Logic and Set Theory

June 2, 2024

Contents

Lectures

Start of

lecture 1

1 Propositional Logic

We build a language consisting of statements / propositions; we will assign truth values to statements; we build a deduction system so that we can prove statements that are true (and only those).

These are also the features of more complicated languages.

Definition (Language of Propositional Logic)**.** Our language consists of a set P of *primitive propositions* and a set $L = L(P)$ of *propositions* defined inductively as follows:

(i) $P \subset L$

(ii) $\perp \in L$ ($\perp$ is called 'false' or 'bottom')

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

Often $P = \{p_1, p_2, p_3, \dots\}$.

Example. $(p_1 \Rightarrow p_2)$, $((p_1 \Rightarrow \perp) \Rightarrow p_2)$, $((p_1 \Rightarrow p_2) \Rightarrow (p_1 \Rightarrow p_3))$. If $p \in L$ then we must always have $((p \Rightarrow \perp) \Rightarrow \perp) \in L$.

Remark.

(1) "Defined inductively" means that $L = \bigcup_{n \in \mathbb{N}} L_n$ where

$$\begin{aligned} L_1 &= P \cup \{\perp\} \\ L_{n+1} &= L_n \cup \{(p \Rightarrow q) \mid p, q \in L_n\} \qquad n \in \mathbb{N} \end{aligned}$$

(2) Every $p \in L$ is a finite string in $P \cup \{\perp, \Rightarrow, (,)\}$. Can prove that L is the smallest (with respect to inclusion) subset of the set Σ of all finite strings in $P \cup \{\perp, \Rightarrow, (,)\}$ such that (i) - (iii) above hold. Note $L \subsetneq \Sigma$. For example, $\Rightarrow p_1 p_3 (\in \Sigma \setminus L$.

(3) Every $p \in L$ is uniquely determined by (i) - (iii) above, i.e. either $p \in P$ or $p = \perp$ or there exists unique $q, r \in L$ such that $p = (q \Rightarrow r)$.

What about $\wedge$, $\vee$ etc? We introduct symbols $\wedge$ ('and'), $\vee$ ('or'), $\top$ ('true' or 'top') and $\neg$ ('not') as abbreviations as follows:

- $\top = (\bot \Rightarrow \bot)$
- $\neg p = (p \Rightarrow \bot)$
- $p \vee q = (\neg p \Rightarrow q)$
- $p \wedge q = \neg(p \Rightarrow \neg q)$

1.1 Semantic Entailment

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

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

(ii) if $p, q \in L$ then

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

Example. $v(p_1) = 1$, $v(p_2) = 0$. Then

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

Proposition 1.

(i) If v, v' are valuations on L and $v|_P = v'|_P$ then $v = v'$.

(ii) For any $w : P \to \{0, 1\}$, there is a valuation $v : L \to \{0, 1\}$ such that $v|_P = w$.

Proof.

(i) So $v(p) = v'(p)$ $\forall p \in P$ and $v(\bot) = v'(\bot) = 0$, so $v|_{L_1} = v'|_{L_1}$. If $v|_{L_n} = v'|_{L_n}$ then $\forall p, q \in L_n$, $v(p \Rightarrow q) = v'(p \Rightarrow q)$ and thus $v|_{L_{n+1}} = v'|_{L_{n+1}}$. So by induction, v and v' agree on $\bigcup_n L_n = L$.

(ii) We define v on L_n by induction: Let $v(p) = w(p)\ \forall p \in P$ and $v(\perp) = 0$. This defines v on L_1. Assume v is defined on L_n. Given $p \in L_{n+1} \setminus L_n$, write $p = (q \Rightarrow r)$, $q, r \in L_n$ and define

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

This defines v on L_{n+1}. Hence v is defined on $\bigcup_n L_n = L$. By construction, v is a valuation on L and $v|_P = w$. □

Definition (Tautology)**.** $t \in L$ is a *tautology* if $v(t) = 1$ for all valuations v.

Example.

(1) $(p \Rightarrow (q \Rightarrow p))$, $p, q \in L$ (a true statement is implied by any statement). We check:

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

Start of lecture 2

(2) $(\neg\neg p \Rightarrow p)$ for any $p \in L$. This can also be written as $(((p \Rightarrow \perp) \Rightarrow \perp) \Rightarrow p)$, and this can also be rewritten as $\neg p \vee p$. This is called 'law of excluded middle'.

$v(p)$	$v(p \Rightarrow \perp)$	$v((p \Rightarrow \perp) \Rightarrow \perp)$	$v(((p \Rightarrow \perp) \Rightarrow \perp) \Rightarrow p)$
0	1	0	1
1	0	1	1

(3) $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$ $(p, q, r \in L)$. If not a tautology, then there exists a valuation v such that $v(p \Rightarrow (q \Rightarrow r)) = 1$, $v((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) = 0$. So $v(p \Rightarrow q) = 1$, $v(p \Rightarrow r) = 0$. Hence $v(p) = 1$, $v(r) = 0$ and $v(q) = 1$. Then $v(p \Rightarrow (q \Rightarrow r)) = 0$ ※ .

Definition (Semantic entailment)**.** Let $S \subset L$, $t \in L$. Say S *entails* t (or S *semantically entails* t), written $S \models t$, if for every valuation v on L, $v(s) = 1\ \forall s \in S$ implies $v(t) = 1$.

Example.

(1) $\{p, p \Rightarrow q\} \models q$.

(2) $\{p \Rightarrow q, q \Rightarrow r\} \models (p \Rightarrow r)$. If $v(p \Rightarrow r) = 0$ then $v(p) = 1$, $v(r) = 0$. Then either $v(q) = 0$ and $v(p \Rightarrow q) = 0$ or $v(q) = 1$ and $v(q \Rightarrow r) = 0$.

Note. t is a tautology if and only if $\emptyset \models t$. We write this as $\models t$.

Definition (Model)**.** Given $t \in L$, say a valuation *is a model for* t (or t *is true in* v) if $v(t) = 1$. Given $S \subset L$, say a valuation v *is a model of* S if $v(s) = 1$ for all $s \in S$.

Remark. So $S \models t$ says that t is true in every model of S.

We will have one rule of deduction called *modus ponens* (MP): from p and $p \Rightarrow q$ we can deduce q.

Definition (Axiom)**.** The axioms we will use for proofs in proprositional logic are the following:

(11) $(p \Rightarrow (q \Rightarrow p))$

(22) $(\neg\neg p \Rightarrow p)$

(33) $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$

Definition (Proof)**.** Given $S \subset L$, $t \in L$, a *proof of* t *from* S is a finite sequence $t_1, t_2, \ldots, t_n$ of propositions such that $t_n = t$ and for every i either t_i is an axiom or t_i is a member of S (t_i is a premise or hypothesis) or t_i follows by MP from earlier lines: $\exists j, k < i$ such that $t_k = (t_j \Rightarrow t_i)$.

Say S *proves* t or S *syntactically entails* t if there's a proof of t from S. We denote this by $S \vdash t$. Say t is a theorem if $\emptyset \vdash t$, which we denote $\vdash t$.

Example.

(1) $\{p \Rightarrow q, q \Rightarrow r\} \vdash (p \Rightarrow r)$.

$$
\begin{array}{ll}
(p \Rightarrow (q \Rightarrow r)) \Rightarrow (p \Rightarrow q) \Rightarrow (p \Rightarrow r)) & \text{(A2)} \\
(q \Rightarrow r) \Rightarrow (p \Rightarrow (q \Rightarrow r)) & \text{(A1)} \\
(q \Rightarrow r) & \text{(premise)} \\
p \Rightarrow (q \Rightarrow r) & \text{(MP)} \\
(p \Rightarrow q) \Rightarrow (p \Rightarrow r) & \text{(MP)} \\
p \Rightarrow q & \text{(premise)} \\
p \Rightarrow r & \text{(MP)}
\end{array}
$$

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

$$
\begin{array}{ll}
p \Rightarrow ((p \Rightarrow p) \Rightarrow p) & \text{(A1)} \\
(p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)) & \text{(A2)} \\
(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p) & \text{(MP)} \\
p \Rightarrow (p \Rightarrow p) & \text{(A1)} \\
p \Rightarrow p & \text{(MP)}
\end{array}
$$

Proposition 2 (Deduction Theorem)**.** Given $S \subset L$, $p, q \in L$, we have

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

Note. This shows '$\Rightarrow$' really does behave like implication in formal proofs.

Note. To show $\{p \Rightarrow q, q \Rightarrow r\} \vdash (p \Rightarrow r)$, by Proposition 2, enough to show $\{p \Rightarrow q, q \Rightarrow r, p\} \vdash r$. This is easy: write down all premises and use (MP) twice.

Proof. If $S \vdash (p \Rightarrow q)$, then write down this proof and add two lines:

$$
\begin{array}{ll}
p & \text{(premise in } S \cup \{p\}) \\
q & \text{(MP)}
\end{array}
$$

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

Now assume $S \cup \{p\} \vdash q$. Let $t_1, t_2, \ldots, t_n = q$ be a proof of q from $S \cup \{p\}$. We show by induction that $S \vdash (p \Rightarrow t_i)$. Then done. If t_i is an axiom or $t_i \in S$, then write

$$\begin{array}{ll} t_i & \text{(axiom or premise in } S) \\ t_i \Rightarrow (p \Rightarrow t_i) & \text{(A1)} \\ p \Rightarrow t_i & \text{(MP)} \end{array}$$

to get a proof of $p \Rightarrow t_i$ from S. If $t_i = p$ then $S \vdash (p \Rightarrow p)$ since $\vdash (p \Rightarrow p)$.

Finally, assume there exists $j, k < i$ such that $t_k = (t_j \Rightarrow t_i)$. By induction we can write down proofs of $(p \Rightarrow t_j)$, $(p \Rightarrow (t_j \Rightarrow t_i))$ from S. Now just add

$$\begin{array}{ll} (p \Rightarrow (t_j \Rightarrow t_i)) \Rightarrow ((p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)) & \text{(A2)} \\ (p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i) & \text{(MP)} \\ p \Rightarrow t_i & \text{(MP)} \end{array}$$

□

Aim: $\models$ and $\vdash$ are the same.

This has two parts: *soundness* (if $S \vdash t$, then $S \models t$) and *adequacy* (if $S \models t$, then $S \vdash t$)

Start of lecture 3

Proposition 3 (Soundness theorem)**.** Given $S \subset L$, $t \in L$, if $S \vdash t$, then $S \models t$.

Proof. Let $t_1, t_2, \ldots, t_n = t$ be a proof of t from S. Let v be a model of S. We need: $v(t) = 1$. We prove by induction that $v(t_i) = 1$ for all i.

Case 1: t_i is an axiom. Then $v(t_i) = 1$ since axioms are tautologies.

Case 2: t_i is a premise. Then $v(t_i) = 1$ since v is a model of S.

Case 3: $\exists j, k < i$ such that $t_k = (t_j \Rightarrow t_i)$. Then, by the induction hypothesis, $v(t_j) = v(t_j \Rightarrow t_i) = 1$ and hence $v(t_i) = 1$.

□

Definition (Consistent)**.** Given $S \subset L$, say S is *inconsistent* if $S \vdash \bot$ and S is *consistent* if $S \nvdash \bot$.

Special case of adequacy: if $S \models \bot$ then $S \vdash \bot$, i.e. if S has no model, then S is inconsistent, or equivalently, if S is consistent, then S has a model.

Theorem 4 (Model Existence Lemma)**.** Let $S \subset L$. If S is consistent, then S has a model.

Idea: If $S \vdash t$, then $S \models t$ by Soundness theorem. So try

$$v(t) = \begin{cases} 1 & \text{if } S \vdash t \\ 0 & \text{otherwise} \end{cases}$$

This doesn't work because it's possible to have $t \in L$ such that $S \nvdash t$ and $S \nvdash \neg t$. For example, $S = \emptyset$, $t = (p_1 \Rightarrow \perp)$.

We try to enlarge S to $\overline{S}$ such that $\overline{S}$ is consistent and $\forall t \in L$, t or $\neg t$ is in $\overline{S}$.

Proof. We assume P is countable (we'll do the general case in Section 3). Then L_1 is countable and hence each L_n is countable by induction. Thus L is countable. Enumerate L: $t_1, t_2, t_3, \ldots$.

Note: if $S \subset L$ is consistent and $t \in L$, then one of $S \cup \{t\}$ or $S \cup \{\neg t\}$ is consistent. If not, then $S \cup \{t\} \vdash \perp$ and $S \cup \{\neg t\} \vdash \perp$. By the Deduction Theorem, $S \vdash \neg t$, and so $S \vdash \perp$ ※ .

So now start with a consistent $S \subset L$. Set $S_0 = S$. Using the comment above, we let S_1 be either S_1 be either $S_0 \cup \{t_1\}$ or $S_1 \cup \{\neg t_2\}$, where we pick one such that S_1 is consistent. Similarly, let S_2 be either $S_1 \cup \{t_2\}$ or $S_1 \cup \{\neg t_2\}$, where we pick one such that S_2 is consistent.

