proof for arbitrary infinite cardinals κ, λ that

$$\kappa + \lambda = \kappa\lambda = \max(\kappa, \lambda)$$

no longer works without AoC.

Even worse: our formal definition no longer works.

Can we make a formal definition of cardinals without AoC? (Later)

Interlude : Outstanding Matters

Ch 1 Propositional Logic

Still-to-do : finish proof of completeness theorem when P uncountable.

(End of Ch 1 : Looked at how to do this informally
cf Ch 3 : Proof of WOP - gives an idea of how to formalize
Ch 4 : Do it properly, within general framework of proving things like this.)

Ch 2 Ordinals.

We defined an ordinal as a well-ordered set where 'two are considered to be the same if they are isomorphic'.
What does this really mean? Can't do this by making an ordinal an equivalence class of sets, as well-ordered sets don't form a set).
(Ch 6 (Set Theory) : formalize this defⁿ)

Ch 3 Cardinals.

What really is a cardinal ?
(i) With AC, can define cardinal to be a limit ordinal (or 0, 1, 2, ...). So once we've filled gap in Ch 2, that gives a formal defⁿ of cardinals.
(ii) Without AC, this no longer works. Can we define cardinals formally without using AC ?
Informally : e.g. a cardinal is a set where 'two sets are considered to be the same if there is a bijection between them. Meaning ... ?
(Ch 6 as well to formalize this)

Chapter 4 Zorn's Lemma

4.1 Posets

A partial order on a set X is a relation $\leq$ on X s.t.

(i) $\forall x \in X, \quad x \leq x$

(ii) $\forall x, y, z \in X, \quad (x \leq y, y \leq z) \Rightarrow x \leq z$

(iii) $\forall x, y \in X, \quad (x \leq y, y \leq x) \Rightarrow x = y$

Define $<, \geq, >$ in obvious way as with total orders.

Conditions (i), (ii), (iii) on $\leq$ are equivalent to the following conditions (I), (II) on $<$:

(I) $\forall x \in X, \quad x \not< x$

(II) $\forall x, y, z \in X, \quad (x < y, y < z) \Rightarrow x < z$

(Exercise: check this equivalence)

A poset is a set X with a partial order $\leq$ on X. Often write the poset as an ordered pair $(X, \leq)$.

e.g. 'Let $(X, \leq)$ be a poset.'

Examples 1. Any totally ordered set

2. $(\mathbb{N} \setminus \{0\}, |)$ ↑ 'divides'

Not a total order: $2 \nmid 3$ and $3 \nmid 2$

3. Let Y be any set: $(\mathcal{P}(Y), \subseteq)$ ↑ 'is a subset of'

Not a total order in general: if $a, b \in Y$ with $a \neq b$ then $\{a\} \not\subseteq \{b\}$ and $\{b\} \not\subseteq \{a\}$.

(In fact, total order $\Leftrightarrow |Y| = 0, 1$)

4. If X is any subset of $\mathcal{P}(Y)$, then $(X, \subseteq)$.

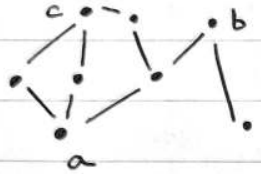
e.g. Let V be a vector space. Then the set of subspaces of V ordered by inclusion is a poset.

We can represent a poset by a Hasse diagram: draw a dot for each $x \in X$ and an upwards line from x to y if y covers x: $x < y$ and $\nexists z, x < z < y$.

By transitivity: if we can start at x and get to y by following (zero or more) upwards lines then $x \leq y$ (N.B. not conversely in general).

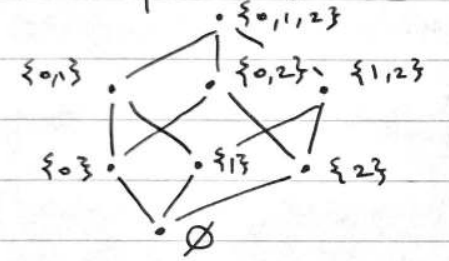
Sometimes useful, sometimes not.

Useful examples 1. If X is finite, Hasse diagram tells us everything

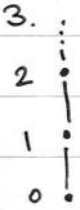


$a \leq b$, $a \leq c$,
$b \not\leq c$, $c \not\leq b$

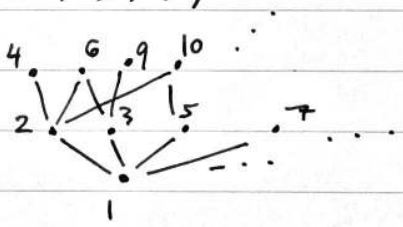
2. Specific finite example: $(\mathcal{P}(\{0,1,2\}), \subseteq)$



3. $(\mathbb{N}, \leq)$



4. $(\mathbb{N} \setminus \{0\}, |)$



Useless Example $(\mathbb{Q}, \leq)$

Let $x, y \in \mathbb{Q}$, $x < y$.

Then $\frac{1}{2}(x+y) \in \mathbb{Q}$, $x < \frac{1}{2}(x+y) < y$.

So nothing covers anything.

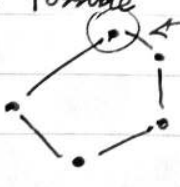
Hasse diagram: infinitely many dots, no lines.

Tells us practically nothing.

Helpful: finite posets, some infinite posets

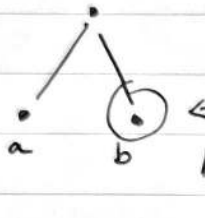
Unhelpful: other infinite posets

E.g. Possible misconception 'Always put things into levels saying how far they are above bottom'.



One way 2 steps above bottom, another way 3 steps

Let $(X, \leq)$ be a poset. We say $x \in X$ is <u>least</u> if $\forall y \in X$, $x \leq y$. Similarly, we say $x \in X$ is <u>greatest</u> if $\forall y \in X$, $y \leq x$.



← not least, but it is minimal $b \not\leq a$

$x \in X$ is <u>minimal</u> (resp <u>maximal</u>) if $\forall y \in X$, $y \leq x \Rightarrow y = x$ (resp. if $\forall y \in X$, $x \leq y \Rightarrow x = y$)

least $\Rightarrow$ minimal but minimal $\not\Rightarrow$ least
Similarly greatest $\Rightarrow$ maximal but maximal $\not\Leftrightarrow$ greatest

A <u>chain</u> is a totally ordered subset of X.

<u>Examples</u> 1. Any totally ordered set X has X as a chain.
e.g. in $(\mathbb{R}, \leq)$, $\mathbb{R}$ is a chain
(And so is any subset: e.g. $\mathbb{Q}$ is a chain in $(\mathbb{R}, \leq)$)
2. Let $(X, \leq)$ be a poset. Then $\emptyset$ is a chain.
3. Let $(\mathbb{N} \setminus \{0\}, \nmid)$ be our poset. Then
$\{2^n \mid n \in \mathbb{N}\} \subset \mathbb{N} \setminus \{0\}$ is a chain.

Let $(X, \leq)$ be a poset and let $Y \subset X$. An <u>upper bound</u> for Y is $x \in X$ s.t. $\forall y \in Y$, $y \leq x$.
The <u>supremum</u> $\sup(Y)$ is the least upper bound of Y. (if it exists)

<u>Remark</u> Given $Y \subset X$, $\sup(Y)$ may or may not exist. If it exists it is unique; however, it may or may not lie in Y.
If $\sup Y \in Y$ then $\sup Y$ is the greatest element of Y.
Similarly, we can define lower bounds, $\inf(Y)$.

A __Lattice__ is a poset X in which every finite subset has a sup and an inf.

For $A \subset X$ finite, write $\bigwedge A = \inf(A)$, $\bigvee A = \sup A$.

If $a, b \in X$, write

$$a \wedge b = \bigwedge \{a, b\}, \quad a \vee b = \bigvee \{a, b\}$$

('meet' under $\bigwedge$, 'join' under $\bigvee$)

Any lattice X has a greatest elt $\top = \bigwedge \emptyset$ and a least elt $\bot = \bigvee \emptyset$.

A lattice is a __Boolean Algebra__ if it satisfies

(i) $\forall x, y, z \in X$, $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$;

(ii) $\forall x, y, z \in X$, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$;

} de Morgan's Laws

(iii) $\forall x \in X, \exists y \in X$ s.t. $x \vee y = \top$ and $x \wedge y = \bot$

__Example__ $(\mathcal{P}Y, \subset)$ is a Boolean Algebra

(i) (ii) ✓ $\wedge = \cap$, $\vee = \cup$

(iii) $y = x^c$, $x \cup x^c = \top = Y$

$x \cap x^c = \bot = \emptyset$

4.2 Zorn's Lemma and Applications

Theorem 31 (Zorn's Lemma) Let $(X, \leq)$ be a poset in which every chain has an upper bound. Then X has a maximal element.

Proof Suppose not. Then for each $x \in X$ there is some $f(x) \in X$ with $f(x) > x$. Also, for each chain $C \subset X$, let $G(C)$ be an upper bound for C.

By Hartogs' Lemma, there is an ordinal β such that there is no injection $\beta \to X$.

For $\alpha \in Ord$ with $\alpha < \beta$ recursively define

$$x_\alpha = f(G(\{x_\delta \mid \delta < \alpha\}))$$

(noting that the set $\{x_\delta \mid \delta < \alpha\}$ is a chain, by induction)

Then $\alpha \mapsto x_\alpha$ is an injection $I_\beta \to X$ ⁂

(as $I_\beta \cong \beta$). □

Remarks 1. Formally ($\ddot\smile$), we should say:

$$x_\alpha = \begin{cases} f(G(\{x_\delta \mid \delta < \alpha\})) & \text{if } \{x_\delta \mid \delta < \alpha\} \text{ is chain} \\ \text{a cabbage} & \text{o/w} \end{cases}$$

Then use induction to show that there aren't any cabbages. This happens frequently; usually don't comment on cabbages.

2. N.B. if $\alpha = 0$, $\{x_\delta \mid \delta < \alpha\} = \emptyset$

This *is* a chain, and any element of X is an upper bound for $\emptyset$. So in case $\alpha = 0$, we've basically saying 'Pick any $x_0 \in X$'.

3. When applying Zorn, we will need to check the condition that every chain has an upper bound. In particular, this means we need to check $\emptyset$ has an upper bound. Often useful to check this explicitly as a special case.

i.e. in practice, often check condition in form

'$X \neq \emptyset$ and if $C \subset X$ is a non-empty chain then C has an upper bound'

4. We have clearly used AC in our proof of Zorn. In fact, this is essential. (See §4.3)

Proofs using ZL tend to follow a typical pattern. We'll do two applications to illustrate this.

First application: every vector space has a basis.

Recall that if V is a vector space and $U \subset V$ then U is

- <u>spanning</u> if every elt of V is a linear combination of (finitely many) elts of U;
- <u>linearly independent</u> if no non-trivial linear combintion of (finitely many) elts of U is zero;
- a <u>basis</u> if it's LI and spanning.

<u>Theorem 32</u> Every vector space has a basis

<u>Proof</u> Let V be a vector space over the field k.

Let $X = \{U \subset V \mid U \text{ LI}\}$ ordered by inclusion, so $(X, \subset)$ is a poset.

First, $\emptyset \subset V$ is LI so $\emptyset \in X$ so $X \neq \emptyset$.

Let $\mathcal{C} \subset X$ be a non-empty chain, say $\mathcal{C} = \{U_i \mid i \in I\}$.

Let $U = \bigcup_{i \in I} U_i$.

Is $U \in X$?

Clearly $U \subset V$.

Suppose U is not LI.

Then $\exists\, u_1, \dots, u_n \in U$ distinct and $\lambda_1, \dots, \lambda_n \in k$ not all zero with $\lambda_1 u_1 + \cdots + \lambda_n u_n = 0$.

But each of the u_j is in some U_{i_j}; there are finitely many of them and the U_is are nested, so there is some $i \in I$ s.t. for all j $(1 \leq j \leq n)$ we have $u_j \in U_i$.