Continue inductively and set $\overline{S} = \bigcup_{n=0}^{\infty} S_n$. Then $\forall t \in L$, either $t \in \overline{S}$ or $\neg t \in \overline{S}$. Also, $\overline{S}$ is consistent since proofs are finite, so if $\overline{S} \vdash \perp$, then $\exists n$ such that $S_n \vdash \perp$ ※ .

It follows that $\overline{S}$ is *deductively closed*: if $\overline{S} \vdash t$, then $t \in \overline{S}$. If not, then $\neg t \in \overline{S}$, so $\overline{S} \vdash \neg t$ and also $\overline{S} \vdash t$ and hence $\overline{S} \vdash \perp$ (MP) ※ .

We now define $v : L \to \{0, 1\}$ by

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

Claim: v is a valuation. Then v is a model of S, and we are done.

Firstly: $v(\perp) = 0$ since $v \notin \overline{S}$ s $\overline{S}$ is consistent. Now we check $v(p \Rightarrow q)$ for $p, q \in L$.

Case 1: $v(p) = 1$, $v(q) = 0$. We need $(p \Rightarrow q) \notin \overline{S}$. By assumption, $p \in \overline{S}$, $q \notin \overline{S}$, so $\neg q \in \overline{S}$. If $(p \Rightarrow q) \in \overline{S}$, then by (MP), $\overline{S} \vdash q$ and hence $q \in \overline{S}$ ($\overline{S}$ deductively closed) ※ (as $\neg q \in \overline{S}$, so $\overline{S} \vdash \perp$).

Case 2: $v(q) = 1$. We need $(p \Rightarrow q) \in \overline{S}$. We have $q \in \overline{S}$. Write down

$$\begin{array}{ll} q & \text{(premise)} \\ q \Rightarrow (p \Rightarrow q) & \text{(A1)} \\ (p \Rightarrow q) & \text{(MP)} \end{array}$$

so $\overline{S} \vdash (p \Rightarrow q)$ and hence $(p \Rightarrow q) \in \overline{S}$.

Case 3: $v(p) = 0$. We need $(p \Rightarrow q) \in \overline{S}$, or equivalently $\overline{S} \vdash (p \Rightarrow q)$ (since $\overline{S}$ is deductively closed). Enough to show that $\overline{S} \cup \{p\} \vdash q$ (by Deduction Theorem). Since $v(p) = 0$, $p \notin \overline{S}$, and hence $\neg p \in \overline{S}$. Now obtain a proof of q from $\overline{S} \cup \{p\}$ as follows:

$$\begin{array}{ll} p & \text{(premise)} \\ \neg p & \text{(premise)} \\ \bot & \text{(MP)} \\ \bot \Rightarrow (\neg q \Rightarrow \bot) & \text{(A1)} \\ \neg\neg q & \text{(MP)} \\ \neg q \neg q \Rightarrow q & \text{(A3)} \\ q & \text{(MP)} \end{array}$$

□

Corollary 5 (Adequacy)**.** Let $S \subset L$, $t \in L$. If $S \models t$ then $S \vdash t$.

Proof. $S \cup \{\neg t\} \models \bot$, so by Theorem 4, $S \cup \{\neg t\} \vdash \bot$. Then by the Deduction Theorem, $S \vdash \neg\neg t$. Take a proof of this, and add the lines:

$$\begin{array}{ll} \neg\neg t \implies t & \text{(A3)} \\ t & \text{(MP)} \end{array}$$

So $S \vdash t$. □

Theorem 6 (Completeness Theorem)**.** Let $S \subset L$, $t \in L$. Then $S \models t$ if and only if $S \vdash t$.

Proof.

$\Rightarrow$ Soundness theorem

$\Leftarrow$ Adequacy $\square$

Corollary 7 (Compactness Theorem)**.** Let $S \subset L$, $t \in L$. If $S \models t$ then $\exists$ finite $S' \subset S$ such that $S' \models t$.

Proof. Trivial for $\vdash$ as proofs are finite. $\square$

Special case:

Corollary 8. Let $S \subset L$. If every finite subset of S has a model, then S has a model.

Proof. If not, then $S \models \perp$, so by Corollary 7 there exists finite $S' \subset S$ with $S' \models \perp$, contradiction. $\square$

Remark. Corollary 8 implies Corollary 7. If $S \models t$ then $S \cup \{\neg t\} \models t$, so by Corollary 8 there exists finite $S' \subset S$ such that $S' \cup \{\neg t\} \models \perp$. So $S' \models t$.

Note. The use of the word 'compactness' is more than a fancified analogy (see Example Sheet 1).

Corollary 9 (Decidability Theorem)**.** Let $S \subset L$, S finite and $t \in L$. Then there's an algorithm that can decide in finite time whether $S \vdash t$ or not.

Proof. Easy to decide if $S \models t$. Just write out a truth table. $\square$

Start of

lecture 4

2 Well-ordering and ordinals

Definition (Linear order)**.** A *linear order* of *total order* on a set X is a relation $<$ on X that is:

(i) *irreflexive*: $\forall x \in X$, $\neg(x < x)$.

(ii) *transitive*: $\forall x, y, z \in X$, $(x < y \wedge y < z) \implies (x < z)$.

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

Remark. In (iii) exactly one holds: for example, if $x < y$ and $y < x$, then $x <$ by (ii) which contradicts (i).

Notation. We say X is linearly ordered by $<$, or simply say X is a linearly ordered set.

Example. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ with their usual order ($\mathbb{N} = \{1, 2, 3, \ldots\}$).

Note. If X is a set of size ≥ 2, then on $\mathbb{P}X = \{Y \mid Y \subset X\}$ (power set of X), defining $a < b$ to mean $a \subset b$, $a \neq b$ is not trichotomous.

Notation. If X is linearly ordered by $<$, then we write $x > y$ for $y < k$, $x \leq y$ for $x < y$ or $x = y$, and $x \geq y$ for $x > y$ or $x = y$.

Note. Note that $\leq$ is:

1. *reflexive*: $\forall x \in X$, $x \leq x$.
2. *antisymmetric*: $\forall x, y \in X$, $(x \leq y \wedge y \leq x) \implies (x = y)$.
3. *transitive*: $\forall x, y, z \in X$, $(x \leq y \wedge y \leq z) \implies (x \leq z)$.
4. *trichotomous*: $\forall x, y \in X$, $x \leq y$ or $y \leq x$.

Note. If X is linearly ordered by $<$, then any $Y \subset X$ is linearly ordered by $<$ (more precisely, by the restriction of $<$ to Y).

Definition (Well-ordering)**.** A *well-ordering* on a set X is a linear order $<$ on X such that every non-empty subset X has a least element: $\forall S \subset X$, $S \neq \emptyset$ implies $\exists x \in S$ such that $\forall y \in S$, $x \leq y$.

Note. This least element is always unique by antisymmetric.

Notation. Say X is *well-ordered* by $<$, or simply say X is a well-ordered set.

Example. $\mathbb{N}$ with the usual linear order is a well-ordering.

$\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are not (they have no least element). $\{x \in \mathbb{R} \mid x \geq 0\}$ is not well-ordered, because for example, $\{x \in \mathbb{R} \mid x > 0\}$ has no least element.

Note. Every subset of a well-ordered set is well-ordered. We'll see that $\mathbb{Q}$ has a rich collection of well-ordered subsets.

Definition (Order isomorphic)**.** Say linearly ordered sets X, Y are *order-isomorphic* if there exists a bijection $f : X \to Y$ which is *order-preserving*: $\forall x < y$ in X, $f(x) < f(y)$. Such an f is called an *order-isomorphism.* Then f^{-1} is also an order-isomorphism.

Note. If linearly ordered sets X, Y are order-isomorphic and X is well-ordered, then so is Y.

Example. $\mathbb{N}$ and $\mathbb{Q}$ are not order-isomorphic.

$\mathbb{Q}$ and $\mathbb{Q} \setminus \{0\}$ are order-isomorphic (see Numbers & Sets Example Sheet).

$A = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\} = \left\{\frac{n}{n+1} \mid n \in \mathbb{N}\right\}$ is order-isomorphic to $\mathbb{N}$ ($n \mapsto \frac{n}{n+1}$).

$B = A \cup \{1\}$ is well-ordered, but not order-isomorphic to $\mathbb{N}$ (it has a greatest element).

$C = A \cup \{2\}$ is order-isomorphic to B.

$D = A \cup (A+1) = A \cup \{\frac{3}{2}, \frac{5}{3}, \frac{7}{4}, \dots\}$ is well-ordered, but not order-isomorphic to A or B.

Definition (Initial segment)**.** A subset I of a linearly ordered set X is an *initial segment* (i.s.) of X is $x \in I, y < x \implies y \in I$ for any $x, y \in X$.

Example. $\{1, 2, 3, 4\}$ is an initial segment of $\mathbb{N}$. $\{1, 2, 3, 5\}$ is not.

$[0, 1]$ is an initial segment of $\{x \in \mathbb{R} \mid x \geq 0\}$.

Notation. In general, for $x \in X$, $I_x = \{y \in X \mid y < x\}$ is an *is* of X by transitive. I_x is a proper initial segment of X (meaning $I_x \neq X$), because it does not contain x.

Note. In general, not every proper initial segment is of this form. For example, $(-\infty, 1]$ is a proper initial segment of $\mathbb{R}$, but $(-\infty, 1] \neq I_x$ for any $x \in \mathbb{R}$.

Remark. If X is well-ordered and I is a proper initial segment of X, then $I = I_x$ where x is the least element of $X \setminus I$.

Indeed, if $y \in I_x$ then $y < x$, so $y \in I$ by choice of x. If $y \in I$ and $y \geq x$, then $x \in I$ as I is an initial segment, contradiction. So $y < x$, i.e. $y \in I_x$.

Lemma 1. Let X, Y be a well-ordered set, I an initial segment of Y and $f : X \to Y$ an order-isomorphism between X and I. Then for each $x \in X$, $f(x)$ is the least element of $Y \setminus \{f(y) \mid y < x\}$.

Proof. The set $A = Y \setminus \{f(y) \mid y < x\}$ is $\neq \emptyset$ since $f(x) \in A$. Let a be the least element of A. Then $a \leq f(x)$ and $f(x) \in I$, and so $a \in I$. Thus $a = f(x)$ for some $z \in X$. Note that $z > x$ implies $a = f(z) > f(x)$, contradiction. So $z \leq x$.

If $z < x$, then $a = f(z) \in \{f(y) \mid y < x\}$, ※ as $a \in A$. So $z = x$ and $a = f(z) = f(x)$. □

Proposition 2 (Proof by induction)**.** Let X be a well-ordered set and $S \subset X$ satisfying the following for every $x \in X$: $\forall y < x, y \in S$ implies $x \in S$. Then $S = X$.

Note. Assume S is given by a property p: $S = \{x \in X \mid p(x)\}$. The above can be written as

$$(\forall x \in X)((\forall y < x, p(y)) \implies p(x)) \implies (\forall x \in X, p(x))$$

(base case is included since the left hand side will be vacuously true for the least element).

Proof. If $S \neq X$, then $X \setminus S$ has a least element x, say. If $y < x$, then $y \in S$ by choice of x. By the assumption on S, $x \in S$, contradiction. □

Start of lecture 5

Proposition 3. Let X, Y be well-ordered sets that are order-isomorphic. Then there exists unique order-isomorphism $X \to Y$.

Remark. Not true in general for linearly ordered sets. For example for $\mathbb{Z} \to \mathbb{Z}$ we can take $n \mapsto n$ or $n \mapsto n + 17$, and for $[0, \infty) \to [0, \infty)$ can take $x \mapsto x$ or $x \mapsto x^2$.

Proof. Let $f, g : X \to Y$ be order-isomorphisms. We prove that $\forall x \in X$, $f(x) = g(x)$ by induction. Let $x \in X$. Assume $f(y) = g(y)$ for all $y < x$ (induction hypothesis). By

Lemma 1,

$$f(x) = \min(Y \setminus \{f(y) \mid y < x\})$$
$$g(x) = \min(Y \setminus \{g(y) \mid y < x\})$$

By induction hypothesis,

$$\{f(y) \mid y < x\} = \{g(y) \mid y < x\}.$$

So $f(x) = g(x)$. □

Remark. Induction proves things. We need a tool to construct things. This will be *recursion*.

Note. A function from a set X to a set Y is a subset f of $X \times Y$ such that:

(i) $\forall x \in X$, $\exists y \in Y$ such that $(x, y) \in f$.

(ii) $\forall x \in X$, $\forall y, z \in Y$ $((x, y) \in f \wedge (x, z) \in f) \implies (y = z)$.

Of course we write '$y = f(x)$' instead of '$(x, y) \in f$'. Note that $f \in \mathbb{P}(X \times Y)$. For $Z \subset X$, the restriction of f to Z is $f|_Z = \{(x, y) \in f \mid x \in Z\}$. $f|_Z$ is a function $Z \to Y$, so $f|_Z \subset Z \times Y \subset X \times Y$, so $f|_Z \in \mathbb{P}(X \times Y)$.

Theorem 4 (Definition by recursion)**.** Let X be a well-ordered set and Y be an arbitrary set. Then for any function $G : \mathbb{P}(X \times Y) \to Y$ there is a unique function $f : X \to Y$ such that $f(x) = G(f|_{I_x})$ for every $x \in X$.