So U_i is not LI ⚔ Hence $U \in X$.

Clearly $\forall i \in I$, $U_i \subset U$. So U is an upper bound for $\mathcal{C}$

Hence by Zorn's Lemma, X has a maximal element B.

By definition of B, B is LI. Suppose B does not span V. Pick $v \in V$ and $v \notin \text{span}(B)$ and let $B' = B \cup \{v\}$. Then $B' \in X$ with $B \subsetneq B'$ ⚡
Hence B is a basis for V. □

Remarks 1. This is typical of an application of Z-L:
(i) Define a poset X
(ii) Check $X \neq \phi$
(iii) Check if $C \subset X$ is a non-empty chain then it has an upper bound x (where we make sure $x \in X$)
(iv) Use ZL to get maximal element M.
(v) Check M is the thing we want.
Usually most of the 'real maths' happens in Step (v).
2. The type of poset used in the proof of Thm 31 is fairly common in applications: X is some collection of sets ordered by inclusion. Often the upper bound we want for $C = \{U_i \mid i \in I\}$ is $\bigcup_{i \in I} U_i$.
Obvious this contains every U_i $(i \in I)$. Sometimes need to do some work to check that it actually lies in the set X.

Second application: Completeness Theorem for Propositional Logic
Recall that our proof of the key Lemma only worked for P countable. Now use Zorn to do it in general.

Lemma 5 Let $S \subset L$ be consistent. Then there exists a consistent $\overline{S}$, s.t. for all $p \in L$ either $p \in \overline{S}$ or $\neg p \in \overline{S}$.

Proof Let X be the poset $\{U \subset L \mid U \text{ consistent}, S \subset U\}$ ordered by inclusion.
$X \neq \phi$ as $S \in X$.
Let $\mathcal{C} = \{U_i \mid i \in I\}$ be a non-empty chain in X.
Let $U = \bigcup_{i \in I} U_i$. Is $U \in X$? Clearly $U \subset L$.
Also, $S \subset U$. Suppose $U \vdash \bot$. Write down a proof of

$\bot$ from U. As proofs are finite, we only use finitely many hypotheses $p_1, \dots, p_n \in U$. Each p_j is in some U_i, there are finitely many of them and the U_i's are nested so $\exists i \in I$ s.t. $\forall j$ $(1 \leq j \leq n)$, $p_j \in U_i$.
So our proof of $\bot$ is actually a proof of $\bot$ from U_i, i.e. $U_i \vdash \bot$. $\lightning$ So $U \nvdash \bot$ and so $U \in X$.
Hence U is an upper bound for $\mathcal{C}$.
Hence by Zorn, there is a maximal element $\overline{S} \in X$.
By def^n of X, $\overline{S} \subset L$, $\overline{S}$ is consistent and $S \subset \overline{S}$.
Suppose $\exists p \in L$ s.t. $p \notin \overline{S}$ and $\neg p \notin \overline{S}$.
By exactly the same argument as we used when proving LS in the ctble case, at least one of $\overline{S} \cup \{p\}$ and $\overline{S} \cup \{\neg p\}$ is consistent.
Assume wlog $\overline{S} \cup \{p\}$ is consistent. Then
$\overline{S} \cup \{p\} \in X$, $\overline{S} \cup \{p\} \supsetneq \overline{S}$ $\lightning$ $\square$

We have now proved the completeness theorem for propositional logic in full.

4.3 Zorn's Lemma and the Axiom of Choice

Our proof of Zorn's Lemma used AC.

Theorem 33 Any proof of Zorn's Lemma must use AC.

Proof We shall prove AC assuming Zorn's Lemma.

Let $\mathcal{A} = \{A_i : i \in I\}$ be a collection of non-empty sets.

Let $X = \{ (J, f) \mid J \subset I, f : J \to \bigcup_{i \in I} A_i, \forall i \in J, f(i) \in A_i \}$,

made into a poset by the ordering

$(J, f) \leq (K, g)$ if $J \subset K$, $g|_J = f$

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

Let $\mathcal{C} = \{(J_h, f_h) \mid h \in H\}$ be a non-empty chain in X.

Let $J = \bigcup_{h \in H} J_h$ and $f = \bigcup_{h \in H} f_h : J \to \bigcup_{i \in I} A_i$

[That is, if $i \in J$ then $f(i) = f_h(i)$ for any $h \in H$ with $i \in J_h$. This is well-defined as $\mathcal{C}$ is a chain.]

Is $(J, f) \in X$?

Clearly $J \subset I$. Next, we check f is indeed a well-defined function $f : J \to \bigcup_{i \in I} A_i$.

Certainly $f \subset J \times \bigcup_{i \in I} A_i$. Suppose $(x, y), (x, z) \in f$.

Then $x \in J_h$ for some $h \in H$ with $f_h(x) = y$ and also $x \in J_g$ for some $g \in H$ with $f_g(x) = z$.

As $\mathcal{C}$ is a chain, wlog $(J_h, f_h) \leq (J_g, f_g)$. Then $J_h \subset J_g$ and $f_g|_{J_h} = f_h$.

In particular, $f_g(x) = f_h(x)$ i.e. $y = z$.

Moreover, given any $x \in J$ then $x \in J_h$ for some $h \in H$ and so $(x, f_h(x)) \in f$.

Hence $f : J \to \bigcup_{i \in I} A_i$ is a well-defined function.

Finally, given any $x \in J$ then $x \in J_h$ for some $h \in H$ and so $f(x) = f_h(x) \in A_x$.

Hence $(J, f) \in X$. Clearly (J, f) is an upper bound for $\mathcal{C}$.

Thus, by Zorn, X has a maximal element (K, e).

Suppose $K \neq I$. Take $i \in I \setminus K$, and any $y \in A_i$ and set

$L = K \cup \{i\}$ and $d = e \cup \{(i,y)\}$.
Then $(L,d) \in X$ and $(L,d) > (K,e)$. ⁂
Hence $K = I$ and e is a choice function for $\mathcal{A}$. □

Remark This proof has not assumed AC.
'AC for finite collections' follows from our usual assumptions about sets.
E.g. Let $\mathcal{A} = \{A_0, \dots, A_{n-1}\}$ where $A_0, \dots, A_{n-1}$ are non-empty sets. Want to find a choice function f for $\mathcal{A}$.

$A_0 \neq \emptyset$ so $\exists\, x_0 \in A_0$
$A_1 \neq \emptyset$ so $\exists\, x_1 \in A_1$
⋮
$A_{n-1} \neq \emptyset$ so $\exists\, x_{n-1} \in A_{n-1}$

Define $f : \{0, 1, \dots, n-1\} \to \bigcup_{i=0}^{n-1} A_i$ by $f(i) = x_i$.

Without AC, we can prove a 'weakened Zorn'.
When proving Zorn, used AC in two different ways:
(i) for each chain $\mathcal{C}$ we picked an upper bound for $\mathcal{C}$
(ii) under the assumption that no maximal element exists, we had to pick for each $x \in X$, some $f(x) > x$.
We shall weaken Zorn by strengthening its hypotheses so these choices are already made.
Definition Let $(X, \leq)$ be a poset. We say $f : X \to X$ is inflationary if $\forall x \in X$, $x \leq f(x)$. We say X is chain-complete if every chain has a supremum.
Theorem 34 (Bourbaki-Witt Theorem) Let $(X, \leq)$ be a chain-complete poset and let $f : X \to X$ be inflationary.
Then f has a fixed point.
Proof By Zorn, X has a maximal element x.
Then $x \leq f(x)$ so by maximality, $f(x) = x$. □

Remarks Bourbaki-Witt can be thought of as the 'choice-free' part of Zorn's Lemma':

1. We can prove B-W without assuming AC — mimic proof of Zorn but not making arbitrary choices - done for us by f and sup.
2. Given B-W, can use AC to deduce Zorn.

CHAPTER 5 FIRST-ORDER LOGIC

Throughout the following definitions, useful to keep in mind two examples.

1. The Theory of Groups What is a group? A set A together with a binary operation ('multiplication'), a 'unary' operation ('inverse') and a constant ('identity') subject to certain axioms.

Write m, i, e for multiplication, inverse, identity.
Then can think of m, i, e as functions:

$m: A^2 \to A$, $i: A^1 \to A$, $e: A^0 \to A$

↑ function of arity 2; arity 1; arity 0

Axioms: $(\forall x,y,z)\ m(x, m(y,z)) = m(m(x,y), z)$

$(\forall x)\ (m(x,e) = x) \wedge (m(e,x) = x)$

$(\forall x)\ (m(x, i(x)) = e) \wedge (m(i(x), x) = e)$

2. The theory of posets What is a poset? A set A together with a binary relation $\leq$ subject to some axioms.

$\leq \subset A^2$ 'predicate of arity 2'

Axioms: $(\forall x)\ (x,x) \in \leq$

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

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

4.1 First-order languages

A first-order signature is an ordered triple (Σ, Π, α) where Σ and Π are disjoint sets with

$(\ ,\ ), \Rightarrow, \bot, =, \forall, w, ' \notin \Sigma \cup \Pi$ and

$\alpha : \Sigma \cup \Pi \to \mathbb{N}$

We call elements of Σ function symbols, elements of Π predicate symbols and α the arity function.

Examples 1. Language of Groups (LG):

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

2. Language of Posets (LP):

$\Sigma = \emptyset$, $\Pi = \{\leq\}$, $\alpha(\leq) = 2$.

Given a signature (Σ, Π, α), we define the Language $\mathcal{L} = \mathcal{L}(\Sigma, \Pi, \alpha)$ associated with it as follows.

The variables of $\mathcal{L}$ are defined inductively by:

- w is a variable;
- if x is a variable then x' is a variable.

(So variables are $w, w', w'', w''', w'''', \ldots$)

Often use informally the variables as $w_0, w_1, w_2, w_3, \ldots$

Or indeed use $x, y, z \ldots$

The terms of $\mathcal{L}$ are defined inductively by:

- each variable is a term;
- if $f \in \Sigma$, $\alpha(f) = n$ and $t_1, \ldots, t_n$ are terms then $f t_1 \cdots t_n$ is a term.

(So, for example, if $c \in \Sigma$ with $\alpha(c) = 0$ then c is a term — called a constant symbol

e.g. in LP the terms are just variables

In LG have terms like e, w_0, w_{29}, $m w_0 w_1$, $i w_0$, $m w_0 m w_1 i w_0 \left(w_0 (w_1 w_0^{-1})\right)$.

Note that in general a term is constructed in a unique way by the definition above so there is no ambiguity. But sometimes convenient to use informal notation to aid readability by introducing brackets and commas.

e.g. terms of LG above: e, w_0, w_{29},
$m(w_0, w_1)$, $i(w_0)$, $m(w_0, m(w_1, i(w_0)))$.)

The formulae of $\mathcal{L}$ are defined inductively by:

- $\bot$ is a formula;
- if s, t are terms then $(s = t)$ is a formula;
- if $\varphi \in \Pi$, $\alpha(\varphi) = n$ and $t_1, \ldots, t_n$ are terms, then $\varphi t_1 \cdots t_n$ is a formula;
- if p, q are formulae then $(p \Rightarrow q)$ is a formula;
- if p is a formula and x is a variable then $(\forall x)\, p$ is a formula.

(Note in last bullet point there is no requirement for the variable x to appear in the formula p.

e.g. $(\forall w_0)(w_1 = w_2)$ is a formula,
or $(\forall w_0)((w_1 = w_0) \Rightarrow (\forall w_0)(w_2 = w_1))$ is a formula.

Again, no ambiguity – each formula is built up in a unique way from the rules. Again, informally we may add/remove brackets, commas, use different sizes/styles of brackets to improve readability.)

Examples 1. LoG. Some formulae:

$(\forall x)(\forall y)(\forall z)\ (mxmyz = mmxyz)$

$(\forall x)(\forall y)\ (mxy = myx)$

$(\forall x)\,(ix = x) \Rightarrow (\forall x)(mxx = e)$

$((x = y) \Rightarrow (\forall x)\,(iz = mzz))$

"$x(yz) = (xy)z$"

2. A formula from LP:

$(\forall x)(\forall y)(\forall z)\ (\leqslant xy \Rightarrow (\leqslant yz \Rightarrow \leqslant xz))$.

Remarks 1. A formula is a finite string of symbols drawn from the alphabet $\Pi \cup \Sigma \cup \{(,), \Rightarrow, \bot, \forall, w, ', =\}$.

2. Define informal abbreviations as in Ch 1: $\neg$, $\wedge$, $\vee$

New one: $(\exists x)p$ means $\neg(\forall x)\neg p$

Definition The **language** $\mathcal{L} = \mathcal{L}(\Sigma, \Pi, \alpha)$ is the set of formulae.

Some further definitions

A term is **closed** if it contains no variables.
(e.g. LG: memiee is a closed term)

The **bound** occurrences of a variable x in a formula p are defined inductively by:

- all occurrences of x in $(\forall x)p$ are bound;
- if an occurrence of x in p or q is bound, then so is the corresponding occurrence in $(p \Rightarrow q)$.

An occurrence of x that is not bound is **free**.

The **free variables** of a formula p are the variables that occur free in p. Write $FV(p)$ for the set of free variables of p.

Examples (LG) 1. $mxx = e \Rightarrow (\exists y)(myy = x)$

(free, free; bound, free)

2. $(mxx = e) \Rightarrow (\forall x)(mxy = myx)$ [STUPID + UNHELPFUL]

(free; bound)

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

If p is a formula, t is a term and x is a variable then $p[t/x]$ is the formula obtained from p by substituting t for every free occurrence of x.

4.2 Semantic entailment

Let $\mathcal{L} = \mathcal{L}(\Sigma, \Pi, \alpha)$ be a language.

An **$\mathcal{L}$-structure** is a set A endowed with

- for each $f \in \Sigma$, a function $f_A : A^{\alpha(f)} \to A$;
- for each $\varphi \in \Pi$, a subset $\varphi_A \subset A^{\alpha(\varphi)}$.

Example If $\mathcal{L} = LG$: an $\mathcal{L}$-structure is a set A with functions $m_A : A^2 \to A$, $i_A : A \to A$, $e_A \in A$ a constant.

N.B. A group can be made into an LG-structure in the obvious way. But an LG-structure need not be a group.

The interpretation of a closed term t in A is defined inductively by: Σ(!)

- if $f \in \Omega$, $\alpha(f) = n$, $t_1, \ldots, t_n$ are closed terms, then
$$t_A = f_A((t_1)_A, \ldots, (t_n)_A)$$
where $t = f t_1 \cdots t_n$

(Remark: if f is a constant symbol, $f_A = f_A$)

⚠ $\mathcal{L}$-structures are non-empty

The interpretation of a formula p in the $\mathcal{L}$-structure A is $p_A \in \{0,1\}$, defined inductively by:

- $\bot_A = 0$,
- if s, t are closed terms then
$$(s=t)_A = \begin{cases} 1 & \text{if } s_A = t_A, \\ 0 & \text{otherwise;} \end{cases}$$
- if $\varphi \in TT$, $\alpha(\varphi) = n$, $t_1, \dots, t_n$ are closed terms, then
$$(\varphi t_1 \cdots t_n)_A = \begin{cases} 1 & \text{if } ((t_1)_A, \dots, (t_n)_A) \in \varphi_A, \\ 0 & \text{otherwise;} \end{cases}$$
- if p, q are sentences then
$$(p \Rightarrow q)_A = \begin{cases} 0 & \text{if } p_A = 1,\ q_A = 0, \\ 1 & \text{otherwise;} \end{cases}$$
- if $(\forall x)p$ is a sentence where x is a variable, p a formula, then
$$((\forall x)p)_A = \begin{cases} 1 & \text{if for all } a \in A,\ p[\bar{a}/x]_A = 1, \\ 0 & \text{otherwise} \end{cases}$$

where, for each $a \in A$, we define a language $\mathcal{L}_a$ by adding a constant symbol $\bar{a}$ to $\mathcal{L}$ and make A into an $\mathcal{L}_a$-structure via interpreting $\bar{a}_A = a$.

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

A theory is a set of sentences $T \subset \mathcal{L}$.

We say A is a model of T if A is a model of p for all $p \in T$. (In particular A is a model of p iff A is a model of $\{p\}$).

Let T be a set of sentences and p a sentence.

We say T semantically entails p if every model of T is a model of p in which case we write $T \models p$.

If $T = \emptyset$, we say p is a tautology and write $\models p$.

Examples of Theories 1. The theory of groups (GT)

Language of groups: $\Omega = \{m, i, e\}$, $\Pi = \emptyset$ rmk $\Omega = \Sigma$

$\alpha(m) = 2$, $\alpha(i) = 1$, $\alpha(e) = 0$

$GT = \{ (\forall x)(\forall y)(\forall z)(mxmyz = mmxyz),$
$(\forall x)((mex = x) \wedge (mxe = x)),$ ← "Group Theory"
$(\forall x)((mxix = e) \wedge (mixx = e)) \}$

The axioms of a theory T are the sentences $p \in T$.

Now a group is just a model of GT.

2. Theory of posets (PT)

Language of posets: $\Omega = \emptyset$, $\Pi = \{\leq\}$, $\alpha(\leq) = 2$

$PT = \{ (\forall x)(x \leq x),$ ↑ i.e. $\leq xx$ $(\forall x)(\forall y)(\forall z)(((x \leq y) \wedge (y \leq z)) \Rightarrow (x \leq z)),$
$(\forall x)(\forall y)(((x \leq y) \wedge (y \leq x)) \Rightarrow (x = y)). \}$

A model of PT is a non-empty poset.

3. Theory of fields (FT)

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

$\alpha(+) = \alpha(\times) = 2$, $\alpha(-) = 1$, $\alpha(0) = \alpha(1) = 0$

$FT = \{ (\forall x)(\forall y)(\forall z)(++xyz = +x+yz),$
$(\forall x)((+x0 = x)), (\forall x)(+x-x = 0),$
$(\forall x)(\forall y)(+xy = +yx), (\forall x)(\forall y)(\times xy = \times yx),$
$(\forall x)(\forall y)(\forall z)(\times x \times yz = \times\times xyz),$
$(\forall x)(\times x1 = x),$
$(\forall x)(\neg(x = 0) \Rightarrow (\exists y)(\times xy = 1)),$
$\neg(0 = 1),$
$(\forall x)(\forall y)(\forall z)(\times x + yz = +\times xy \times xz) \}$ whew!

4. Theory of real vector spaces ($V_{\mathbb{R}}T$)

Language of real vector spaces?

$\Omega = \{a, z, -\} \cup \mathbb{R}$, $\Pi = \emptyset$

$\alpha(a) = 2$, $\alpha(z) = 0$, $\alpha(-) = 1$, $\forall \lambda \in \mathbb{R}$, $\alpha(\lambda) = 1$.

Exercise Work out the axioms. (Hint: uncountably many axioms, but grouped together in a natural way into finitely groups)

E.g. for each $\lambda \in \mathbb{R}$ we have an axiom

$(\forall x)(\lambda axy = a\lambda x\lambda y)$.

5. The theory of graphs (GrT)

A graph is an ordered pair (V, E) where V is a set and E set of unordered pairs of elements of V.
Things in V are vertices; things in E are edges.
Edges join two vertices; these two vertices are adjacent.
N.B. No finiteness condition: V can be infinite.
Language of Graphs: $\Omega = \emptyset$, $\Pi = \{a\}$, $\alpha(a) = 2$
(a: "adjacent")

$$GrT = \{ (\forall x) \neg axx,\ (\forall x)(\forall y)(axy \Rightarrow ayx) \}.$$

Might want both $\Omega \neq \emptyset$ and $\Pi \neq \emptyset$; e.g. theory of ordered fields.

5.3 Syntactic entailment (proof)

Logical axioms

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

Remarks? Definition Let $p \in \mathcal{L}$ with $FV(p) = \{w_{i_1}, \dots, w_{i_n}\}$ $(i_1 < \dots < i_n)$. The universal closure of p is the sentence

$$\breve{p} = (\forall w_{i_1}) \dots (\forall w_{i_n})\, p$$

(In particular, if p is a sentence then $\breve{p} = p$.)

Remarks 1. If p is a logical axiom then $\check{p}$ is a tautology.

2. Suppose we allowed $A = \phi$ as an $\mathcal{L}$-structure. Then we would have $((\forall x)\bot)_\phi = 1$ but $\bot_\phi = 0$.

However, $((\forall x)\bot) \Rightarrow \bot$ is an instance of Axiom 6 and we now have $(((\forall x)\bot) \Rightarrow \bot)_\phi = 0$.

So Axiom 6 would no longer always be a tautology.

This is why we disallow empty $\mathcal{L}$-structures.

Rules of deduction

M.P. From p and $p \Rightarrow q$ we can deduce q.

Gen. From p we can deduce $(\forall x)p$ as long as x doesn't appear free in any hypothesis used in the proof of p.

('Generalisation')

Definition Let $S \subset \mathcal{L}$ and let $p \in \mathcal{L}$. Then a **proof** of p from S is a finite sequence of 'lines'

$l_1, l_2, \ldots, l_n \in \mathcal{L}$ such that each l_i is a logical axiom, or a **hypothesis** (i.e. $l_i \in S$) or follows from earlier lines using a deduction rule.

If such a proof exists we say S **proves** p or S **syntactically entails** p and we write $S \vdash p$.

In the case $S = \emptyset$, write $\vdash p$.

Example proof $\Omega = \emptyset$, $\Pi = \emptyset$. $\{x=y, x=z\} \vdash y=z$.

⌜Try Ax 5 with p being $x=z$⌟

1. $(\forall x)(\forall y)\,((x=y) \Rightarrow ((x=z) \Rightarrow (y=z))$ (Ax 5)
2. $(\forall x)(\forall y)\,((x=y) \Rightarrow ((x=z) \Rightarrow (y=z)))$
   $\Rightarrow (\forall y)\,((x=y) \Rightarrow ((x=z) \Rightarrow (y=z)))$ ⌜Try Ax 6 with $t=x$ i.e.⌟
3. $(\forall y)((x=y) \Rightarrow ((x=z) \Rightarrow (y=z)))$ (MP on 1, 2)
4. $(\forall y)((x=y) \Rightarrow ((x=z) \Rightarrow (y=z)))$
   $\Rightarrow ((x=y) \Rightarrow ((x=z) \Rightarrow (y=z)))$ (Ax 6)
5. $(x=y) \Rightarrow ((x=z) \Rightarrow (y=z))$ (MP on 3, 4)
6. $x=y$ (hyp)
7. $(x=z) \Rightarrow (y=z)$ (MP on 5, 6)
8. $x=z$ (hyp)
9. $y=z$ (MP on 7, 8)

Proposition 35 (Deduction Theorem). Let $S \subset \mathcal{L}$ and $p, q \in \mathcal{L}$. Then $S \vdash (p \Rightarrow q)$ iff $S \cup \{p\} \vdash q$.

Proof "$\Rightarrow$" Exactly as in Ch 1: given a proof of $p \Rightarrow q$ from S, append the lines p, and q (the former being a hypothesis and the latter uses MP on p and $p \Rightarrow q$).