Proof. **Uniqueness:** Assume f, g both satisfy the conclusion. Given $x \in X$, if $f(y) = g(y)$ for all $y < x$, then $f(x) = G(f|_{I_x}) = G(g|_{I_x}) = g(x)$. So by induction, $f = g$.

Existence: Say h is an *attempt* if h is a function $I \to Y$ for some initial segment I of X such that $\forall x \in I$, $h(x) = G(h|_{I_x})$ (note $I_x \subset I$). Let h, h' be attempts. We show that $\forall x \in X$, if $x \in \operatorname{dom}(h) \cap \operatorname{dom}(h')$, then $h(x) = h'(x)$. Here, $\operatorname{dom}(h)$ is the domain of h, i.e. I as above. Fix $x \in \operatorname{dom}(h) \cap \operatorname{dom}(h')$ and assume $h(y) = h'(y)$ for every $y < x$ (note $y < x$ implies $y \in \operatorname{dom}(h) \cap \operatorname{dom}(h')$). Then $h|_{I_x} = h'|_{I_x}$, so $h(x) = G(h|_{I_x}) = G(h'|_{I_x}) = h'(x)$. Then done by induction.

What we have left to show for existence is that $\forall x \in X$ there exists an attempt h such that $x \in \operatorname{dom} h$. We prove this by induction. Fix $x \in X$ and assume that for $y < x$ there is an attempt defined at y, and let h_y be the unique attempt with domain $\{z \in X \mid z \leq$

$y\} = I_y \cup \{y\}$. Then $h = \bigcup_{y<x} h_y$ is a well-defined function on I_x and it is an attempt since foy $< x$, $h(y) = h_y(y) = G(h_y|_{I_x}) = G(h|_{I_y})$. Then $h \cup \{(x, G(h))\}$ is an attempt with domain $I_x \cup \{x\}$. Finally, define $f : X \to Y$, $f(x) = h(x)$ where h is any attempt defined at x. This is well-defined by above and $f(x) = h(x) = G(h|_{I_x}) = G(f|_{I_x})$. □

Proposition 5 (Subset collapse)**.** Let Y be a well-ordered set and $X \subset Y$. Then X is order-isomorphic to a unique initial segment of Y.

Proof. Without loss of generality, $X \neq \emptyset$.

Uniqueness: Assume $f : X \to I$ is an order-isomorphism where I is an initial segment of Y. By Lemma 1, $f(x) = \min(Y \setminus \{f(y) \mid y < x, y \in X\})$. So by induction, f and hence I are uniquely determined.

Existence: Fix $y_0 \in Y$. By Theorem 4, there's a function $f : X \to Y$ such that

$$f(x) = \begin{cases} \min(Y \setminus \{f(y) \mid y \in X, y < x\}) & \text{if it exists} \\ y_0 & \text{otherwise} \end{cases}$$

We first prove that the 'otherwise' clause never occurs. We prove that $\forall x \in X$, $f(x) \leq x$. If $\forall y \in X$, $y < x$ implies $f(y) \leq y$, then $x \in Y \setminus \{f(y) \mid y \in X, y < x\}$, so $f(x) \leq x$. Done by induction. This also shows that f is injective.

f order preserving: Given $y < x$ in X, $f(x) \in Y \setminus \{f(z) \mid z \in X, z < x\} \subset Y \setminus \{f(z) \mid z \in X, z < y\}$. So $f(y) \leq f(x)$, and hence $f(y) < f(x)$ by injectivity.

$\operatorname{Im} f$ is an initial segment of Y: Assume $a \in Y \setminus \operatorname{Im} f$. We show $f(x) < a$ for all $x \in X$. If $f(y) < a$ for all $y \in X$, $y < x$, then $a \in Y \setminus \{f(y) \mid y \in X, y < x\}$, so $f(x) \leq a$ and hence $f(x) < a$. Done by induction. □

Remark. A well-ordered set X is not order-isomorphic to a proper initial segment of X (by uniqueness). But X is of course order-isomorphic to X.

Notation. Let X, Y be well-ordered sets. Write $X \leq Y$ if X is order-isomorphic to an initial segment of Y.

Example. If $A = \left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\} \cup \{1\}$. Then $\mathbb{N} \leq A$.

Theorem 6. Let X, Y be well-ordered sets. Then $X \leq Y$ or $Y \leq X$.

Proof. Assume $Y \not\leq X$. Then $Y \neq \emptyset$ and we can fix $y_0 \in Y$. We recursively define $f : X \to Y$ by

$$f(x) = \begin{cases} \min(Y \setminus \{f(y) \mid y < x\}) & \text{if it exists} \\ y_0 & \text{otherwise} \end{cases}$$

If the 'otherwise' clause occurs, let x be the least element of X when this happens. Then $f(I_x) = Y$ and as in Proposition 5, f is an order-isomorphism $I_x \to Y$, which contradicts $Y \not\leq X$. So the 'otherwise' clause never occurs. So as in proof of Proposition 5, f is an order-isomorphism to an initial segment of Y, i.e. $X \leq Y$. □

Proposition 7. Let X, Y be well-ordered sets. If $X \leq Y$ and $Y \leq X$ then X and Y are order-isomorphic.

Proof. Let $f : X \to Y$, $g : Y \to X$ be order-isomorphisms onto initial segment of Y, X respectively. Then $g \circ f$ is an order-isomorphism between X and an order-isomorphism of X, so $g \circ f = \mathrm{id}_X$ by uniqueness in Proposition 5. Similarly $f \circ g = \mathrm{id}_Y$. □

Start of

lecture 6

Remark. Theorem 6 and Proposition 7 together show that $\leq$ is a linear order (reflexive, antisymmetric, transitive and trichotomous), provided we identify well-ordered sets that are order-isomorphic to each other.

Notation. We introduce '$X < Y$' to mean $X \leq Y$ and X is not order-isomorphic to Y. So $X < Y$ if and only if X order-isomorphic to a proper initial segment of Y.

Question: Do the well-ordered sets form a set? If so, is it a well-ordered set?

First we construct new well-ordered sets from old ones.

'there's always another one':

Definition (Successor ordinal)**.** Let X be a well-ordered set, fix $x_0 \notin X$, and set $X^+ = X \cup \{x_0\}$, which we well-order by extending $<$ on X to X^+ by letting $x < x_0$ for all $x \in X$. This is unique up to order-isomorphism and $X < X^+$.

Upper bounds: Given a set $\{X_i \mid i \in I\}$ of well-ordered sets, we seek a well-ordered set X such that $X_i \leq X$ for all $i \in I$.

Definition (Extends)**.** Given well-ordered sets $(X, <_X)$ and $(Y, <_Y)$, say Y *extends* X if $X \subset Y$, $<_X$ is the restriction to X of $<_Y$ and X is an initial segment of Y.

Definition (Nested)**.** We say $\{X_i \mid i \in I\}$ is *nested* if $\forall i, j \in I$ either X_j extends X_i or X_i extends X_j.

Proposition 8. Let $\{X_i \mid i \in I\}$ be a nested set of well-ordered sets. Then there exists a well-ordered set X such that $X_i \leq X$ for all $i \in I$.

Proof. Let $X = \bigcup_{i \in I} X_i$ and define $<$ on X as follows: $x < y$ if and only if $\exists i \in I$ such that $x, y \in X_i$ and $x <_i y$ where $<_i$ is the well-ordering of X_i. Since the X_i are nested, this is well-defined, is a linear order and each X_i is an initial segment of X.

Given $S \subset X$, $S \neq \emptyset$, since $S = \bigcup_{i \in I}(S \cap X_i)$, there exists $i \in I$ such that $S \cap X_i \neq \emptyset$. Let x be a least element of $S \cap X_i$ (since X_i is well-ordered). Then x is a least element of S since X_i is an initial segment of X. □

Remark. Proposition 8 holds even if the X_i are not nested (see Section 5).

Ordinals

Definition (Ordinal)**.** An *ordinal* is a well-ordered set but we consider two ordinals the same if they're order-isomorphic.

Remark. A formal definition will be given in Section 5. You could think of the term 'ordinal' as a shorthand (for now).

Definition (Order type)**.** The *order type* of a well-ordered set X is the unique ordinal α order-isomorphic to X. Write 'α is the order type (O.T.) of X'.

Example. For $k \in \mathbb{N} \cup \{0\}$, we let k be the order type of a well-ordered set of size k (this is unique). Let ω be the order type of $\mathbb{N}$ (also the order type of $\mathbb{N} \cup \{0\}$). The set $A = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\}$ in $\mathbb{Q}$ also has order type ω.

Notation. We write ω for the order type of any set which is order-isomorphic to $\mathbb{N}$.

Notation. For ordinals α, β we write $\alpha \leq \beta$ is $X \leq Y$ where X is a well-ordered set of order type α, Y is a well-ordered set of order type β. This is well-defined. We also write $\alpha < \beta$ is $X < Y$. We let α^+ be the order type of X^+.

Remark. $\leq$ is a linear order; if $\alpha \leq \beta$ and $\beta \leq \alpha$ then $\alpha = \beta$.

Theorem 9. Let α be an ordinal. The ordinals $< \alpha$ form a well-ordered set of order type α.

Proof. Fix a well-ordered set X with order type α. Let

$$\tilde{X} = \{Y \subset X \mid Y \text{ is a proper initial segment of } X\}.$$

Then $<$ (defined for well-ordered sets) is a linear order on $\tilde{X}$. Note that $x \mapsto I_x : X \to \tilde{X}$ is an order-isomorphism. So $\tilde{X}$ is a well-ordered set of order type α. So

$$\{\mathrm{OT}(Y) \mid Y \in \tilde{X}\}$$

is the set of ordinals $< \alpha$ and $Y \mapsto \mathrm{OT}(Y)$ is an order-isomorphism from $\tilde{X}$ to this set. □

Notation. $I_\alpha = \{\beta \mid \beta < \alpha\}$ 'A nice example of a well-ordered set of order type α'.

Proposition 10. A non empty set S of ordinals has a least element.

Proof. Pick $\alpha \in S$. If α is not a least element of S, then $S \cap I_\alpha \neq \emptyset$, and hence (by Theorem 9) it has a least element β. Then β is a least element of S: if $\gamma \in S$, $\gamma < \alpha$, then $\gamma \in I_\alpha \cap S$, and so $\beta \leq \gamma$. □

Theorem 11 (Burati-Forti paradox)**.** The ordinals do not form a set.

Proof. Assume otherwise and let X be the set of ordinals. Then X is a well-ordered set by Proposition 10 (and earlier results). Let α be the order type of X. Then X is order-isomorphic to I_α, which is a proper initial segment of X, ※ . □

Remark. Let $S = \{\alpha_i \mid i \in I\}$ be a set of ordinals. Then by Proposition 8 the nested set $\{I_{\alpha_i} \mid i \in I\}$ has an upper bound. So there exists an ordinal α such that $\alpha_i \le \alpha$ for all $i \in I$. By Theorem 9 we can take the least such α. We take the least element of

$$\{\beta \in I_\alpha \cup \{\alpha\} \mid \forall i \in I, \alpha \le \beta\}.$$

We denote by $\sup S$ the least upper bound on S. Note if $\alpha = \sup S$ then $I_\alpha = \bigcup_{i \in I} I_\alpha$.

A list of some ordinals

$$0, 1, 2, 3, \dots, \omega, \omega, \omega^+ = \omega + 1, \omega + 2, \omega + 3, \dots,$$
$$\omega + \omega = \omega \cdot 2 = \sup\{\omega + n \mid n < \omega\}, \omega \cdot 2 + 1, \omega \cdot 2 + 2, \dots, \omega \cdot 3, \dots, \omega \cdot 4, \dots$$
$$\omega \cdot \omega = \omega^2 = \{\omega \cdot n \mid n < \omega\}, \omega^2 + 1, \omega^2 + 2, \dots \omega^2 + \omega, \dots, \omega^2 + \omega \cdot 2, \dots, \omega^2 + \omega \cdot 3, \dots$$
$$\omega^2 \cdot 2, \dots, \omega^2 \cdot 3, \dots, \omega^3, \dots, \omega^4, \dots, \omega^\omega = \sup\{\omega^n \mid n < \omega\}, \omega^\omega + 1, \dots$$
$$\omega^\omega + \omega, \dots, \omega^\omega + \omega^2, \dots, \omega^\omega \cdot 2, \dots, \omega^\omega \cdot \omega = \omega^{\omega+1}, \dots, \omega^{\omega+2}, \dots, \omega^{\omega \cdot 2}, \dots, \omega^{\omega \cdot 3}, \dots$$
$$\omega^{\omega^2}, \dots, \omega^{\omega^3}, \dots, \omega^{\omega^\omega}, \dots, \omega^{\omega^{\omega^\omega}}, \dots, \varepsilon_0 = \sup\{\underbrace{\omega^{\omega^{\cdot^{\cdot^{\cdot^\omega}}}}}_{n} \mid n < \omega\}, \dots$$
$$\varepsilon_1, \dots, \varepsilon_2, \dots, \varepsilon_\omega, \dots, \varepsilon_{\varepsilon_0}, \dots, \varepsilon_{\varepsilon_{\varepsilon_0}}, \dots$$

Remarkably, all of these are countable! This can be seen by checking that each of them is a countable supremum of countable ordinals, hence must be countable.