"<=" As Ch 1: given a proof of q from $S \cup \{p\}$, say $l_1, \ldots, l_n$ then we want to show $S \vdash (p \Rightarrow l_i)$ for all i.
The only new case is that l_i follows from an earlier line by (Gen).
So assume $l_i = (\forall x)\, r$ where $\exists\, j < i$ with $l_j = r$.
By ind hyp, we have already constructed a proof of $p \Rightarrow r$ from S.
Note first that l_i was obtained from earlier line r by (Gen.) in our proof from $S \cup \{p\}$. Hence no hypothesis used in proof of r from $S \cup \{p\}$ had a free occurrence of x. Thus no hypothesis used in proof of $p \Rightarrow r$ from S had a free occurrence of x. So by (Gen.),

$$S \vdash (\forall x)(p \Rightarrow r).$$

Case (i) $x \notin FV(p)$: add

$$((\forall x)(p \Rightarrow r)) \Rightarrow (p \Rightarrow (\forall x) r) \qquad (Ax\ 7)$$
$$p \Rightarrow (\forall x) r \qquad (M.P.)$$

Case (ii) $x \in FV(p)$: we know no hypothesis used in proof of r from $S \cup \{p\}$ had a free occurrence of x, and so p was not used. So we have a proof of r from S. Then add:

$$(\forall x) r \qquad (Gen.)$$
$$(\forall x) r \Rightarrow (p \Rightarrow (\forall x) r) \qquad (Ax\ 1)$$
$$p \Rightarrow (\forall x) r \qquad (M.P.)$$

□

5.4 Completeness

N.B. Proofs in §5.4 are all non-examinable.
Everything else is examinable (statements, definitions, discussion)
Aim: Prove that if $S \subset \mathcal{L}$ is a set of sentences and $p \in \mathcal{L}$ is a sentence then $S \models p$ iff $S \vdash p$.
Parts of some proofs in §5.4 will be fairly sketchy.

Proposition 36 (Soundness Theorem) Let S be a set of sentences and $p \in \mathcal{L}$ a sentence with $S \vdash p$. Then $S \models p$.

*Proof Let $l_1, \dots, l_n$ be a proof of p from S.
It is easy to show by induction that $S \models \check{l}_i$ for $i = 1, \dots, n$
Then $p = \check{p} = l_n = \check{l}_n$. □*

Now for adequacy. Next, and main step: if $S \nvdash \bot$ then $S \not\vdash \bot$.

We say S is consistent if $S \nvdash \bot$ and inconsistent if $S \vdash \bot$.
We aim to show that if S is consistent then S has a model.

Idea: the set of all closed terms 'in' an $\mathcal{L}$-structure is itself an $\mathcal{L}$-structure with the obvious interpretation.

E.g. Theory of Fields (FT) $\Omega = \{+, 0, -, 1, \times\}$
Use some informal notation, e.g. write $1+1$ for $+11$ etc.
Let A be the set of closed terms: e.g.
$(1+1)+1$. Must interpret the function symbols in A.
E.g. $0_A = 0$, $1_A = 1$
$(1+1) +_A (1+1) = (1+1)+(1+1)$

This makes A into an $\mathcal{L}$-structure. But in general, A is not a model of FT.

Problem 0 $FT \vdash 0+0=0$. But in A, $0+0$ and 0 are distinct closed terms. Hence $(0+0=0)_A = 0$.

Solution 0 Just 'make closed terms equal if we can prove they are equal'.
Define equivalence relation $s \sim t$ iff $S \vdash (s=t)$ on A; quotient out by this.

Problem 1 Let $FT_{2,3}$ be the theory of fields of characteristic 2 or 3, i.e.
$$FT_{2,3} = FT \cup \{(1+1=0) \vee (1+1+1=0)\}.$$
Then $FT_{2,3} \vdash 1+1=0$ and $FT_{2,3} \vdash 1+1+1=0$.
So even after quotienting out operation in sol^n 0, A is still

A is still not a model of $FT_{2,3}$.

Solution 1 The problem arises because $FT_{2,3} \nvdash 1+1=0$ but also $FT_{2,3} \nvdash \neg(1+1)=0$.

Define a theory S to be complete if for every sentence p we have $S \vdash p$ or $S \vdash \neg p$.

Sol^n to this problem is to extend S to a complete, consistent theory.

Problem 2 Let $FT_{\sqrt{2}} = FT \cup \{(\exists x)(x \times x = 1+1)\}$

Problem: no closed term t s.t. $FT_{\sqrt{2}} \vdash (t \times t = 1+1)$.

Solution 2 Problem 2 arises because $FT_{\sqrt{2}}$ does not 'have witnesses'

Say a theory T has witnesses if whenever we have a formula p with $T \vdash (\exists x)p$ then there exists a closed term s.t. $T \vdash p[t/x]$

Solve Problem 2 by adding witnesses: whenever $T \vdash (\exists x)p$ add a new constant symbol t and a new axiom $p[t/x]$.

Now we have a consistent theory with witnesses in an expanded language.

Problem ω Adding witnesses can make a complete theory incomplete, while completing a theory with witnesses may give a theory without witnesses.

Solution ω Repeatedly apply Solutions 1, 2 alternately forever

Theorem 37 (Model Existence Lemma)

Let T be a consistent theory. Then T has a model.

* Proof By Zorn, we can extend T to a complete consistent theory T_0 in the same language $\mathcal{L} = \mathcal{L}_0$.

For each formula p of $\mathcal{L}_0$ and each variable x s.t. $T \vdash (\exists x)p$, add a constant symbol t to $\mathcal{L}_0$ and an axiom $p[t/x]$ to T_0. This gives a consistent theory S_1 with witnesses in a language $\mathcal{L}_1$ s.t. $T_0 \subset S_1$ and $\mathcal{L}_0 \subset \mathcal{L}_1$ (i.e. if $\mathcal{L}_0 = (\Omega_0, \Pi_0, \alpha_0)$ and $\mathcal{L}_1 = (\Omega_1, \Pi_1, \alpha_1)$ then $\Omega_0 \subset \Omega_1$, $\Pi_0 \subset \Pi_1$ and $\alpha_1|_{\Omega_0 \cup \Pi_0} = \alpha_0$)

By Zorn, we can extend S_1 to a complete, consistent theory T_1 in the language $\mathcal{L}_1$.

Repeat the above indefinitely to obtain languages

$\mathcal{L} = \mathcal{L}_0 \subset \mathcal{L}_1 \subset \cdots$

and theories $T \subset T_0 \subset S_1 \subset T_1 \subset S_2 \subset T_2 \subset \cdots$
with theories S_i, T_i in language $\mathcal{L}_i$, s.t.
- each S_i is consistent and has witnesses; and
- each T_i is consistent and complete.

Let '$\mathcal{L}_\omega = \bigcup_{i=0}^{\infty} \mathcal{L}_i$' and $T_\omega = \bigcup_{i=0}^{\infty} T_i$.
Now T_ω is a theory in the language $\mathcal{L}_\omega$.
Moreover, it is consistent, complete, and it has witnesses.

↑ proofs are finite — ↑ $T_\omega = \bigcup T_i$, each T_i complete — ↑ $T_\omega = \bigcup S_i$, each S_i has witnesses

Let A_ω be the set of closed terms in the language $\mathcal{L}_\omega$.
Define an equivalence relation on A_ω by $s \sim t$ iff $T_\omega \vdash (s = t)$.
Now let $A = A_\omega / \sim$ (i.e. A is the set of equivalence classes.)
Make A into an $\mathcal{L}_\omega$-structure by defining:
- if $f \in \Omega_\omega$ with $\alpha_\omega(f) = n$, define $f_A : A^n \to A$ by $f_A([t_1]_\sim, \ldots, [t_n]_\sim) = [f t_1 \cdots t_n]_\sim$;
- if $\varphi \in \Pi_\omega$ with $\alpha_\omega(\varphi) = n$, define $\varphi_A \subset A^n$ by $\varphi_A = \{ ([t_1]_\sim, \ldots, [t_n]_\sim) \in A^n \mid T_\omega \vdash \varphi t_1 \cdots t_n \}$.

(N.B. This is well-defined.)
We show A is a model of T_ω (and hence of T) by showing that if p is a sentence in $\mathcal{L}_\omega$ then $p_A = 1$ iff $T_\omega \vdash p$.
Proceed inductively.
- $\bot$: $T_\omega \nvdash \bot$ (consistent) and $\bot_A = 0$ (defn)
- $(s = t)$: $T_\omega \vdash (s = t) \iff s \sim t$ (s, t closed terms)
$\iff [s]_\sim = [t]_\sim$
$\iff s_A = t_A$
$\iff (s = t)_A = 1$
- $\varphi t_1 \cdots t_n$: $T_\omega \vdash \varphi t_1 \cdots t_n \iff ([t_1]_\sim, \ldots, [t_n]_\sim] \in \varphi_A$
$\iff (\varphi t_1 \cdots t_n)_A = 1$

($\varphi \in \Pi_\omega$, $\alpha(\varphi) = n$, $t_1, \ldots, t_n$ closed terms)

- $(q \Rightarrow r)$: $T_\omega \vdash (q \Rightarrow r) \iff T_\omega \vdash \neg q$ or $T_\omega \vdash r$ (yes ☺)
  (q, r sentences) $\iff q_A = 0$ or $r_A = 1$ (ind hyp)
  $\iff (q \Rightarrow r)_A = 1$
- $(\forall x)\, q$: $T_\omega \vdash (\forall x)\, q \iff T_\omega \vdash \neg(\exists x) \neg q$
  (x variable, $FV(q) = \{x\}$ or q sentence) $\iff T_\omega \vdash (\exists x) \neg q$ (complete)
  $\iff$ ~~for all~~ closed terms t, $T_\omega \vdash \neg q[t/x]$ (witnesses)
  for all
  $\iff$ for all closed terms t, $T_\omega \vdash q[t/x]$ (complete)
  $\iff$ for all closed terms t, $q[t/x]_A = 1$ (ind hyp)
  $\iff ((\forall x) q)_A = 1$ (as A is set of closed terms modulo $\sim$)

□*

Corollary 38 (Adequacy theorem)
Let T be a theory and p a sentence with $T \models p$. Then $T \vdash p$.
Proof Follows from Model Existence Lemma exactly as in Ch 1. □

Theorem 39 (Gödel's completeness theorem for first-order logic)
Let T be a theory and p a sentence. Then $T \vdash p$ iff $T \models p$.
Proof Soundness and Adequacy. □

§5.5 Applications of Completeness

Corollary 40 (Compactness Theorem)
Let T be a theory such that every finite subset of T has a model. Then T has a model.
Proof Exactly as in Ch 1. □
However, no decidability theorem — no equivalent of truth tables.
One important consequence of compactness:

Corollary 41 Let T be a theory with arbitrarily large finite models. Then T has an infinite model.

Proof Expand our language by introducing infinitely many constants $c_0, c_1, c_2, \ldots$

Let $T' = T \cup \{ \neg(c_i = c_j) \mid 0 \leq i < j \}$.

Any finite subset of T' only mentions finitely many of the c_i, so has as a model any sufficiently large finite model of T. Hence, by compactness, T' has a model, and this model is an infinite model of T. □

Remark This says e.g. there is no first-order theory of finite groups.

We can go further:

Corollary 42 (Upward Lowenheim-Skolem Theorem)
Let T be a theory with an infinite model and let κ be a cardinal. Then T has a model A with $\text{card}(A) \geq \kappa$.

Proof As Cor. 41 but use constants $\{ c_i \mid i \in I \}$ where $\text{card}(I) = \kappa$. □

Remark It is impossible to axiomatise a specific infinite structure – e.g. there is no first-order theory of $\mathbb{N}$, as any theory with $\mathbb{N}$ as a model also has an uncountable model.

Corollary 43 (Downward Lowenheim-Skolem Theorem)
Let T be a theory in language (Ω, Π, α).
Let κ be an infinite cardinal s.t. $\text{card}(\Omega \cup \Pi) \leq \kappa$.
Then, if T is consistent, it has a model A with $\text{card}(A) \leq \kappa$.

Proof Take A to be the model constructed in the proof of the Model Existence Lemma. □

5.6 Peano Arithmetic

Recall (from IA?) that the Peano axioms specify $\mathbb{N}$ uniquely:

1. $0 \in \mathbb{N}$.
2. Each $n \in \mathbb{N}$ has a successor $n^+ \in \mathbb{N}$.
3. If $m, n \in \mathbb{N}$ and $m \neq n$ then $m^+ \neq n^+$.
4. For all $n \in \mathbb{N}$, $0 \neq n^+$.
5. Induction.