Start of

lecture 7

Question: Does there exist an uncountable ordinal, i.e. does there exist an uncountable well-ordered set? Can we well order $\mathbb{R}$?

Theorem 12. There exists an uncountable ordinal.

Idea: Assume α is an uncountable ordinal. Then there is a least such α:

$$\{\beta \in I_\alpha \cup \{\alpha\} \mid \beta \text{ uncountable}\} \neq \emptyset,$$

so has a least element γ, say. So I_γ is exactly the set of all countable ordinals. If X is a countable well-ordered set, then there exists an injection $f : X \to \mathbb{N}$. Then $Y = f(X)$ is well-ordered by $f(x) < f(y) \iff x < y$ in X. Then Y is order-isomorphic to X.

Proof. Let

$$A = \{(Y, <) \in \mathbb{P}\mathbb{N} \times \mathbb{P}(\mathbb{N} \times \mathbb{N}) \mid Y \text{ is well-ordered by } <\}.$$

Let $B = \{\mathrm{OT}(Y, <) \mid (Y, <) \in A\}$. By above, B is exactly the set of all countable ordinals. Let $\omega_1 = \sup B$. If $\omega_1 \in B$ then ${\omega_1}^+ \notin B$, so ${\omega_1}^+$ is an uncountable ordinal. In fact, ω_1 is uncountable, since if ω_1 is countable, then ${\omega_1}^+$ must be countable as well (countable set union with a single element is still countable). □

Notation. ω_1 in the proof is the least uncountable ordinal. In general, when we write ω_1, we mean the least uncountable ordinal (which may be constructed as in the previous proof).

Remark. Every proper initial segment of ω_1 is countable. If $\alpha_1, \alpha_2, \alpha_3, \ldots \in \omega_1$, then

$$\sup\{\alpha_1, \alpha_2, \ldots\} = \mathrm{OT}\left(\bigcup_{i \in \mathbb{N}} I_{\alpha_i}\right)$$

is countable, hence not equal to ω_1.

Theorem 13 (Hartog's Lemma)**.** For any set X, there exists an ordinal α such that α does not inject into X.

Proof. Repeat the proof of Theorem 12 replacing $\mathbb{N}$ with X. □

Notation. The least such α in Hartog's Lemma is denoted by $\gamma(X)$. For example $\gamma(\omega) = \omega_1$.

$$0, 1, 2, \ldots, \omega, \ldots, \varepsilon_0 = \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}}, \ldots, \varepsilon_1, \ldots, \varepsilon_{\varepsilon_{\varepsilon_{\ddots}}}, \ldots, \omega_1, \ldots, \omega_1 \cdot 2, \ldots, \omega_2 = \gamma(\omega_1), \ldots$$

Types of ordinals

Definition (Successor / limit ordinal)**.** Let α be an ordinal, and consider whether α has a greatest element (i.e. if X has order type α, does X have a greatest element).

If yes: Let β be the greatest element of I_α. Then $I_\alpha = I_\beta \cup \{\beta\}$. So $\alpha = \beta^+$, and $\alpha = (\sup I_\alpha)^+$. We call such an α a *successor ordinal*.

If no: Then $I_\alpha = \sup I_\alpha$, i.e. $\alpha = \sup\{\beta \mid \beta < \alpha\}$. We say α is a *limit ordinal*.

Example. $1 = 0^+$ is a successor ordinal, $\omega = \sup\{n < \omega\}$ is a limit ordinal, ω^+ is a successor ordinal, ω_1 is a limit ordinal.

Weirdly, 0 is a limit ordinal. Some people prefer to add a special category for 0, defining it as neither a successor ordinal nor a limit ordinal.

Ordinal Arithmetic

Definition (Ordinal addition)**.** We define $\alpha + \beta$ for α, β ordinals by recursion on β with α fixed. We define:

$\beta = 0$: $\alpha + 0 = \alpha$,

$\beta = \gamma^+$: $\alpha + \gamma^+ = (\alpha + \gamma)^+$,

$\beta \neq 0$ limit: $\alpha + \beta = \sup\{\alpha + \gamma \mid \gamma < \beta\}$.

Remark. Technically, we fix α, β and define $\alpha + \gamma$ for all $\gamma \leq \beta$ by Definition by recursion as above. We do this for all β. This gives a well-defined '+' by uniqueness in the Definition by recursion.

Similarly, we can prove things by induction: Let $p(\alpha)$ be a statement for each ordinal α. Then

$$(\forall \alpha)((\forall \beta)((\beta < \alpha) \implies p(\beta)) \implies p(\alpha)) \implies (\forall \alpha) p(\alpha)$$

If not, then there exists α with $p(\alpha)$ false. Then there exists least such α ($\{\beta \leq \alpha \mid p(\beta) \text{ false}\} \neq \emptyset$). Then $p(\beta)$ is true for all $\beta < \alpha$. By assumption, $p(\alpha)$ is true, ※ .

Example. For any α, $\alpha + 1 = \alpha + 0^+ = (\alpha + 0)^+ = \alpha^+$.

If $m < \omega$, then we have $m + 0 = m$ and for $n < \omega$,

$$m + (n + 1) = m + +n^+ = (m + n)^+ = (m + n) + 1 = m + n + 1$$

So on ω, ordinal addition is the usual addition.

More examples:

$$\begin{aligned}\omega + 2 &= \omega + 1^+ = (\omega + 1)^+ = \omega^{++} \\ \omega + \omega &= \sup\{\omega + n \mid n < \omega\} = \sup\{\omega, \omega + 1, \omega + 2, \ldots\} \\ 1 + \omega &= \sup\{1 + n \mid n < \omega\} = \sup\{1, 2, 3, \ldots\} = \omega \neq \omega + 1\end{aligned}$$

So '+' is not commutative.

Proposition 14. $\forall \alpha, \beta, \gamma$ ordinals, $\beta \leq \gamma \implies \alpha + \beta \leq \alpha + \gamma$.

Proof. We prove this by induction on γ (with α, β fixed).

- $\gamma = 0$: If $\beta \leq \gamma$, then $\beta = 0$, so result is true.
- $\gamma = \delta^+$ If $\beta \leq \gamma$, then either $\beta = \gamma$ and we're done or $\beta \leq \delta$ and so $\alpha + \beta \leq \alpha + \delta < (\alpha + \delta)^+ = \alpha + \delta^+ = \alpha + \gamma$.
- $\gamma \neq 0$ limit If $\beta \leq \gamma$, then without loss of generality $\beta < \gamma$, so $\alpha + \beta \leq \sup\{\alpha + \delta \mid \delta < \gamma\} = \alpha + \gamma$. □

Remark. From Proposition 14, we get $\beta < \gamma \implies \alpha + \beta < \alpha + \gamma$. Indeed,

$$\alpha + \beta < (\alpha + \beta)^+ = \alpha + \beta^+ \leq \alpha + \gamma.$$

Note that $1 < 2$ but $1 + \omega = 2 + \omega = \omega$, the proposition is not true when the order is swapped.

Lemma 15. Let α be an ordinal and S a nonempty set of ordinals. Then

$$\alpha + \sup S = \sup\{\alpha + \beta \mid \beta \in S\}.$$

Proof. If $\beta \in S$, then $\alpha + \beta \leq \alpha + \sup S$ (by Proposition 14). Hence

$$\sup\{\alpha + \beta \mid \beta \in S\} \leq \alpha + \sup S.$$

For the reverse inequality, consider two cases. If S has a greatest element, β say, then

$$\alpha + \sup S = \alpha + \beta.$$

For all $\gamma \in S$, $\gamma \leq \beta$, so by Proposition 14, $\alpha + \gamma \leq \alpha + \beta$. It follows that

$$\sup\{\alpha + \gamma \mid \gamma \in S\} = \alpha + \beta.$$

If S has no greatest element, then $\lambda = \sup S$ is a $\neq 0$ limit ordinal (if $\lambda = \gamma^+$ then $\gamma < \lambda$, so there exists $\delta \in S$ with $\gamma < \delta$, then $\lambda = \gamma^+ \leq \delta$, so $\lambda \in S$, contradiction). So

$$\alpha + \sup S = \sup\{\alpha + \beta \mid \beta < \lambda\}$$

by definition. If $\beta < \gamma$, then there exists $\delta \in S$, $\beta < \delta$. By Proposition 14, $\alpha+\beta \leq \alpha+\delta$. It follows that

$$\sup\{\alpha + \beta \mid \beta < \lambda\} \, wle \sup\{\alpha + \delta \mid \delta \in S\}$$ □

Proposition 16. $\forall \alpha, \beta, \gamma$, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

Proof. By induction on γ.

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

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

$\gamma \neq 0$ limit:

$$\begin{aligned} (\alpha + \beta) + \gamma &= \sup\{(\alpha + \beta) + \delta \mid \delta < \gamma\} \\ &= \sup\{\alpha + (\beta + \delta) \mid \delta < \gamma\} \\ &= \alpha + \sup\{\beta + \delta \mid \delta < \gamma\} \\ &= \alpha + (\beta + \delta) \end{aligned}$$ □

Start of lecture 8

Remark. The definition of $\alpha + \beta$ we gave last time is called the "induction definition".

Definition (Synthetic ordinal addition)**.** Given well-ordered sets X, Y, the disjoint union $X \sqcup Y$ is the well-ordered set $\overset{X}{\leftrightarrow}\overset{Y}{\leftrightarrow}$. Formally, it is the set $X \times \{0\} \cup Y \times \{1\}$ with ordering:

$$(x, i) < (y, j) \iff \begin{cases} \text{either } i = j = 0 \text{ and } x < y \text{ in } X \\ \text{or } i = j = 1 \text{ and } x < y \text{ in } Y \\ \text{or } i = 0, j = 1 \text{ and } x \in X,\, y \in Y \end{cases}$$

So this is a well-ordered set Z which has an initial segment X' to X and $Z \setminus X'$ is order-isomorphic to Y. This is unique up to order-isomorphism.

For ordinals $\alpha, \beta + \beta = \alpha \sqcup \beta$ (more precisely, $\alpha + \beta$ is the order type of $X \sqcup Y$ where $\alpha = \mathrm{OT}(X)$, $\beta = \mathrm{OT}(Y)$).

Note. $\alpha^+ = \alpha \sqcup 1$. With this definition, it's easy to see that $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ since $(\alpha \sqcup \beta) \sqcup \gamma$ is order-isomorphic to $\alpha \sqcup (\beta \sqcup \gamma)$.

Also, we can easily prove $\beta \le \gamma \implies \alpha + \beta \le \alpha + \gamma$ as $\alpha \sqcup \beta$ is an initial segment of $\alpha \sqcup \gamma$.

Proposition 17. The inductive and synthetic definitions of ordinal addition coincide.

Proof. Temporarily, let $\alpha \dot{+} \beta$ denote the synthetic addition, and $\alpha + \beta$ denote the inductive addition. We prove $\forall \alpha, \beta\ \alpha + \beta \dot{+} \beta$ by induction on β (with α fixed).

$\beta = 0$: $\alpha + 0 = \alpha = \alpha \sqcup 0$.

$\beta = \delta^+$: $\alpha + \beta = (\alpha + \delta)^+ = (\alpha \dot{+} \delta)^+ = (\alpha \sqcup \delta) \sqcup 1 = \alpha \sqcup (\delta \sqcup 1) = \alpha \dot{+} \delta^+ = \alpha \dot{+} \beta$.

$\beta \neq 0$ limit:

$$\begin{aligned} \alpha + \beta &= \sup\{\alpha + \gamma \mid \gamma < \beta\} \\ &= \sup\{\alpha \dot{+} \gamma \mid \gamma < \beta\} \\ \bigcup_{\gamma < \beta} \alpha \sqcup \gamma & \\ &= \alpha \sqcup \bigcup_{\gamma < \beta} \gamma \\ &= \alpha \sqcup \beta \\ &= \alpha \dot{+} \beta \end{aligned}$$

(as $\alpha \sqcup \gamma$, $\gamma < \beta$ are nested). □

Ordinal Multiplication

We give two definitions: inductive and syntetic.

Definition (Inductive multiplication)**.** Define $\alpha \cdot \beta$ by recursion on β (α fixed):

- $\alpha \cdot 0 = 0$
- $\alpha \cdot \beta^+ = \alpha \cdot \beta + \alpha$
- $\alpha \cdot \beta = \sup\{\alpha \cdot \gamma \mid \alpha < \beta\}$ (for $\beta \neq 0$ limit ordinal)

Example. For $m, n < \omega$, we have $m \cdot 0 = 0$, $m \cdot (n+1) = m \cdot n^+ = m \cdot n + m$. This gives the usual multiplication.

$$\omega \cdot 2 = \omega \cdot 1^+ = \omega \cdot 1 + \omega = \omega \cdot 0^+ + \omega = (\omega \cdot 0 + \omega) + \omega = \omega + \omega$$

$$2 \cdot \omega = \sup\{2 \cdot n \mid n < \omega\} = \omega \neq \omega \cdot 2$$

So multiplication is not commutative.

Definition (Synthetic multiplication)**.** Given well-ordered sets X, Y, we well-order $X \times Y$ by

$$(x, y) < (w, z) \iff \begin{cases} \text{either } y = z \text{ and } x < w \text{ in } X \\ \text{or } y < z \text{ in } Y \end{cases}$$

For ordinals α, β define $\alpha \cdot \beta = \alpha \times \beta$ (the order type of $X \times Y$ where X has order type α, Y has order type β).

Note. As before, the two definitions coincide (proof by inductionon β).

Properties:

$$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$$

$$\beta \leq \gamma \implies \alpha \cdot \beta \leq \alpha \cdot \gamma$$

On Example Sheet 2, you will check whether the following are true:

$$(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$$

$$\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$$

Ordinal Exponentiation

Define α^β by recursion on β (α fixed):

- $\alpha^0 = 1$
- $\alpha^{\beta^+} = \alpha^\beta \cdot \alpha$
- $\alpha^\beta = \sup\{\alpha^\gamma \mid \gamma < \beta\}$ (for $\beta \neq 0$ limit ordinal)

Example. For $m, n < \omega$, m^n has usual meaning.

$$\omega^2 = \omega^{1^+} = \omega^1 \cdot \omega = \omega^{0^+} \cdot \omega = (\omega^0 \cdot \omega) \cdot \omega = \omega \cdot \omega$$

$$2^\omega = \sup\{2^n \mid n < \omega\} = \omega$$

which is countable!

** Non-examinable **

Let X be a separable Banach space, then $X \hookrightarrow C[0,1]$ (universal property for separable Banach spaces).

Question: Does there exist a universal space for separable reflexive spaces?

Answer: No (Szlenk).

To each Banach space X you associate an ordinal $\mathrm{Sz}(X)$ (Szlenk index of X). For all separable X, $\mathrm{Sz}(X) \le \omega_1$.

$$\mathrm{Sz}(X) < \omega_1 \iff X^* \text{ separable}$$

$$X \hookrightarrow Y \implies \mathrm{Sz}(X) \le \mathrm{Sz}(Y)$$

$\forall \alpha < \omega_1$, there exists separable reflexive X_α such that $\mathrm{Sz}(X_\alpha) > \alpha$. If Z is separable reflexive and for all separable reflexive X, $X \hookrightarrow Z$ then $X_\alpha \hookrightarrow Z$ for all $\alpha < \omega_1$, so $\mathrm{Sz}(Z) \ge \mathrm{Sz}(X_\alpha) > \alpha$. So $\mathrm{Sz}(Z) = \omega_1$, contradiction.

This is the end of the non-examinable part.

3 Posets and Zorn's Lemma

Definition (Partial order)**.** A *partial order* on a set X is a relation $\leq$ that is:

reflexive: $\forall x \in X,\ x \leq x$

antisymmetric: $\forall x, y \in X,\ (x \leq y \wedge y \leq x) \implies x = y$

transitive: $\forall x, y, z \in X,\ (x \leq y \wedge y \leq z) \implies x \leq z$

We will write $x < y$ for "$x \leq y$ and $x \neq y$". This is:

irreflexive: $\forall x,\ \neg(x < x)$.

transitive: $\forall x, y, z,\ ((x < y) \wedge (y < z)) \implies x < z$.

Definition (Partially ordered set)**.** A *partially ordered set* or *poset* is a set X with a partial order.

Examples:

(1) Every linearly ordered set.

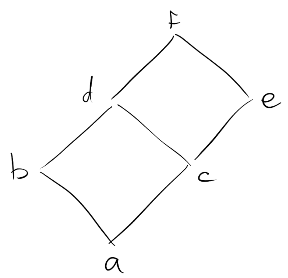
(2) $\mathbb{N}$ with $a \leq b \iff a \mid b$.

(3) For a set X $\mathbb{P}X$ with $a \leq b \iff a \subset b$.

(4) Every subset of a partially ordered set: for example, if G is a group, then

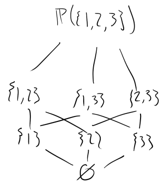
$$\{H \in \mathbb{P}G \mid H \text{ is a subgroup of } G\}$$

(5) Posets given by Hasse diagrams. For example

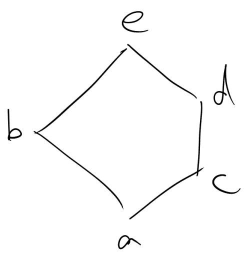


$X = \{a, b, c, d, e, f\}$. $b, c > a$, $d > b, c$, $e > c$, $f > d, e$ and all relations that follow by transitivity. ($e \not> b$, $f > a$).

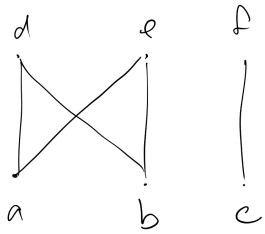
In general, a Hasse diagram for a partially ordered set X is a grawing of elements of X where we join x to y with an upward line if $y > x$ and $\nexists z$ with $y > z > x$. For example:



(6)



(7)



(8)

Start of lecture 9

Definition (Chain)**.** A subset S of a partially ordered set X is a *chain* if it is linearly ordered by the partial order on X.

Example.

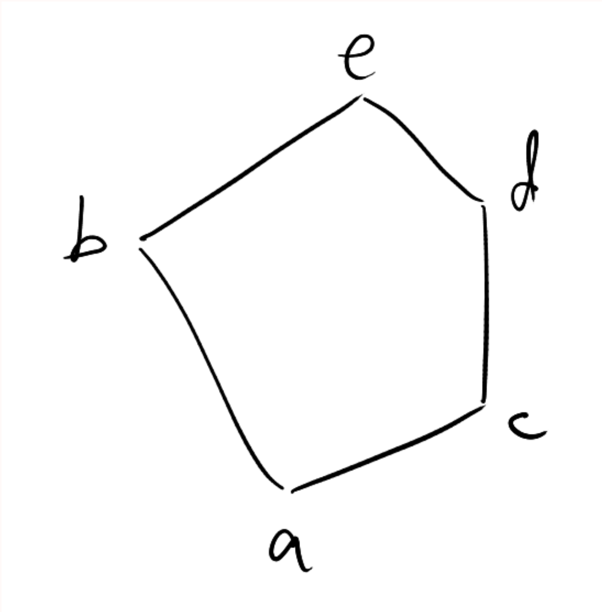
(1) Every linearly ordered set is a chain in itself.

(2) Any subset of a chain in a partially ordered set.

(3) In $\mathbb{N}$ with $a \leq b \iff a \mid b$, $\{2^n \mid n = 0, 1, 2, \ldots\}$ is a chain.

(4) In $\mathbb{P}(\{1, 2, 3\})$ with $\subset$, $\{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}\}$ is a chain.

(5) $\{a, c, d, e\}$ is a chain in



(6) In $\mathbb{PQ}$, $\{(-\infty, x) \cap \mathbb{Q} \mid x \in \mathbb{R}\}$ is an uncountable chain in $\mathbb{PQ}$.

Definition (Antichain)**.** A subset S of a partially ordered set X is an *antichain* if no two distinct members of S are related, i.e. $\forall x, y \in S, x \leq y \implies x = y$.

Example.

(1) In a linearly ordered set there is no antichain of size > 1.

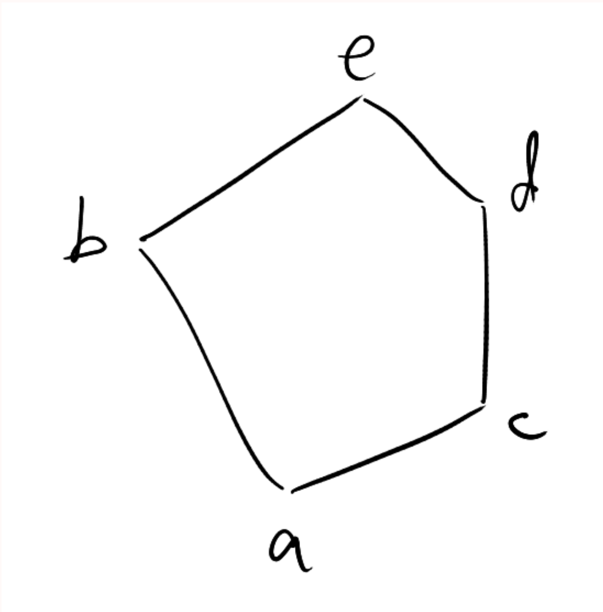
(2) In $\mathbb{N}$ with $a \leq b \iff a \mid b$, the set of primes is an antichain.

(3) In $\mathbb{P}(\{1, 2, \ldots, n\})$ with $\subset$, for any k, $0 \leq k \leq n$,

$$\mathcal{F}_k = \{A \subset \{1, \ldots, n\} \mid |A| = k\}$$

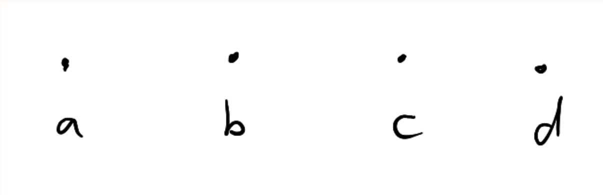
is an antichain.

(4) In



$\{b, d\}$ and $\{b, c\}$ are antichains.

(5) In



the whole set is an antichain.

Definition (Upper bound)**.** Let S be a subset of a partially ordered set X. Say $x \in X$ is an *upper bound* for S if $\forall y \in S$, $y \leq x$.

Definition (Least upper bound)**.** Say $x \in X$ is a *least upper bound* or *supremum* for S if x is an upper bound for S and $x \leq y$ for all upper bounds y for S.

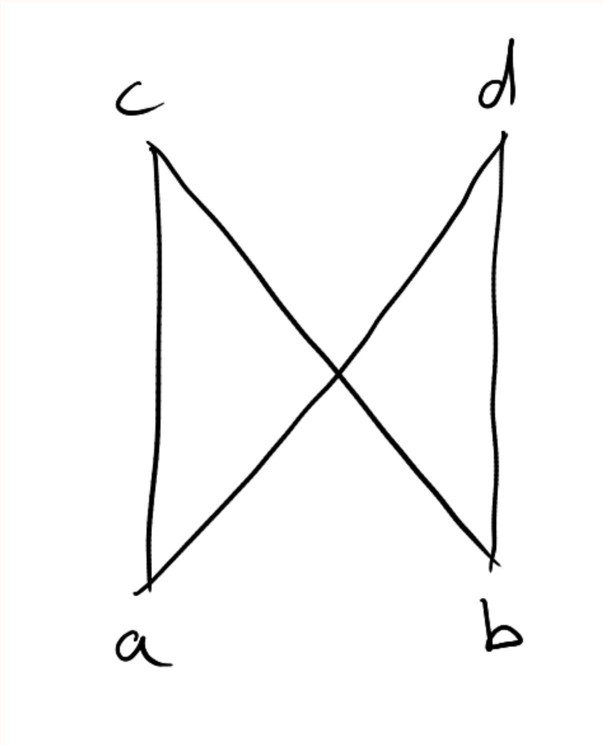
If it exists, we denote this by $\sup S$ or $\bigvee S$ ('join' of S).

Example.

(1) In $\mathbb{R}$, $\sup[0,1] = 1$, $\sup(0,1) = 1$.

(2) $\mathbb{Q}$ has no supremum in $\mathbb{Q}$, as it doesn't even have any upper bound.

(3) In



$\{a, b\}$ has upper bounds, for example c, d, but no least upper bound.

(4) If $X = \mathbb{P}A$, A any set, $S \subset X$, then $\sup S = \bigcup\{B \subset A \mid B \in S\}$.

Definition (Complete Partial Order)**.** A partially ordered set X is *complete* if every $S \subset X$ has a supremum.

Example.

1. $\mathbb{P}A$ for any A is complete.

2. $[0, 1]$ is complete.

3. $\mathbb{R}$ is not complete.

4. $\mathbb{Q} \cap [0, 2]$ is not complete.

Remark. A complete partially ordered set X has a greatest element $\sup X$ and a least element $\sup \emptyset$. In particular, $X \neq \emptyset$.

Definition (Order-preserving function)**.** Let $f : X \to Y$ be a function between partially ordered sets X, Y. Say f is *order-preserving* if $\forall x, y \in X$, $x \leq y \iff f(x) \leq f(y)$.

Note. f need not be injective. But f is order-preserving injective if and only if $\forall x, y \in X$, $x < y \iff f(x) < f(y)$.

Example. $f : \mathbb{N} \to \mathbb{N}$, $f(n) = n + 1$ (with the usual order).