Define addition and multiplication inductively:

$m + 0 = m$	$m \cdot 0 = 0$
$m + n^+ = (m+n)^+$	$mn^+ = mn + m$

We aim to turn this into a first-order theory.

The language of arithmetic is $\mathcal{L} = (\Omega, \Pi, \alpha)$ where $\Omega = \{0, s, a, m\}$, $\Pi = \emptyset$ and

$$\alpha(0) = 0,\ \alpha(s) = 1,\ \alpha(a) = \alpha(m) = 2.$$

Peano arithmetic is the theory PA in the language of arithmetic with the following axioms:

1. $(\forall x)(\forall y)\ ((sx = sy) \Rightarrow (x = y))$
2. $(\forall x)\ \neg(sx = 0)$
3. $(\forall x)(\ axO = x)$
4. $(\forall x)(\forall y)\ (axsy = saxy)$
5. $(\forall x)(\ mx0 = 0)$
6. $(\forall x)(\forall y)\ (mxsy = amxyx)$
7. $(\forall y_1) \cdots (\forall y_n)\ ((p[0/x] \wedge (\forall x)(p \Rightarrow p[sx/x])) \Rightarrow (\forall x)p)$

   for formulas p with $FV(p) = \{x, y_1, \ldots, y_n\}$

Remarks on Axiom 7

1. The variables $y_1, \ldots, y_n$ are parameters: e.g. we might want to prove 'addition is commutative'. We will want to prove by induction on x

   $$(\forall y)(\forall x)(axy = ayx)$$

2. Axiom 7 is not a single axiom. It is an axiom scheme consisting of infinitely many axioms.

Note PA is consistent since it has $\mathbb{N}$ as a model.
$\mathbb{N}$ is infinite. So by Upward L-S, PA has an uncountable model. But we thought the Peano axioms uniquely specified $\mathbb{N}$. What's gone wrong?

Genuine induction: This says if $S \subseteq \mathbb{N}$ is s.t. $0 \in \mathbb{N}$ and $n \in \mathbb{N} \Rightarrow n^+ \in \mathbb{N}$ then we must have $S = \mathbb{N}$.

First order induction (Axiom 7): Same as genuine induction, but only for sets $S \subseteq \mathbb{N}$ that can be specified by a formula in the language of arithmetic.

Now $\mathcal{P}(\mathbb{N})$ is uncountable. But only countably many formulae
So First-order induction says less than genuine induction. So maybe unsurprising that $\mathbb{N}$ is not the only model of PA.
So PA does not specify $\mathbb{N}$ uniquely.

* Next hope: maybe PA can prove every true statement about $\mathbb{N}$ that it can express in the language of arithmetic. (*)

Suppose $p \in \mathcal{L}$ is a sentence. Then $p_{\mathbb{N}} = 1$, or $p_{\mathbb{N}} = 0$ and in latter case, $(\neg p)_{\mathbb{N}} = 1$.
So (*) is equivalent to saying: for any sentence $p \in \mathcal{L}$, either $PA \vdash p$ or $PA \vdash \neg p$. That is, PA is complete.
Unfortunately, PA is not complete.
Liar paradox: 'This sentence is false'.
Aim: find a sentence in the language of arithmetic saying 'This sentence is unprovable'.
Suppose $p \in \mathcal{L}$ saying 'This sentence is unprovable'. Then
$PA \vdash p \Rightarrow p$ true $\Rightarrow PA \nvdash p$ ⁂, and if
$PA \vdash \neg p \Rightarrow p$ false $\Rightarrow PA \vdash p$ ⁂
(as then $PA \vdash \bot$ but PA is consistent)
So $PA \nvdash p$ and $PA \nvdash \neg p$. So PA is incomplete.
How can PA talk about itself?

Our alphabet is $\mathcal{A} = \{ (_1,)_2, \Rightarrow_3, \bot_4, \forall_5, =_6, w_7, '_8, 0_9, s_{10}, a_{11}, m_{12} \}$

Define $c : \not{\mathcal{A}} \longrightarrow \{1, \dots, 12\}$ by $c(() = 1$, $c()) = 2$ etc.

The Gödel number or code of a formula $p \in \mathcal{L}$, where $p = a_1 a_2 \cdots a_n$ $(a_i \in \not{\mathcal{A}})$, is

$$G(p) = \prod_{i=1}^{n} p_i^{c(a_i)}$$

where $p_1, p_2, p_3, \dots$ are the primes in increasing order.

If $\ell = (\ell_1, \dots, \ell_n)$ is a proof in PA, the code of ℓ is

$$\mathcal{G}(\ell) = \prod_{i=1}^{n} p_i^{G(\ell_i)}$$

Now define a (partial) function $f : \mathbb{N} \to \mathbb{N}$ by

$$f(m) = \begin{cases} n & \text{if } m = \mathcal{G}(\ell_1, \dots, \ell_k) \text{ with } G(\ell_k) = n, \\ & \text{undefined otherwise.} \end{cases}$$

That is, $f(m) = n$ iff m codes a proof of a formula whose code is n.

Can we talk about the function f within PA?

A function $f : \mathbb{N} \to \mathbb{N}$ is (PA)-definable if there is some formula $p \in \mathcal{L}$ with two free variables, say $FV(p) = \{x, y\}$, s.t. for all $m, n \in \mathbb{N}$,

$$PA \vdash p[\bar{m}/x, \bar{n}/y] \quad \text{iff} \quad f(m) = n$$

(where for $a \in \mathbb{N}$, $\bar{a}$ is the closed term $\underbrace{ss\cdots s}_{a}0$ of PA)

Fact If f is computable (by an algorithm) then f is PA-definable. (See Johnstone, Ch. 4 – earlier parts of Ch 4 more slowly in AFL)

Remark The converse of this fact is not true.

(Why? (AFL) If $S \subseteq \mathbb{N}$, define the partial indicator φ of S as $\varphi_S : \mathbb{N} \to \mathbb{N}$ given by $\varphi_S(n) = \begin{cases} 1, & n \in S \\ \text{undefined} & \text{o/w} \end{cases}$

Suppose S is recursively enumerable but not recursive.

Then φ_S is computable, so definable by some $p \in \mathcal{L}$, with $FV(p) = \{x, y\}$.

Then $q = (\neg p) \wedge (y = S0)$ defines φ_{S^c}. But φ_{S^c} is not computable.)

Recall $f: \mathbb{N} \to \mathbb{N}$ is defined by

$f(m) = n$ iff $\exists$ proof $(\ell_1, \dots, \ell_k)$
s.t. $\mathcal{G}(\ell_1, \dots, \ell_k) = m$ and $G(\ell_k) = n$.

(N.B. if m doesn't code a proof then $f(m)$ is undefined).

Now f is computable and hence is PA-definable.

Let $\theta \in \mathcal{L}$ with $FV(\theta) = \{x, y\}$ be s.t. for all $m, n \in \mathbb{N}$,
$PA \vdash \theta[\bar{m}/x, \bar{n}/y]$ iff $f(m) = n$.

So θ: 'x codes the proof of a formula with code y'.

For $p \in \mathcal{L}$ define $Pr(p)$ to be the formula
$(\exists x)\, \theta[\overline{G(p)}/y]$.

$Pr(p)$: 'p is provable'.

Lemma 44 Let p be a sentence in $\mathcal{L}$.
Then $PA \vdash p$ iff $PA \vdash Pr(p)$.

Proof Let $n = G(p)$.

"$\Rightarrow$" Suppose $PA \vdash p$. Let m be the code for a proof of p.
Then $f(m) = n$ so $PA \vdash \theta[\bar{m}/x, \bar{n}/y]$
so $PA \vdash (\exists x)\, \theta[\bar{n}/y]$
i.e. $PA \vdash Pr(p)$.

"$\Leftarrow$" Suppose $PA \vdash Pr(p)$, i.e. $PA \vdash (\exists x)\, \theta[\bar{n}/y]$.
Now, $\mathbb{N}$ is a model of PA, so $(\exists x)\, \theta[\bar{n}/y]$ is true in $\mathbb{N}$.
That is, there is some $m \in \mathbb{N}$ s.t. $f(m) = n$.
This means precisely m codes a proof of p.
In particular, there is a proof of p. So $PA \vdash p$. □

Main idea: Look for a formula $p(w)$ with $FV(p(w)) = \{w\}$
s.t. $p(\bar{n})$ says 'n codes a formula $q(w)$ with $FV(q(w)) = \{w\}$
and $q(\bar{n})$ is unprovable'.

Let $n = G(p(w))$ and let $g = p(\bar{n})$.

Then g says: 'n codes a formula $q(w)$ with $FV(q(w)) = \{w\}$
and $q(\bar{n})$ is unprovable',

i.e. '$p(\bar{n})$ is unbrovable', i.e. 'g is unprovable'.

Theorem 45 (Gödel's First Incompleteness Theorem)
PA is incomplete

Proof Let $h : \mathbb{N} \to \mathbb{N}$ be defined by $h(m) = n$ iff there is some formula $q \in \mathcal{L}$ s.t. $FV(q) = \{w\}$, $m = G(q)$ and $n = G(q[\overline{m}/w])$
Then h is computable and thus definable. So let $\psi \in \mathcal{L}$ with $FV(\psi) = \{w, y\}$ be s.t. for all $m, n \in \mathbb{N}$,
$PA \vdash \psi[\overline{m}/w, \overline{n}/y]$ iff there exists $q \in \mathcal{L}$ with $FV(q) = \{w\}$ s.t. $G(q) = m$ and $G(q[\overline{m}/w]) = n$.
Recall that there is a formula $\theta \in \mathcal{L}$ with $FV(\theta) = \{x, y\}$ s.t. for all $m, n \in \mathbb{N}$, $PA \vdash \theta[\overline{m}/x, \overline{n}/y]$ iff m codes the proof of a formula with Gödel number n.
Let φ be the formula $(\exists y)(\psi \wedge \neg(\exists x)\theta)$.
Note $FV(\varphi) = \{w\}$.
For $m \in \mathbb{N}$, $\varphi[\overline{m}/w]$ says:
'm codes a formula q with $FV(q) = \{w\}$,
$G(q[\overline{m}/w]) = y$,
and the formula with Gödel number y is unprovable'.
That is: 'm codes a formula q with $FV(q) = \{w\}$,
and $q[\overline{m}/w]$ is unprovable.'
Let $g = \varphi[\overline{G(\varphi)}/w]$.
So g says: '$\varphi[\overline{G(\varphi)}/w]$ is unprovable'
i.e. 'g is unprovable'.
Then g is a sentence, and
$PA \vdash \neg g \iff PA \vdash Pr(g) \iff PA \vdash g$.
Hence $PA \nvdash g$, $PA \nvdash \neg g$ (since PA is consistent). □

So PA is incomplete — can we complete it?
⊛ Note g is true in $\mathbb{N}$: if g is false then g is provable, so as $\mathbb{N}$ is a model of PA, g is true ※
So let's take $PA^+ = PA \cup \{g\}$.

This still has model $\mathbb{N}$ and so is consistent — is it complete?
No! — just run the same proof as before on PA^+.
So PA is incomplete in an 'irreparable' way.
What about the theory:

$$PA! = \{ p \in \mathcal{L} \text{ is a sentence} \mid P_{\mathbb{N}} = 1 \}.$$

This is consistent, with $\mathbb{N}$ as a model.
Obviously complete. Where does the proof of Theorem 45 fail?
Functions are not definable.
No algorithm to determine if a given formula p is an axiom of PA!
'Truth is undefinable'.
At ⊛ we proved that g, thought of as a statement about $\mathbb{N}$, is true. Why does this not formalise into a proof within PA?
We used the fact that PA is consistent.
Let $Con(PA)$ to be the formula $\neg Pr(\bot)$.
So $Con(PA)$ says 'PA is consistent'.
Seems: $PA \vdash (Con(PA) \Rightarrow g)$.
But $PA \nvdash g$.
So $PA \nvdash Con(PA)$.
'PA does not prove its own consistency'.

Theorem 46 (Gödel's Second Incompleteness Theorem)
$PA \nvdash Con(PA)$

Proof Given a sentence $p \in \mathcal{L}$, given a proof of p, then we can algorithmically construct a proof of $Pr(p)$. Thus:

$$PA \vdash (Pr(p) \Rightarrow Pr(Pr(p))).$$

Now take $p = g$ as in the first incompleteness thm.
Now: $PA \vdash (Pr(g) \Rightarrow Pr(Pr(g)))$
i.e. $PA \vdash (Pr(g) \Rightarrow Pr(\neg g))$
Thus: $PA \vdash (Pr(g) \Rightarrow Pr(\bot))$
so $PA \vdash (\neg Pr(\bot) \Rightarrow \neg Pr(g))$

i.e. $PA \vdash (Con(PA) \Rightarrow g)$.

But $PA \vdash g$ so $PA \vdash Con(PA)$. □

Remarks 1. All of this works if PA is replaced by any 'suitable' theory T: want

- T is consistent;
- T has $\mathbb{N}$ as a model;
- T is 'sufficiently powerful';
- T is 'recursively presented' — there is an algorithm to determine if a given formula is an axiom of T.

2. Gödel proved rather more than we have done in this section. *

Ch 6 Set Theory

6.1 Introduction

Example The principle of comprehension says if $p(x)$ is a property then the collection of all things x that have the property $p(x)$ is a set: $\{x \mid p(x)\}$.

This leads to paradoxes, e.g.

Russell's Paradox Then $A = \{x \mid x \notin x\}$ is a set.

But $A \in A \iff A \notin A$ ※

Maybe we should forbid a set being a member of itself?

Still get paradoxes:

Burali-Forti Paradox $\{x \mid x$ is an ordinal$\}$ is not a set.

- What can we validly do with sets without creating paradoxes?
- What does the universe of sets look like?

We shall develop set theory as a first-order theory as in Chapter 5.

The language of sets is $\mathcal{L} = (\Omega, \Pi, \alpha)$ with $\Omega = \emptyset$, $\Pi = \{\in\}$, $\alpha(\in) = 2$. "epsilon"

If V is a model of set theory, we want to think of every element of V as being a set.

Suppose $A = \{f \mid f : \mathbb{N} \to \mathbb{N}\}$.

Looks like A is a set. (Comprehension doesn't work but when we state our axioms, will be enough to make A a set).

Suppose $f : \mathbb{N} \to \mathbb{N}$. Is f a set? A priori, no:

a function from $\mathbb{N}$ to $\mathbb{N}$ is something that associates with each $n \in \mathbb{N}$ some element $f(n) \in \mathbb{N}$

But can define a function as a set: could say

$f : \mathbb{N} \to \mathbb{N}$ iff $f \subset \mathbb{N} \times \mathbb{N}$ s.t. for all $n \in \mathbb{N}$, there is a unique $m \in \mathbb{N}$ with $(n, m) \in f$.

Write $m = f(n)$.

Are elements of f sets? i.e. is an ordered pair a set?

A priori, no. But:

How should ordered pairs behave? Need

$(x,y) = (z,t) \iff (x = z \text{ and } y = t)$. ⊛

So define the Wiener–Kuratowski ordered pair:

$(x,y) = \{\{x\}, \{x,y\}\}$.

Exercise Check that with this definition, ⊛ holds.

Could have made lots of other definitions that work.

But will stick with this one. Henceforth, 'ordered pair' means 'Wiener–Kuratowski' ordered pair.

Can we make the natural numbers into sets?

Yes: e.g. $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{1\} = \{\{\emptyset\}\}$,

$3 = \{2\} = \{\{\{\emptyset\}\}\}$, ...

Lots of different ways to do this. Another turns out to be more convenient.

The <u>von Neumann natural numbers</u> are defined by

$0 = \emptyset$; $n^+ = n \cup \{n\}$, the successor of n.

So $n = \{0, 1, \ldots, n-1\}$.

Henceforth, 'natural number' will mean 'von Neumann natural number'.

6.2 The axioms of Zermelo–Fraenkel set theory

<u>Zermelo–Fraenkel set theory</u> is the theory ZF in the language of sets with the following axioms.

I. <u>The axiom of extensionality</u> says 'if two sets have the same members then they are the same set':

$(\text{Ext}): (\forall x)(\forall y)\left((\forall z)\left((z \in x) \iff (z \in y)\right)\right) \Rightarrow (x = y)$

<u>Remark</u> The converse is not necessary as an axiom of ZF as it follows from the logical axioms.

II. <u>The axiom of separation</u> is like the principle of comprehension but restricted to subsets of a set: 'given a set, the collection of its elements with a given property is also a set'.

$(\text{Sep}): (\forall t_1)\cdots(\forall t_n)(\forall x)(\exists y)(\forall z)\left((z \in y) \iff ((z \in x) \wedge p)\right)$

where $p \in \mathcal{L}$ with $FV(p) = \{z, t_1, \ldots, t_n\}$.

Remarks 1. The $t_1, \ldots, t_n$ are parameters (cf. induction in PA)

2. This is an axiom scheme rather than a single axiom.

3. Notation: denote the set y defined in (Sep) by

$$\{z \in x \mid p\}$$

III. The empty set axiom says 'there is a set with no elements'

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

Remark (Ext) tells us that there is a unique empty set (once we know one exists)

Denote the empty set by $\emptyset$. N.B. is not formally part of the language but is an abbreviation.

So, e.g., '$\emptyset \in z$' formally means

$$(\exists x)\big((\forall y)\, \neg(y \in x) \wedge (x \in z)\big).$$

IV. The pair-set axiom says 'given any two sets x, y the set $\{x, y\}$ exists'.

(Pair) $(\forall x)(\forall y)(\exists z)(\forall u)\Big((u \in z) \Leftrightarrow \big((u = x) \vee (u = y)\big)\Big)$.

Remarks 1. Write $\{x, y\}$ for the set z constructed by the axiom.

2. Putting $y = x$, for any x there is a set whose only element is x. Write $\{x\}$ for this set.

3. What about e.g. $\{x, y, z\}$? Will follow from (Pair) and later axioms.

Now, given any x, y, can define the ordered pair

$$(x, y) = \{\{x\}, \{x, y\}\}$$

using (Pair) thrice.

Write 'x is an ordered pair' to mean

$$(\exists y)(\exists z)\,(x = (y, z))$$

Write 'f is a function' to mean

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

Write '$f(x) = y$' for $(f \text{ is a function}) \wedge ((x, y) \in f)$.

Write '$x = \mathrm{dom} f$' for

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

Write '$f : x \to y$' for

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

V. The union axiom says 'we can take the union of all sets that are elements of a given set'.

(Un) $(\forall x)(\exists y)(\forall z)\big((z \in y) \Leftrightarrow (\exists w)((z \in w) \wedge (w \in x))\big)$.

Remarks 1. Write $\bigcup x$ for the set y constructed in (Un). If e.g. $x = \{a, b\}$ write $a \cup b = \bigcup x = \bigcup \{a, b\}$.

2. Don't need an intersection axiom. If $x \neq \emptyset$ can form $\bigcap x$ as a subset of any $y \in x$ using (Sep).

good example of (Sep) with param →

VI. The power set axiom says 'we can form power sets'

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

where '$z \subset x$' means $(\forall w)((w \in z) \Rightarrow (w \in x))$.

Remarks. 1. Write $\mathcal{P} x$ for the set y in (Pow).

2. Can now define the Cartesian product

$x \times y \subset \mathcal{P}\mathcal{P}(x \cup y)$ using (Sep).

Similarly, can define $\{f \mid f : x \to y\} \subset \mathcal{P}(x \times y)$

Recall We want to be able to consider $\mathbb{N}$ within our set theory.

For x a set, define $x^+ = x \cup \{x\}$. This is a set:

$$x^+ = \bigcup \{ x, \{x\} \} \qquad \circledast$$

(by applying (Pair) twice and (Un) once).

Using this and (Emp) we now have the von Neumann natural numbers: $0 = \emptyset$ and $n+1 = n^+$.

So in general, $n = \{0, 1, 2, \dots, n-1\}$.

(Remark ~~This~~ A similar approach would also show if $x_1, \dots, x_k$ are sets then $\{x_1, \dots, x_k\}$ is a set.

e.g. $\{x, y, z\} = \bigcup \{ \{x, y\}, \{z\} \}$)

But the von Neumann natural numbers form a set?

Not necessarily, given only Axioms I to VI.

Indeed, need not have any infinite sets at all. (†)

Now, by $\circledast$, any model of axioms I to VI must be infinite.

Doesn't that contradict (†)?

No: · External observer: 'V is an infinite set'.

· Internally: 'V is not a set'

So need another axiom to say there is an infinite set.

Say 'x is a successor set' to mean

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

VII The axiom of infinity says 'There exists a successor set':

(Inf) $(\exists x)$ (x is a succesor set)

Remarks 1. Any intersection of successor sets is a successor set, so once (Inf) has told us there is some successor set then there must be a smallest successor set ω, the set of natural numbers.

2. Can check that (with suitable interpretation), ω satisfies the axioms of PA. Moreover,

$$ZF \vdash (\forall x)((x \text{ is a successor set}) \wedge (x \subset \omega)) \Rightarrow (x = \omega)$$

— genuine induction, not just first-order induction

3. Say 'x is a natural number' to mean '$x \in \omega$';
'x is finite': $(\exists y \in \omega)$(x bijects with y);
'x is infinite': $\neg$(x is finite);
'x is countable': (x is finite) $\vee$ (x bijects with ω).
(Used some new abbreviations here: 'x bijects with y'
and $(\exists u \in v)p$, the latter meaning
$(\exists u)((u \in v) \wedge p)$.)

Defined $0, 1, 2, \ldots$ and also ω as sets.
Could try to define <u>von Neumann ordinals</u> with this as our starting point.
<u>Recall</u> $\alpha \in Ord$ then $\alpha \cong I_\alpha = \{\beta \in Ord \mid \beta < \alpha\}$.
So <u>aim</u> define v.N. ordinals s.t. $\alpha = \{\beta \mid \beta \text{ Ord}, \beta < \alpha\}$.
Works so far: if $n \in \omega$, $n = \{0, 1, \ldots, n-1\}$ ✓
and $\omega = \{0, 1, 2, \ldots\}$ ✓
Continue $\omega + 1 = \omega \cup \{\omega\} = \omega^+$ ✓
$\omega + 2 = (\omega + 1)^+$ ✓
In general: $\omega + n + 1 = (\omega + n)^+$ $\quad (n \in \omega)$
Note each such x is <u>transitive</u>:
$(\forall y)(\forall z)\big((z \in y) \wedge (y \in x)\big) \Rightarrow (z \in x)$
(Equivalently $(\forall y)\big((y \in x) \Rightarrow (y \subset x)\big)$)
and totally ordered by $\in$.
<u>Definition</u> Say 'x is a von Neumann ordinal' to mean 'x is transitive and totally ordered by $\in$'.
Eventually: show that these correspond to ordinals.
For now: have v.N. ordinals n and $\omega + n$ for all $n \in \omega$.
What about $\omega + \omega$?
Want to define $\omega + \omega = \omega \cup \{\omega + n \mid n \in \omega\}$
But even (Inf) is not enough to ! say this is a set.

Idea Start with $\omega = \{n \mid n \in \omega\}$ and for each n, 'replace' it with $\omega + n$ to get the set we want.

So our axiom should say: 'The image of a set is a set' ← under f^n ? Not quite — need some sort of 'function-type' thing

(Since before we know the image is a set, also don't know this thing is a function).