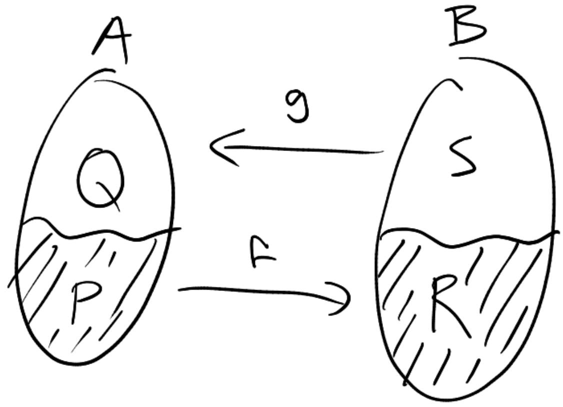
$g : \mathbb{P}(A) \to \mathbb{P}(A)$, $A \mapsto A \cup B$, B fixed.

Definition (Fixed point)**.** Let X be any set. Then a *fixed point* for a function $f : X \to X$ is an element $x \in X$ such that $f(x) = x$.

Theorem 1 (Knaster-Tarski Fixed Point Theorem)**.** If X is a complete partially ordered set and $f : X \to X$ is order-preserving, then f has a fixed point.

Proof. Let $S = \{x \in X \mid x \leq f(x)\}$. Let $z = \sup S$. Let $x \in S$. Then $x \leq z$, so $f(x) \leq f(z)$. Since $x \in S$, $x \leq f(x)$, so by transitivity, $x \leq f(z)$. Thus $f(z)$ is an upper bound for S, so $z \leq f(z)$. It follows that $f(z) \leq f(f(z))$. So $f(z) \in S$, and thus $f(z) \leq z$. So z is a fixed point. □

Corollary 2 (Schröder-Bernstein Theorem)**.** Let A, B be sets and assume there exist injections $f : A \to B$ and $g : B \to A$. Then there exists a bijection $h : A \to B$.



Proof. We seek partitions $A = P \cup \mathbb{Q}$, $B = R \cup S$ such that ($P \cap Q = \emptyset$, $R \cap S = \emptyset$), $f(P) = R$, $g(S) = Q$. Then we will have that

$$h : A \to B, \qquad h = \begin{cases} f & \text{on } P \\ g^{-1} & \text{on } Q \end{cases}$$

Such partitions exist if and only if there exists $P \subset A$ such that

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

Let $X = \mathbb{P}A$ with ordering by $\subset$. Define $H : X \to X$,

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

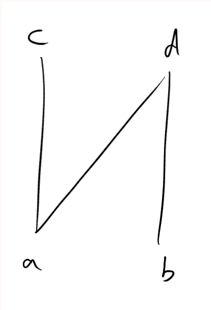
H is order-preserving and X is complete, so by Knaster-Tarski Fixed Point Theorem, we can find such P. □

Zorn's Lemma

Definition (Maximal element)**.** Say an element x in a partially ordered set X is *maximal* if $\forall y \in X$, $x \le y \implies x = y$. In other words, there is no $y \in X$ with $y > x$.

Example. In $\mathbb{P}A$, A is maximal, A is even a greatest element. In general, "greatest" $\implies$ maximal, but the other way round does not hold.

Example. In:



c, d are both maximal, but there does not exist a greatest element.

Theorem 3 (Zorn's Lemma)**.** Let X be a (non-empty) partially ordered set such that every chain in X has an upper bound in X. Then X has a maximal element.

Remark. $\emptyset$ is a chain in X, so it has an upper bound, so $X \neq \emptyset$. Often we check the chain condition by checking it for $\emptyset$ (i.e. that $X \neq \emptyset$) and then for $\neq \emptyset$ chains.

Proof. Assume X has no maximal element. For each $x \in X$, fix $x' > x$. We also fix an upper bound $u(C)$ for every chain $C \subset X$. Let $\gamma = \gamma(X)$ (from Hartog's Lemma). Define $f : \gamma \to X$ by Definition by recursion:

- $f(0) = u(\emptyset)$.
- $f(\alpha + 1) = f(\alpha)'$.
- $f(\lambda) = u(\{f(\alpha) \mid \alpha < \lambda\})'$ ($\lambda \neq 0$ limit ordinal).

An easy induction shows that $\forall \alpha < \beta$ (in γ), $f(\alpha) < f(\beta)$ (on β, α, α fixed). This also shows $\{f(\alpha) \mid \alpha < \beta\}$ is a chain for all $\beta < \alpha$. Hence f is an injection. This contradicts the definition of $\gamma(X)$. □

Remark. Technically, for $\lambda \neq 0$ a limit ordinal, $f(\lambda)$ should be defined as above if $\{f(\alpha) \mid \alpha < \gamma\}$ is a chain and $f(\lambda) = u(\emptyset)$ otherwise. Then by induction, $\alpha < \beta \implies f(\alpha) < f(\beta)$, so the 'otherwise' clause never happens.

Start of lecture 10

Warning. Recall that when studying linearly ordered sets, we noted that

$$f \text{ is order-preserving and injective} \iff \forall x, y \in A, x < y \implies f(x) < f(y).$$

The $\Rightarrow$ direction is true for partially ordered sets, but the $\Leftarrow$ direction is not true in general for a partially ordered set.

Applications of Zorn's Lemma

Theorem 4. Every vector space V (over some field) has a basis.

Proof. We seek a maximal linearly independent set $B \subset V$. Then we're done: if $V \neq \langle B \rangle$, then for any $x \in V \setminus \langle B \rangle$, $B \cup \{x\}$ is also linearly independent, which would contradict maximality of B.

Let $X = \{A \subset v \mid A \text{ is linearly independent}\}$ ordered by inclusion. Let $\{A_i \mid i \in I\}$ be a chain in X. Then this has upper bound $A = \bigcup_{i \in I} A_i$. We first need to check that A is linearly independent. Assume $\sum_{j=1}^n \lambda_j x_j = 0$ is a linear relation on A (where $x_1, \dots, x_n \in A$, and $\lambda_1, \dots, \lambda_n$ are scalars). For each $1 \le j \le n$, pick $i_j \in I$ such that $x_j \in A_{i_j}$. Since the A_i form a chain, there exists $1 \le m \le n$ such that $A_{i_j} \subset A_{i_m}$ for all $1 \le j \le n$. Then $\sum_{j=1}^n \lambda_j x_j = 0$ is a linear relation on the linearly independent set A_{i_m}, so $\lambda_1 = \cdots = \lambda_n = 0$. Thus A is linearly independent. □

Remark.

(1) A very similar proof shows that if $B_0 \subset V$ is linearly independent, then V has a basis B such that $B \supset B_0$.

(2) $\mathbb{R}$ is a vector space over $\mathbb{Q}$, so has a basis (Hamel basis). This can be used to show the existence of non-Lebesgue-measurable sets (see Probability & Measure).

(3) $\mathbb{R}^{\mathbb{N}}$ the real vector space of real sequences has no countable basis, but we now know it has a basis.

(4) In topology: Tychonoff's Theorem. In Functional Analysis: Hahn-Banach Theorem. In algebra: maximal ideals in rings with 1.

The next application of Zorn's Lemma completes the proof of Model Existence Lemma:

Theorem 5. Let P be any set of primitive proposition, $S \subset L = L(P)$ be consistent. Then there exists a consistent set $\overline{S} \subset L$ such that $S \subset \overline{S}$ and $\forall t \in L$ either $t \in \overline{S}$ or $\neg t \in \overline{S}$.

Proof. We seek a maximal consistent set $\overline{S} \supset S$. Then we're done as follows: given $t \in L$, one of $S \cup \{t\}$ and $S \cup \{\neg t\}$ is consistent, otherwise $S \cup \{t\} \vdash \perp$, $\overline{S} \cup \{\neg t\} \vdash \perp$, and so by the Deduction Theorem, $\overline{S} \vdash \neg t$, $\overline{S} \vdash \neg\neg t$ and hence $\overline{S} \vdash \perp$ by MP, contradiction. Hence by maximality of $\overline{S}$, either $t \in \overline{S}$ or $\neg t \in \overline{S}$.

Let $X = \{T \subset L \mid S \subset T, T \text{ is consistent}\}$, partially ordered by $\subset$. $X \neq \emptyset$ since $S \in X$. Let $C = \{T_i \mid i \in I\}$ be a non-empty chain in X. Let $T = \bigcup_{i \in I} T_i$. Then $S \subset T$ ($I \neq \emptyset$). If $T \vdash \perp$ then as proofs are finite, there exists finite $J \subset I$ such that $\bigcup_{j \in J} T_j \vdash \perp$. Since C is a chain, there exists $j_0 \in J$ such that $\bigcup_{j \in J} T_j = T_{j_0}$, so $T_{j_0} \vdash \perp$, contradiction. By Zorn's Lemma, X has a maximal element. □

Theorem 6 (Well-ordering principle)**.** Every set can be well-ordered.

Example. $\mathbb{R}$ can be well-ordered. Think about this for a bit. This feels very unnatural!

Proof. Let A be a set. Let

$$X = \{(B, R) \in \mathbb{P}A \times \mathbb{P}(A \times A) \mid R \text{ is a well-ordering of } B\}$$

partially ordered by extension: $(B_1, R_1) \leq (B_2, R_2)$ if and only if $B_1 \subset B_2$, $R_1 = R_2 \cap (B_1 \times B_1)$ (R_1 is the restriction of R_2 to B_1) and B_1 is an initial segment of B_2. Note $X \neq \emptyset$, since $(\emptyset, \emptyset) \in X$.

Let $C = \{(B_i, R_i) \mid i \in I\}$ be a chain in X, i.e. a nested set of well-ordered sets. Then

$$\left(\bigcup_{i \in I}, \bigcup_{i \in I} R_i \right)$$

is an upper bound as in Section 2.

By Zorn's Lemma, X has a maximal element (B, R). We need $B = A$. If not, pick $x \in A \setminus B$, then

$$(B, R)^+ = (B \cup \{x\}, R \cup \{(b, x) \mid b \in B\}) \in X$$

and $(B, R) < (B, R)^+$, contradiction. □

Remark. Often in applications of Zorn's Lemma, the maximal object whose existence it asserts cannot be described explicitly ("magical").

The Axiom of Choice (AC)

In the proof of Zorn's Lemma we used two functions:

$$\begin{aligned} X &\to X \\ x &\mapsto x' \in \{y \mid y > x\} \\ u : \{C \subset X \mid C \text{ is a chain}\} &\to X \\ u(C) &\in \{x \in X \mid x \text{ is an upper bound for } C\} \end{aligned}$$

These are known as choice functions.

Axiom of Choice says:

For any set $\{A_i \mid i \in I\}$ of non-empty sets, there exists a function $f : I \to \bigcup_{i \in I} A_i$ such that $f(i) \in A_i$ for all $i \in I$. We call this a *choice function.*

This is different in character from other rules for building sets ($\cup$, $\mathbb{P}$ etc) in the sense that choice functions need not be unique. For this reason, we're often interested in proving things without axiom of choice.

Note. When I is finite, we can prove existence of choice functions by induction on $|I|$.

Theorem 7. The following are equivalent:

(i) Axiom of choice.

(ii) Zorn's Lemma.

(iii) Well-ordering principle.

Proof.

AC $\Rightarrow$ ZL See proof of Theorem 3.

ZL $\Rightarrow$ WO See proof of Theorem 6.

WO $\Rightarrow$ AC Let $\{A_i \mid i \in I\}$ be a set of non-empty sets. Let $A = \bigcup_{i\in I} A_i$. Well order A and define $f : I \to A$ by setting $f(i)$ to be the least element of A_i. $\square$

Exercise: Prove the implications directly.

Start of lecture 11

**** Non-examinable ****

Definition (Chain-complete)**.** A partially ordered set X is *chain-complete* if $X \neq \emptyset$ and every chain has a supremum.

Example. Every complete partially ordered set is chain-complete. Finite non-empty partially ordered sets are chain-complete. If S is a partially ordered set, then

$$X = \{C \subset C \mid C \text{ is a chain}\}$$

ordered by $\subset$ is chain-complete, but not complete in general.

Definition (Inflationary function)**.** A function $f : X \to X$, X a partially ordered set is *inflationary* if $x \leq f(x)$ for all $x \in X$.

Theorem (Bourbak-Witt fixed point theorem)**.** If X is chain-complete and $f : X \to X$ is inflationary, then f has a fixed point.

Proof 1 (with axiom of choice). By Zorn's Lemma, X has a maximal element. Then $x \leq f(x)$, so $x = f(x)$. $\square$

Proof 2 (without axiom of choice). Fix $x_0 \in X$. Let $\gamma = \gamma(X)$. Define $g : \gamma \to X$ by recursion:

- $g(0) = x_0$
- $g(\alpha + 1) = f(g(\alpha))$
- $g(\lambda) = \sup\{g(\alpha) \mid \alpha < \lambda\}$ ($\lambda \neq 0$ limit)

By induction $\forall \alpha < \gamma$, $g(\alpha) \leq g(\alpha + 1)$

Either there exists $\alpha < \gamma$ with $g(\alpha + 1) = g(\alpha)$. Then $g(\alpha)$ is a fixed point of f. Otherwise g is injective, which would contradict Hartog's Lemma. $\square$

Remark. Axiom of Choice and Bourbak-Witt fixed point theorem implies Zorn's Lemma. Bourbak-Witt fixed point theorem is sometimes called "the choice-free part of the proof of Zorn's Lemma".

Proof of Remark. Let X be a partially ordered set in which every chain has an upper bound.