Digression on classes

Let V be a model of ZF. We shall define a class to be a 'subset' of the model given by a particular property (but may not correspond to a set in theory)

A class is a subset $W \subset V$ defined by a formula $p \in \mathcal{L}$. That is to say, p has free variable, x, say, and given $a \in V$, we have $a \in W$ iff $p[\bar{a}/x]$ is true in the model V

where we extend $\mathcal{L}$ to a language $\mathcal{L}'$ by adding a constant symbol $\bar{a}$ and make V an $\mathcal{L}'$-structure by interpreting $\bar{a}_V = a$.

We also allow parameters — so can have $FV(p) = \{x, t_1, \ldots, t_n\}$.

Important note A class is not a 'thing' in the theory of ZF. Theorems of ZF would talk about the formula p rather than the class W.

Examples: 1. V is a class: take p to be $(x = x)$

2. For each $b \in V$, the collection of sets of which b is a member is a class: take p to be $(t \in x)$

($FV(p) = \{x, t\}$ where t is a parameter, intended to be interpreted as b)

3. Every set 'is' a class: take p to be $(x \in t)$

(t parameter)

However, converse is not true — e.g. V is not a set.

A class defined by a formula p with $FV(p) = \{x\}$ is a function-class if

$$ZF \vdash (\forall x)\big(p \Rightarrow (x \text{ is an ordered pair})\big)$$
$$\wedge\ (\forall u)(\forall v)(\forall w)\Big(\big(p[(u,v)/x] \wedge p[(u,w)/x]\big) \Rightarrow v = w\Big)$$

Equivalently, and more conveniently, we can consider a function-class F to be defined by the formula φ with $FV(\varphi) = \{x, y\}$ where φ is

$$((\exists z)(p[z/x]) \wedge (z = (x, y)))$$

(So φ 'says' $F(x) = y$)

Again, allow parametes $t_1, \ldots, t_n$.

VIII The axiom of replacement says 'the image of a set under a function class is a set'.

$$(\text{Rep}) \quad \underbrace{(\forall t_1)\cdots(\forall t_n)}_{\text{params}}\Big[\Big(\underbrace{(\forall x)(\forall y)(\forall z)\Big((p \wedge p[z/y]) \to (y = z)\Big)}_{p \text{ is fn class}} \Big) \Rightarrow \underbrace{(\forall x)(\exists y)(\forall z)\left((z \in y) \Leftrightarrow (\exists t \in x)\, p[t/x, z/y]\right)}_{\substack{y \text{ is image of } x \\ \text{under } F \text{ defined by } p}}\Big]$$

where $p \in \mathcal{L}$ with $FV(p) = \{t_1, \ldots, t_n, x, y\}$.

Recall The axiom of replacement says that the image of a function-class is a set under a set. ☆

Example $\omega + \omega$ For $n \in \omega$, define $F(n) = \omega + n$. Then $\omega + \omega$ is the set $\omega \cup y$ where y is the set obtained by applying replacement to the set ω and the function-class F, provided F is a function class.

↖ IMPORTANT !!!

Is F a function-class on ω? Yes — $F(x) = y$ is defined by the formula $\varphi \in \mathcal{L}$ given by

$(x \in \omega) \wedge (y \text{ is an ordinal}) \wedge (\omega \subset y) \wedge (y \setminus \omega \text{ is IM to } x)$

Finally, we want to ban things like '$\underbrace{x \in x}_{*}$' or more generally e.g. '$x \in y \in z \in x$'. } **

Our picture of the set-theoretic universe is that it starts with $\emptyset$ and new sets are built up from old sets using our set-building rules.

The final axiom aims to capture this.

IX The axiom of foundation says 'every non-empty set has an $\in$-minimal member':

(Fdn) $(\forall x)\left(\neg(x = \emptyset) \Rightarrow (\exists y \in x)(y \cap x = \emptyset)\right)$

Remark For example, (*) is banned by applying (Fdn) to the set $\{x\}$, while (**) is banned by applying (Fdn) to the set $\{x, y, z\}$.

We're now done: ZF is the theory consisting of axioms I to IX.

6.3 The Axiom of Choice

Recall the axiom of choice:

(AC) $(\forall x)\Big(\big((\forall y \in x)\neg(y=\emptyset)\big) \Rightarrow \big((\exists f)(f \text{ is a function}) \wedge (\operatorname{dom} f = x) \wedge (\forall y \in x)(f(y) \in y)\big)\Big)$

Note AC is not part of ZF : $AC \notin ZF$

Indeed, $ZF \nvdash AC$.

The theory ZFC is defined to be $ZFC = ZF \cup \{AC\}$.

It can be shown that while $ZF \nvdash AC$, the theory ZFC is consistent.

(Well, assuming ZF is consistent : if $ZF \vdash \bot$ then $ZF \vdash AC$ and $ZFC \vdash \bot$.

Is ZF consistent? See later. But assume for now that it is.)

Recall ZL: Zorn's Lemma

WOP : Well-Ordering Principle

Inj : $(\forall x)(\forall y)(\exists f)\ (f \text{ is an injection}) \wedge \big((f: x \to y) \vee (f: y \to x)\big)$

We showed : $ZFC \vdash ZL, WOP, Inj$.

We also 'showed' AC is required to prove any of ZL, WOP or Inj by assuming one of them and proved AC.

So we showed : $ZF \cup \{p\} \vdash AC$

if $p \in \{ZL, WOP, Inj\}$.

For rest of this chapter : unless otherwise stated, we are working in ZF, not ZFC.

6.4 $\in$-induction and $\in$-recursion (:)

(Fdn) suggests that we should be able to 'induct' over $\in$. That is to say, if we want to prove $(\forall x)p$, $FV(p)=\{x\}$ then when proving it for x, we can assume it for all $y \in x$ (i.e. when proving p can assume $(\forall y \in x)\, p[y/x]$.

Why does this work? Essentially the following.

Suppose have $p \in \mathcal{L}$ with $FV(p)=\{x\}$ for which

$$(\forall x)\,((\forall y \in x)\, p[y/x]) \Rightarrow p$$

but $\neg(\forall x)\, p$. ⚠

Let $z = \{x \mid \neg p\}$. Then $z \neq \phi$ so by (Fdn), z has an $\in$-minimal member. Then $\neg p[y/x]$ but $(\forall w \in y)\, p[w/x]$. ※

Problem This is nonsense as there is no reason to believe the set z exists.

Solution For each x, restrict to working within a set that has everything we need to prove p for that particular x.

First, we show such a set exists.

Theorem 47 Let x be a set. Then there is a set

$$TC(x) = x \cup \bigcup x \cup \bigcup\bigcup x \cup \bigcup\bigcup\bigcup x \cup \dots$$

Remark We mean 'this is a theorem of ZF' or, equivalently, this holds in every model of ZF'.

Proof $TC(x) = \bigcup \{x, \bigcup x, \bigcup\bigcup x, \bigcup\bigcup\bigcup x, \dots\}$

obtained by (Un) applied to $\{x, \bigcup x, \bigcup\bigcup x, \bigcup\bigcup\bigcup x, \dots\}$, which is obtained by applying (Rep) to the set ω and the function-class $F(n) = \underbrace{\bigcup \cdots \bigcup}_{n} x$

provided F is indeed a function-class.

We shall prove that there is a unique function-class defined on ω with $F(0)=x$ and $F(n^+) = \bigcup F(n)$ for all $n \in \omega$.

Say 'f is an attempt' to mean

$$(f \text{ is a function}) \wedge (\operatorname{dom} f \in \omega) \wedge (f(0) = x) \wedge (\forall y \in \omega)\big((y^+ \in \operatorname{dom} f) \Rightarrow (f(y^+) = \bigcup f(y))\big).$$

We need:

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

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

These are both easy by ω-induction.

So can define $F(x) = y$ by the formula

$$(\exists f)\big((f \text{ is an attempt}) \wedge (x \in \operatorname{dom} f) \wedge (f(x) = y)\big)$$

Finally, if G is another such function-class, it is easy to show by ω-induction that

$$(\forall x \in \omega)\ (F(x) = G(x)). \qquad \square$$

Recall that x is <u>transitive</u> if $(\forall y \in x)(\forall z \in y)(z \in x)$

Now $TC(x)$ is transitive and $x \subset TC(x)$.

Moreover, if $x \subset y$ and y is transitive then $TC(x) \subset y$.

So $TC(x)$ is the <u>smallest</u> transitive set containing x as a subset, the transitive closure of x.

<u>Theorem 48</u> (Principle of $\in$-induction)

Let $p \in \mathcal{L}$, with $FV(p) = \{x, t_1, t_2, \ldots, t_n\}$. Then

$$(\forall t_1) \cdots (\forall t_n)\big((\forall x)\big((\forall y \in x)\, p[y/x] \Rightarrow p\big) \Rightarrow (\forall x)\, p\big).$$

<u>Remark</u> This is a family of theorems of ZF, one for each $p \in \mathcal{L}$ with $FV(p) = \{x, t_1, \ldots, t_n\}$.

<u>Proof</u> Given $t_1, \ldots, t_n$, suppose

$$(\forall x)\big((\forall y \in x)\, p[y/x] \Rightarrow p\big) \quad \text{but} \quad \neg(\forall x)\, p.$$

Then there is some set x for which $\neg p$.

Let $z = \{ y \in TC(\{x\}) \mid \neg p[y/x] \}$.

Now $z \neq \emptyset$ (as $x \in z$) so by (Fdn) z has an ε-minimal member y. Now $\neg p[y/x]$ but

$(\forall w \in y)\ p[w/x]$ ⁂ □

Theorem 49 ($\in$-recursion) Let G be an everywhere-defined function-class. Then there is a unique everywhere-defined function-class F s.t. $(\forall x)(F(x) = G(F|_x))$.

Remark This is again many theorems of ZF, one for each $p \in \mathcal{L}$ defining a function class G.

Proof Note the statement of the theorem does make sense, $F|_x$ is a set by (Rep).

Define 'f is an attempt' to mean

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

It is easy to show by $\in$-induction that

$$(\forall x)(\forall f)(\forall g)\Big(\big((f \text{ is an attempt}) \wedge (g \text{ is an attempt}) \wedge (x \in \operatorname{dom} f \wedge \operatorname{dom} g)\big) \Rightarrow (f(x) = g(x))\Big) \quad (*)$$

We next show by $\in$-induction that

$(\forall x)(\exists f)((f \text{ is an attempt}) \wedge (\operatorname{dom} f = TC(\{x\})))$:

Fix x. By ind hyp and $(*)$, for each $y \in x$ there is a unique attempt f_y with $\operatorname{dom} f_y = TC(\{y\})$.

Now define $\bar{f} = \bigcup_{y \in x} f_y$ and $f = \bar{f} \cup \{(x, G(\bar{f}|_x))\}$.

Hence we may define $F(x) = y$ by the formula:

$(\exists f)((f \text{ is an attempt}) \wedge (f(x) = y))$

Suppose finally that the function-class E also works. Then it is easy to show by $\in$-induction that

$(\forall x)(F(x) = E(x))$. □

We can think of $\in$ as a relation-class, i.e. a class of ordered pairs. What properties of the relation-class $\in$ did we use in proving Theorems 48, 49?

Definition Let R be a relation-class. We say R is

- local if $(\forall x)(\exists y)(\forall z)((z \in y) \Leftrightarrow (z R x))$;
- well-founded if $(\forall x)\big((\neg(x = \emptyset)) \Rightarrow (\exists y \in x)(\forall z \in x)\neg(z R y)\big)$

Observe that $\in$ is local (trivial) and well-founded (by (Fdn))

Theorem 50 (R-induction and R-recursion)

Let R be a local, well-founded relation class. Then:

(a) Let $p \in \mathcal{L}$ with $FV(p) = \{x, t_1, \ldots, t_n\}$. Then

$$(\forall t_1)\cdots(\forall t_n)\Big(\big((\forall x)\big((\forall y)((yRx) \Rightarrow p[y/x])\big) \rightarrow p\big) \Rightarrow (\forall x)\, p\Big);$$

(b) Let G be an everywhere-defined function class.

Then there is a unique everywhere-defined function class F s.t. $(\forall x)\big(F(x) = G(F|_{\{y \mid yRx\}})\big)$.

Proof As R is local, we can mimic the proof of Lemma 47 to show that there is a unique function class F defined on ω with $F(\phi) = x$ and $(\forall n \in \omega)\big(F(n^+) = \{y \mid (\exists z \in F(n))(yRz)\}\big)$.

Then define $TC_R(x) = \bigcup y$ where y is the image of ω under F.

(a) Next, copy the proof of Theorem 48, replacing $\in$ by R, and $TC(\{x\})$ by $TC_R(\{x\})$, and the use of (Fdn) by well-foundedness of R.

(b) Finally, $\{y \mid yRx\}$ is a set (as R local), so if F is a function class, then $F|_{\{y \mid yRx\}}$ is a set by (Rep), so $G(F|_{\{y \mid yRx\}})$ does make sense as before.

So now copy the proof of Theorem 49, replacing transitivity by 'R-transitivity' (where 'x is R-transitive' means $(\forall y \in x)(\forall z)\big((zRy) \Rightarrow (z \in x)\big)$), and $\in$-induction by R-induction. □

Recall that a von-Neumann ordinal is a transitive set that is totally ordered by $\in$.

equivalently, 'well-ordered' since $\in$ well-founded

We want: every well-ordered set x is isomorphic to a unique v.N. ordinal.

We prove something more general.

Definition Let R be a relation-class. We say R is *extensional* if $(\forall x)(\forall y)\big(((\forall z)((zRx) \iff (zRy))) \Rightarrow (x = y)\big)$

N.B. By (Ext), $\in$ is extensional.

Theorem 51 (Mostowski's collapsing theorem)
Let r be a well-founded, extensional relation on a set a. Then there exists a transitive set b and a bijection $f : a \to b$ s.t. $(\forall x, y \in a)\big((x \,r\, y) \iff (f(x) \in f(y))\big)$.
Moreover, b and f are unique.

Proof Note r is local by (Rep) and (Sep) since it is a relation, not just a relation-class:
$\{y \mid y \,r\, x\} = \{y \in TC_R(x) \mid y \,r\, x\}$. ← what ??
Now, by r-recursion, define for $x \in a$,
$f(x) = \underbrace{\{f(y) \mid y \,r\, x\}}_{\text{a set by (Rep)}}$

Note f is a function, not just a function-class, by (Rep).
(f is an image of the set a)
Let $b = \{f(x) \mid x \in a\}$ (a set by (Rep)).
Clearly b is transitive and f is surjective.
Claim $(\forall x \in a)(\forall y \in a)\big((f(x) = f(y)) \Rightarrow (x = y)\big)$.
Proof By r-induction on x. Suppose $x, y \in a$, with $f(x) = f(y)$. Now:

$\{f(u) \mid u \,r\, x\} = \{f(v) \mid v \,r\, y\}$ (by def^n of f)
so $\{u \mid u \,r\, x\} = \{v \mid v \,r\, y\}$ (by r-induction)
so $x = y$ (r extensional) $=$

Thus f is injective, so also bijective, and then
$x \,r\, y \iff f(x) \in f(y)$.
Finally, suppose b' and $f' : a \to b'$ also work. Then by r-induction
$(\forall x \in a)\big(f(x) = f'(x)\big)$
so $f = f'$, and so $b' = \{f'(x) \mid x \in a\} = \{f(x) \mid x \in a\} = b$. □

Corollary 52 Any well-ordered set is isomorphic to a unique von-Neumann ordinal.

Proof Apply Theorem 51 to a set a with a well-ordering r on a (noting that a well-ordering is well-founded and extensional).
Then, as a poset, (a, r) is isomorphic to a unique transitive set b ordered by $\in$. As a is totally ordered by r, thus b is totally ordered by $\in$. So b is a v.N. ordinal. □

Remark This shows that the v.N. ordinals have the property required by the ordinals.
Thus we can define the ordinals to be the v.N. ordinals, and everything in the previous theory works within ZF.
Thus if we are working in ZFC, we now also have a formal definition of cardinals as initial ordinals.
Still outstanding: formal definition of cardinals within just ZF (i.e. not assuming AC).

5.6 The Universe of Sets

Recall: our picture of the universe of sets is 'start with $\emptyset$ and keep building'.

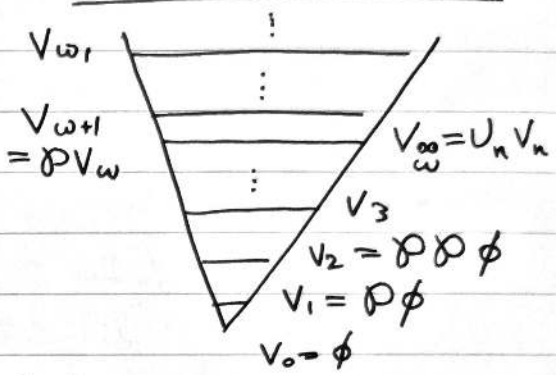
$\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}\}$, ...

$\wp(\emptyset)$

$\wp\wp(\emptyset)$

$\wp\wp\wp(\emptyset)$

Picture of the Universe



Write Ord for the class of ordinals.

Recursively define sets V_α for $\alpha \in \text{Ord}$ by:

$\alpha = 0$	$V_0 = \emptyset$
$\alpha = \delta^+$	$V_{\delta^+} = \wp V_\delta$
α ltd	$V_\alpha = \bigcup_{\delta<\alpha} V_\delta$

Recursively? Is this $\in$-recursion or ordinal recursion?

Doesn't matter: if $\alpha, \beta \in \text{Ord}$ then $\alpha < \beta \iff \alpha \in \beta$.

So ordinal induction / recursion are special case of $\in$-induction and $\in$-recursion.

Write V for the class of all sets

Aim: '$V = \bigcup_{\alpha \in \text{Ord}} V_\alpha$'

↑ not a set

Definition If $x \subset V_\alpha$ for some $\alpha \in \text{Ord}$, then there is a least such $\alpha = \text{rk}(x)$, the rank of x.

Then rk is a function-class; we shall eventually show that rk is everywhere-defined.

Lemma 53 $\forall \alpha \in Ord$, V_α is transitive

Pf Induction on α.

Clearly $\emptyset$ is transitive, and any union of transitive sets is transitive.

So assume $\alpha = \delta^+$.

Now, by ind hyp, V_δ is transitive, and $V_\alpha = \wp V_\delta$.

Let $x \in y \in \wp V_\delta$. Then $y \subset V_\delta$ so $x \in V_\delta$.

But V_δ is transitive so $x \subset V_\delta$, i.e. $x \in \wp V_\delta$.

So V_α is transitive. □

Lemma 54 $(\forall \alpha, \beta \in Ord)((\alpha \leq \beta) \Rightarrow (V_\alpha \subset V_\beta))$

Pf Induction on $\beta \geq \alpha$ (α fixed).

Obvious if $\beta = \alpha$ or β nzl. So assume $\beta = \delta^+$, where $\delta \geq \alpha$.

Then $V_\beta = \wp V_\delta$ where $V_\alpha \subset V_\delta$ by ind. hyp.

Let $x \in V_\alpha$. Then $x \in V_\delta$, but V_δ is transitive so $x \subset V_\delta$, i.e. $x \in \wp V_\delta = V_\beta$.

Hence $V_\alpha \subset V_\beta$. □

Theorem 55 $(\forall x)(\exists \alpha \in Ord)(x \subset V_\alpha)$

(and so $x \in \wp V_\alpha = V_{\alpha+}$).

Thus the function-class rk is everywhere-defined.

Moreover, $(\forall x)\left(rk(x) = \sup\{rk(y)^+ \mid y \in x\}\right)$. (*)

Proof By $\in$-induction. Fix x.

By ind hyp., for all $y \in x$, $rk(y)$ is defined, with $y \subset V_{rk(y)}$ and so $y \in V_{rk(y)^+}$.

Now set $\alpha = \sup\{rk(y)^+ \mid y \in x\}$.

We have $(\forall y \in x)(y \in V_\alpha)$ and so $x \subset V_\alpha$.

Suppose $\beta < \alpha$. Then $(\exists y \in x)(rk(y)^+ > \beta)$.

So $rk(y) \geq \beta$. Then $y \notin V_\beta$ but $y \in x$ so $x \notin V_\beta$.

yes! Hence $rk(x) = \alpha$. □

Remark In general, (*) tends to be the best method for calculating ranks in practice.

Example $(\forall \alpha \in \mathrm{Ord})(\mathrm{rk}(\alpha) = \alpha)$

inductively $\mathrm{rk}(\alpha) = \sup\{\mathrm{rk}(\beta)^+ \mid \beta \in \alpha\}$ (yes)

$= \sup\{\beta^+ \mid \beta \in \alpha\} = \alpha$ ∴

Aside on cardinals. We want to define cardinals in ZF, i.e. without using AC.

Required property: $(\forall x)(\forall y)\big((\mathrm{card}\, x = \mathrm{card}\, y) \Leftrightarrow (x, y \text{ biject})\big)$.

Given x, let α be the least ordinal s.t. some set of rank α bijects with x. Now define

$$\mathrm{card}(x) = \{y \in V_{\alpha^+} \mid y \text{ bijects with } x\}$$

* 6.6 Completeness and Consistency of ZF ∴

(Sketch - Ch 9 of Johnstone)

Is ZF consistent? Is it a complete theory?

As in Ch. 5, can Gödel-number etc

Given p in language of sets, get sentence $\mathrm{Pr}_{ZF}(p)$ in language of arithmetic saying '$ZF \vdash p$'.

Also, PA can be interpreted within ZF, so can think of $\mathrm{Pr}_{ZF}(p)$ as a sentence in the language of sets.

So can construct a sentence g in language of sets s.t. g says '$ZF \nvdash g$'.

Now, $ZF \vdash g \Rightarrow PA \vdash \mathrm{Pr}_{ZF}(g)$

$\Rightarrow ZF \vdash \mathrm{Pr}_{ZF}(g)$

$\Rightarrow ZF \vdash \neg g$ ※

(assuming for now that $ZF \nvdash \bot$)

But does $ZF \vdash \neg g$? Maybe...

$ZF \vdash \neg g \Rightarrow ZF \vdash \mathrm{Pr}_{ZF}(g)$

$\Rightarrow ZF \vdash (\exists x \in \omega)(x \text{ codes a proof of } g)$

But this doesn't say $ZF \vdash g$ (as PA doesn't have witnesses)

Could be: for any closed term t of PA then
$ZF \vdash \neg(t \text{ codes a proof of } g)$
Just assuming ZF consistent doesn't seem to rule this out.

Definition Let T be a theory containing an interpretation of PA. Say T is ω-consistent if there is no formula p in the language of PA with $FV(p) = \{x\}$ s.t. $T \vdash (\exists x)\, p$ but for every closed term t of PA, $T \vdash \neg p[t/x]$.

Gödel 1 (ZF): If ZF is ω-consistent then
ZF is incomplete.

Still have:
Gödel 2 (ZF): If ZF is consistent then $ZF \nvdash Con(ZF)$
Fact $ZF \nvdash AC$ and $ZF \nvdash \neg AC$ (assuming $ZF \nvdash \bot$)

As $ZF \nvdash \neg AC$ then $ZFC \nvdash \bot$.
Above all works for ZFC as well (G1, G2).
Indeed, G1 and G2 for any suitable theory T:
- T is recursively presented, (i.e. algorithm to determine axioms);
- T contains a recursive interpretation of PA (done algorithmically)
- T consistent
- (for G1 only) T ω-consistent

Are there any 'normal maths sentences' p s.t.
$ZFC \nvdash p$ and $ZFC \nvdash \neg p$.
Problem Does there exist an uncountable subset of $\mathbb{R}$ that doesn't biject with $\mathbb{R}$?
Equiv. (CH) $2^{\aleph_0} = \aleph_1$ ← is this true?
Fact $ZFC \nvdash CH$ and $ZFC \nvdash \neg CH$