Case 1: X is chain-complete. Assume X has no maximal element. Fix a choice function $g; (\mathbb{P}X) \setminus \{\emptyset\} \to X$. Define

$$f : X \to X, f(x) = g(\{y \in X \mid x < y\}).$$

Then $x < f(x)$ $\forall x \in X$, contradicting Bourbak-Witt fixed point theorem.

Case 2: General case. We first prove that $\mathcal{C} = \{C \subset X \mid C \text{ is a chain}\}$ has a maximal element. (This is the Hausdorff Maximality Principle). Follows from Case 1, since $\mathcal{C}$ is chain-complete.

Let C be a maximal chain in X. Let x be an upper bound of C. If $x < y$ in X, then $C \cup \{y\}$ is a chain which is $\supsetneq C$, contradicting maximality. So x is maximal element. $\square$

Lattices, Boolean algebras – not covered (for now)

This is the end of the non-examinable part.

4 First-order Predicate Logic

In Propositional Logic we had a set P of primitive propositions and then we combined them using logical connectives $\Rightarrow, \perp$ (and shorthands $\wedge, \vee, \neg, \top$) to form the language $L = L(P)$ of all (compound) propositions. We attached no meaning to primitive propositions.

Aim: To develop languages to describe a wide range of mathematical theorems. We will replace primitive propositions with mathematical statements.

Example. In language of groups:

$$m(x, m(y, z)) = m(m(x, y), z), \qquad m(x, i(x)) = e.$$

In language of partially ordered sets:

$$x \leq y.$$

This will need variables $(x, y, z, \ldots)$, operation symbols (m, i, e with arities $2, 1, 0$ respectively) and predicates (for example $\leq$ with arity 2). Note that "arity" means the number of elements that the function takes as input.

We will then combine these to build formulae:

Example. In the language of groups:

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

In the language of partially ordered sets:

$$(\forall x)(\forall y)(\forall z)((x \leq y \wedge y \leq z) \implies (x \leq z)).$$

Valuations will be replaced by a structure, a set A and "truth-functions" $p_A : A^n \to \{0, 1\}$ for every formula p.

If we have a set S of formulae, a model of S is a structure satisfying all $p \in S$. Then we will define $S \models t$ in the same way as in Section 1. $S \vdash t$ will be the same as in Section 1 but more complex.

Definition (Language in first-order logic)**.** A *language* in first-order logic is specified by two disjoint sets Ω (the *set of operation symbols*) and Π (the *set of predicates*) together with an arity function $\alpha : \Omega \cup \Pi \to \mathbb{N}_0 = \{0\} \cup \mathbb{N}$.

The language $L = L(\Omega, \Pi, \alpha)$ consists of the following:

Variables: Countably infinite sets disjoint from Ω and Π. We denote variables as $x_1, x_2, x_3, \dots$ (or $x, y, z, \dots$).

Terms: Defined inductively:

(i) Every variable is a term

(ii) If $\omega \in \Omega$, $n = \alpha(\omega)$ and $t_1, \dots, t_n$ terms, then $\omega t_1 \dots t_n$ is a term (could write $\omega(t_1, \dots, t_n)$).

Example. The language of groups consists of $\Omega = \{m, i, e\}$, $\Pi = \emptyset$, $\alpha(m) = z$, $\alpha(i) = 1$, $\alpha(e) = 0$. Some terms:

$$m \underbrace{x}_{t_1} \underbrace{myz}_{t_2}, \qquad mmxyz, \qquad mxix, \qquad e.$$

Note. Every operation symbol of arity 0 is a term, called a *constant*.

Definition (Atomic formula)**.** There are two types of *atomic formula*:

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

(ii) If $\varphi \in \Pi$ with $\alpha(\varphi) = n$ and $t_1, \dots, t_n$ are terms, then $\varphi t_1 t_2 \dots t_n$ is an atomic formula.

Example. The language of partially ordered sets consists of $\Omega = \emptyset$, $\Pi = \{\leq\}$, $\alpha(\leq) = 2$. Some atomic formulae:

$$x = y, \qquad x \leq y \quad (\text{officially} \leq xy)$$

Definition (Formula)**.** We define *formulae* inductively:

(i) atomic formulae are formulae.

(ii) $\perp$ is a formula.

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

(iv) If p is a formula and the variable x has a *free occurrence* in p, then $(\forall x)p$ is a formula.

Note. A formula is a finite string of symbols from the set of variables, Ω, Π and $\{(,), \Rightarrow, \perp, =, \forall\}$.

Start of lecture 12

Notation. We also introduce the symbols $\wedge$, $\vee$, $\neg$ and $\top$ as in Section 1, and we also introduce the new symbol $(\exists x)p$ for $\neg(\forall x)\neg p$.

Definition (Free occurence)**.** An occurence of a variable x in a formula p is always *free* except if $p = (\forall x)q$, in which case the $\forall x$ quantifier *binds* every free occurence of x, and then such occurences of x are called *bound* occurences (the formal definition is by induction in L). Note that since the symbol $\exists$ implicitly uses a $\forall$, this symbol can also bind free occurences of a variable.

Example. In the language of groups:

$$(\exists x)(mxx = y) \Rightarrow (\forall z)\neg(mmzzz = y)$$

Here the occurences of x and z are bound, while the occurences of y are free.

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

has no free variables.

$$(\exists x)(mxx = y) \Rightarrow (\forall y)(\forall x)(myz = mzy)$$

Technically the above is a correct formula, where y occurs both as a free variable and a bound variable, but in practise we avoid this.

In the language of partially ordered sets:

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

has no free variables.

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

Definition (Free variables)**.** A variable x in a formula is *free* if it has a free occurence in p. Let $\mathrm{FV}(p)$ denote the set of free variables in p.

Definition (L-structure)**.** Let $L = L(\Omega, \Pi, \alpha)$ be a first-order folang. A *structure* in L (or L-*structure*) is a non-empty set A together with a function $\omega_A : A^n \to A$ for every $\omega \in \Omega$ where $n = \alpha(\omega)$ and subsets $\varphi_A \subset A^n$ for every $\varphi \in \Pi$ where $n = \alpha(\varphi)$ (or equivalently $\varphi_A : A^n \to \{0, 1\}$ by identifying a set with its indicator function).

Example. In language of groups: a structure is a non-empty set A with functions $m_A : A^2 \to A$, $i_A : A \to A$, $e_A \in A$ (A^0 is the singleton set). (An operation symbol with arity 0 is called a *constant*). This is not a group yet!

In the language of partially ordered sets: a structure is a non-empty set A with $\leq_A \subset A^2$, i.e. a relation on A. This is not yet a partially ordered set.

Next step: to define for a formula p what it means that "p is satisfied in A".

Example. $p = (\forall x)(mxix = e)$ in language of partially ordered sets. p satisfied in a structure A should mean that for all $a \in A$ we have $m_A(a, i_A(a)) = e_A$.

Here is the formal definitoion in a language $L = L(\Omega, \Pi, \alpha)$:

Definition (Interpretation of a term)**.** Let A be an L-structure. A term t in L with $\mathrm{FV}(t) \subset \{x_1, \dots, x_n\}$ has *interpretation* $t_A : A^n \to A$ defined as follows:

- If $t = x_i$, $1 \le i \le n$, then $t_A(a_1, \dots, a_n) = a_i$.
- If $t = \omega t_1 \cdots t_m$ ($\omega \in \Omega$, $m = \alpha(\omega)$, $t_1, \dots, t_m$ terms), then

$$t_A(a_1, \dots, a_n) = \omega_A((t_1)_A(a_1, \dots, a_n), \dots, (t_m)_A(a_1, \dots, a_n))$$

Example. In groups,

$$t = m \underbrace{x_1}\underbrace{mx_2x_3}$$

has interpretation

$$t_A(a_1, a_2, a_3) = m_A(a_1, m_A(a_2, a_3)).$$

Definition (Interpretation of a formula)**.** We interpret a formula p with $\mathrm{FV}(p) \subset \{x_1, \ldots, x_n\}$ as a subset $p_A \subset A^n$ (or equivalently as a function $p_A : A^n \to \{01\}$).

- If $p = (s = t)$, then

$$p_A(a_1, \ldots, a_n) = 1 \iff s_A(a_1, \ldots, a_n) = t_A(a_1, \ldots, a_n)$$

- If $p = \varphi t_1 \cdots t_m$ ($\varphi \in \Pi$, $m = \alpha(\varphi)$, $t_1, \ldots, t_m$ terms), then

$$p_A(a_1, \ldots, a_n) = 1 \iff \varphi_A((t_1)_A(a_1, \ldots, a_n), \ldots, (t_m)_A(a_1, \ldots, a_n)) = 1$$

- $\perp_A$ is the constant 0 function.
- $p = (q \implies r)$:

$$p_A(a_1, \ldots, a_n) = 0 \iff q_A(a_1, \ldots, a_n) = 1 \quad \text{and} \quad r_A(x_1, \ldots, a_n) = 0$$

- $p = (\forall x_{n+1})q$ where $\mathrm{FV}(q) \subset \{x_1, \ldots, x_{n+1}\}$:

$$p_A = \{(a_1, \ldots, a_n) \in A^n \mid (a_1, \ldots, a_n, a_{n+1}) \in q_A \text{ for all } a_{n+1} \in A\}$$

Example. In groups, if $p = (mmxyz = mxmyz)$ has interpretation

$$p_A = \{(a, b, c) \in A^3 \mid m_A(m_A(a, b), c) = m_A(a, m_A(b, c))\}.$$

The formula $q = (\forall x)(\forall y)(\forall z)p$ has interpretation $q_A = 1$ if and only if $p_A = A^3$.

Definition (Satisfied formula)**.** A formula p in a language L is *satisfied* in an L-structure A if $p_A = A^n$ (n is the number of free variables in p), or equivalently p_A is the constant 1 function. We also say *p holds in A* or *p is true in A* or *A is a model for p*.

Definition (Theory)**.** A *theory* in a language L is a set of sentences in L.

Definition (Model-defn)**.** A *model* for a theory T is an L-structure A that is a model for all $p \in T$.

Examples

(1) Theory of groups: the language is specified by $\Omega = \{m, i, e\}$ (with arities $2, 1, 0$ respectively) and $\Pi = \emptyset$. The theory is

$$\begin{aligned} T = \{&(\forall x)(\forall y)(\forall z)(mmxyz = mxmyz), \\ &(\forall x)((mxe = x) \wedge (mex = x)), \\ &(\forall x)((mxix = e) \wedge (mixx = e))\} \end{aligned}$$

Then models for T are precisely groups. So we can axiomatise groups as a first-order theory.

(2) Partially ordered sets $\Omega = \emptyset$, $\Pi = \{\le\}$ (with arity 2).

$$\begin{aligned} T = \{&(\forall x)(x \le x), \\ &(\forall x)(\forall y)(((x \le y) \wedge (y \le x)) \Rightarrow (x = y), \\ &(\forall x)(\forall y)(\forall z)(((x \le y) \wedge (y \le z)) \Rightarrow (x \le z))\} \end{aligned}$$

Then models are precisely partially ordered sets.

Start of lecture 13

(3) Theory of rings with 1: Language:

$$\Omega = \{+, 0, -, \times, 1\}, \quad \Pi = \emptyset,$$

with arities $2, 0, 1, 2, 0$. Theory:

$$\begin{aligned} &(\forall x)(\forall y)(\forall z)((x + y + z = x + (y + z)) \\ &(\forall x)(x + 0 = x \wedge 0 + x = x) \\ &(\forall x)((x + (-x) = 0) \wedge ((-x) + x = 0)) \\ &(\forall x)(\forall y)(x + y = y + x) \\ &(\forall x)(\forall y)(\forall z)((x \times y) \times z = x \times (y \times z)) \\ &(\forall x)(1 \times x \wedge x \times 1 = x) \\ &(\forall x)(\forall y)(\forall z)((x \times (y + z) = x \times y + x \times z) \wedge ((x + y) \times z = x \times z + y \times z)) \end{aligned}$$

The models are exactly rings with 1.

(4) Fields: Language: same as for rings with 1. Theory: same as for rings with 1, plus the additional sentences:

$$\begin{aligned} &(\forall x)(\forall y)(x \times y = y \times x) \\ &\neg(0 = 1) \\ &(\forall x)(\neg(x = 0) \Rightarrow (\exists y)(xy = 1)) \end{aligned}$$

The models are exactly fields.

(5) Graph theory: Language:

$$\Omega = \emptyset, \quad \Pi = \{a\}$$

with arity 2 (a will mean "is adjacent to"). Theory:

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

The models are exactly graphs.

(6) Propositional theories: Language:

$$\Omega = \emptyset, \quad \Pi = \text{some set}$$

with $\alpha(p) = 0 \ \forall p \in \Pi$. A structure is a non-empty set A together with $p_A \subset A^0$ for all $p \in \Pi$ (equivalently $p_A : A^0 \to \{0,1\}$, equivalently $p_A \in \{0,1\}$, since A^0 is a set of size 1). A structure is a non-empty set A together with a function $v : \Pi \to \{0,1\}$. Every $p \in \Pi$ is an atomic formula. Formulae without variables are precisely elements of $L(\Pi)$ as defined in Section 1, i.e. they are propositions in Π.

Interpreting these in a structure A is just a function $v : L(\Pi) \to \{0,1\}$ obtained from $v : \Pi \to \{0,1\}$ as in Section 1, i.e. a valuation. A *propositional theory* is a set S of formulae not using variables. A model for S is a non-empty set A with a valuation $v : L(\Pi) \to \{0,1\}$ such that $v(s) = 1 \ \forall s \in S$ (here A is irrelevant).

Definition (Semantic entailment of sentences)**.** For a set S of sentences and a sentence t (in a first-order language L), we say S *(semantically) entails* t if t is satisfied in every model of S. In this case we write $S \models t$.

Example.

Let S be the theory of groups (in the language of groups). Then

$$S \models ((\forall x)(x \cdot x = e) \Rightarrow (\forall x)(\forall y)(xy = yx))$$

Let S be the theory of fields (in the language of rings with 1). Then

$$S \models ((\forall x)(\neg(x = 0) \Rightarrow (\forall y)(\forall z)((xy = 1 \wedge xz = 1) \Rightarrow (y = z)))$$

Next, we want to define $S \models t$ for formulae.

Example. Let T be the theory of fields (in the language of rings with 1). Let $S = T \cup \{\neg(x = 0)\}$, $t = (\exists y)(xy = 1)$. Does $S \models t$? Yes.

Suppose F is a structure in which all members of S are true. So F is a field and for $u = \wedge(x = 0)$,

$$u_F = \{a \in F \mid a \neq 0_f\} = F,$$

contradiction. Also, we'll soon define "$S \vdash t$", then $S \vdash t$ if and only if $T \vdash \neg(x = 0) \Rightarrow (\exists y)(xy = 1)$.

Definition (Semantic entailment of formulae)**.** Let S be a set of formulae and t be a formula in a language L. For every variable that occurs free in $S \cup \{t\}$, introduce a constant c_x (add it to Ω). Let L' be our new language. For a formula p, let p' be the formula obtained from p by replacing free occurences of x in p by c_x, for every x. Let $S' = \{s' \mid s \in S\}$. Say S *(semantically) entails* t, written $S \models t$, if $S' \models t'$.

Notation (Substitutions)**.** If x occurs free in a formula p and t is a term that contains no variable that occurs bound in p, we let $p[t/x]$ be the formula obtained from p by replacing free occurences of x in p by t.

Example. In the language of groups: let $p = (\forall y)(mxx = y)$. Then:

$t = mzz$	$p[t/x] = (\forall y)(mmzzmzz = y)$
$t = mzy$	cannot be used
$t = mxx$	$p[t/x] = (\forall y)(mmxxmxx = y)$

Syntactic entailment

Definition (Axioms of first-order logic)**.**

(A1) $p \Rightarrow (q \Rightarrow p)$ (p, q are formulae).

(A2) $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$ (p, q, r any formulae).

(A3) $\neg\neg p \Rightarrow p$ (p any formula).

(A4) $(\forall x)(x = x)$.

(A5) $(\forall x)(\forall y)((x = y) \Rightarrow (p \Rightarrow p[y/x]))$ (x, y distinct variables, p a formula, $x \in \mathrm{FV}(p)$, y does not occur bound in p).

(A6) $((\forall x)p) \Rightarrow p[t/x]$ (p formula $x \in \mathrm{FV}(p)$, t a term, no variable in t occurs bound in p).

(A7) $(\forall x)(p \Rightarrow q) \Rightarrow (p \Rightarrow (\forall x)q)$ (p, q formulae, $x \notin \mathrm{FV}(p)$, $x \in \mathrm{FV}(q)$).

Note. Every axiom is a tautology (t is a tautology if $\emptyset \models t$, i.e. t holds in every structure).

Rules of deduction

Modus ponens (MP) From p and $p \Rightarrow q$, can deduce q.

Generalisation (Gen) From p such that $x \in \mathrm{FV}(p)$, can deduce $(\forall x)p$ provided x did not occur free in any of the premises used in the proof of p.

Start of

lecture 14

Definition (Proof (in first-order logic))**.** Let S be a set of formulae, and p a formula. A *proof of* p *from* S is a finite sequence $t_1, \ldots, t_n$ of formulae such that $t_n = p$ and for every i, we have one of:

- $t_i \in S$ or t_i is an axiom.
- $\exists j, k < i$ with $t_k = (t_j \Rightarrow t_i)$.
- $\exists j < i$ with $t_i = (\forall x)t_j$, $x \in \mathrm{FV}(t_j)$ and for all $k < j$ if $t_k \in S$ then x does not occur free in t_k.

In this case we say S *proves* p and write $S \vdash p$.

(If S is a theory and p is a sentence then we say p *is a theorem of* S).

Remark. Suppose we allow $\emptyset$ as a structure. Note that $(\forall x)\neg(x = x)$ is satisfied in $\emptyset$, whereas $\bot$ is not. So $\{(\forall x)\neg(x = x)\} \not\models \bot$. However, $\{(\forall x)\neg(x = x)\} \vdash$:

$(\forall x)\neg(x = x)$	(premise)
$((\forall x)\neg(x = x)) \Rightarrow (\neg(x = x))$	(A6)
$\neg(x = x)$	(MP)
$(\forall x)(x = x)$	(A4)
$(x = x)$	(A6 + MP)
$\bot$	(MP)

Example. $\{x = y\} \vdash (y = x)$.

$(\forall x)(\forall y)((x = y) \Rightarrow ((x = z) \Rightarrow (y = z))$	(A5)
$(x = y) \Rightarrow ((x = z) \Rightarrow (y = z))$	((A6 + MP) twice)
$x = y$	(premise)
$(x = z) \Rightarrow (y = z)$	(MP)
$(\forall z)((x = z) \Rightarrow (y = z))$	(Gen)
$((\forall z)((x = z) \Rightarrow (y = z))) \Rightarrow ((x = x) \Rightarrow (y = z))$	(A6)
$(x = x) \Rightarrow (y = x)$	(MP)
$(\forall x)(x = x)$	(A4)
$(x = x)$	(A6 + MP)
$(y = x)$	(MP)

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

Proof.

$\Rightarrow$ Write down a proof of $p \Rightarrow q$ from S and add the lines:

p	(premise)
q	(MP)

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

$\Leftarrow$ Let $t_1, \ldots, t_n = q$ be a proof of q from $S \cup \{p\}$. We proe $S \vdash (p \Rightarrow t_i)$ by induction on i.

Our induction hypothesis at step i will be: for $j < i$, $S \vdash (p \Rightarrow t_j)$ such that if the proof of t_j from $S \cup \{p\}$ did not use any premise in which a variable x occurs free, then the proof of $(p \Rightarrow t_j)$ from S does not use any premise in which a variable x occurs free.

To see $S \vdash (p \Rightarrow t_i)$, we consider cases:

- If $t_i \in S$ or t_i an axiom, write

t_i	(premise or axiom)
$t_i \Rightarrow (p \Rightarrow t_i)$	(A1)
$p \Rightarrow t_i$	(MP)

  is a proof of $(p \Rightarrow t_i)$ from S.

- If $t_i = p$, then write down a proof of $p \Rightarrow p$ from $\emptyset$.

- If $\exists j, k < i$ with $t_k = (t_j \Rightarrow t_i)$ then write

$(p \Rightarrow (t_j \Rightarrow t_i)) \Rightarrow ((p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i))$	(A2)
$p \Rightarrow (t_j \Rightarrow t_i)$	(by induction hypothesis)
$(p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)$	(MP)
$p \Rightarrow t_j$	(by induction hypothesis)
$p \Rightarrow t_i$	(MP)

- Finally, if $\exists j < i$ such that $x \in \mathrm{FV}(t_j)$ and $t_i = (\forall x)t_j$, then the proof of t_j from $S \cup \{p\}$ did not use any premise in which x occurs free.

If x occurs free in p, then p did not occur in proof of t_j from $S \cup \{p\}$, i.e. it is a proof of t_j from S. By (Gen), $S \vdash (\forall x)t_j$, i.e. $S \vdash (\forall x)t_j$, i.e. $S \vdash t_i$. Add the lines

$$
\begin{array}{ll}
t_i \Rightarrow (p \Rightarrow t_i) & \text{(A1)} \\
p \Rightarrow t_i & \text{(MP)}
\end{array}
$$

If x does not occur free in p, then we have a proof of $p \Rightarrow t_j$ from S by induction hypothesis, which does not use any premise in which x occurs free. So we can add:

$$
\begin{array}{ll}
(\forall x)(p \Rightarrow t_j) & \text{(Gen)} \\
((\forall x)(p \Rightarrow t_j)) \Rightarrow (p \Rightarrow (\forall x)t_j) & \text{(A7)} \\
\underbrace{p \Rightarrow (\forall x)t_j}_{= p \Rightarrow t_i} & \text{(MP)}
\end{array}
$$

In all cases the condition about free variables remains true. □

Aim: $S \vdash p$ if and only if $S \models p$.

Proposition 2 (Soundness Theorem)**.** Let S be a set of formulae and p be a formula. If $S \vdash p$ then $S \models p$.

Proof (non-examinable). Write down a proof $t_1, \dots, t_n$ of p from S. Verify thet $S \models t_i$ by an easy induction. □

Theorem 3 (Model Existence Lemma)**.** Let S be a consistent theory in the language $L = L(\Omega, \Pi, \alpha)$ (i.e. $S \nvdash \bot$). Then S has a model.

Assuming this, we have:

Corollary 4 (Adequacy Theorem)**.** Let S be a set of formulae and p be a formula. If $S \models p$, ten $S \models p$.

Proof (non-examinable). Without loss of generality S is a theory and p is a sentence (by using the definition of $\models$ in the case where we have formulae rather than sentences). Since $S \models p$, $S \cup \{\neg p\} \models \bot$. So by Theorem 3, $S \cup \{\neg p\} \vdash \bot$. So $S \vdash \neg\neg p$ (by Proposition 1), so $S \vdash p$ by (A3) and (MP). □

Theorem 5 (Gödel's Completeness Theorem for first-order logic)**.** If S is a set of formulae and p is a formula, then $S \vdash p$ if and only if $S \models p$.

Idea of proof of Theorem 3: We build a model from $L = L(\Omega, \Pi)$. Let A be the set of *closed* terms in L, i.e. terms with no variables. For example S = theory of fields (in language of commutative rings with 1). A consists of

$$1 + 1, (((1 + 0) + 0) + 1), 1 \cdot 1, 1 \cdot 0, \dots, 1 + (-1), \dots$$

We will define the interpretation of + (and other symbols similarly) using:

$$(1 + 1) +_A (1 + 0) = (1 + 1) + (1 + 0)$$

If S is the theory of fields, then A is not a model:

Start of lecture 15

$$1 + 1 = (1 + 0) + 1$$

is provable from S, but not satisfied in A:

$$(1 + 1)_A = 1 + 1, \qquad ((1 + 0) + 1)_A = (1 + 0) + 1.$$

Easy remedy: define $s \sim t$ on A if and only if $S \vdash (s = t)$, and then replace A with $A/\sim$. Two issues remain.

Let S be the theory of fields plus the sentence $(1 + 1 = 0 \vee (1 + 1) + 1 = 0)$ (the theory of fields of characteristic 2 or 3). $S \nvdash 1 + 1 = 0$, so in our new A

$$1_A +_A 1_A = [1] +_A [1] = [1 + 1] \neq [0]_A = 0_A.$$

Similarly

$$(1_A +_A 1_A) +_A 1_A \neq 0_A.$$

So A is not a model of S. Remedy: extend S to a consistent theory $\overline{S} \supset S$ such that for every sentence p, either $\overline{S} \vdash p$ or $\overline{S} \vdash \neg p$. Such a theory is called *complete*.

Now consider S being the theory of fields plus $((\exists x)(xx = 1 + 1))$. A is not a model since there's no closed term t such that

$$[t] \cdot [t] = [1] +_A [1] = 1_A +_A 1_A$$

because $S \nvdash (t \cdot = 1 + 1)$. We say S has *witnesses* if for every sentence of the form $(\exists p)$, where $\mathrm{FV}(p) = \{x\}$, such that $S \vdash (\exists x)p$, there exists a closed term t such that $S \vdash p[t/x]$. We will enlarge S to a consistent theory $\overline{S}$ such that $\overline{S}$ will have witnesses for S.

Proof of Theorem 3 (non-examinable). We start with two observations. Let S be a first-order consistent theory in a language $L = L(\Omega, \Pi)$. For any sentence p, at least one of $S \cup \{p\}$ or $S \cup \{\neg p\}$ is consistent. Otherwise they both $\vdash \perp$, so by Deduction Theorem, $S \vdash \neg p$ and $S \vdash \neg\neg p$. Hence $S \vdash \perp$ by MP, contradiction. An argument using Zorn's Lemma gives a consistent $\overline{S} \supset S$ such that for every sentence p, either $p \in \overline{S}$ or $\neg p \in \overline{S}$. So $\overline{S}$ is complete.

Now assume S is consistent and $S \vdash (\exists x)p$ for some p with $\text{FV}(p) = \{x\}$. We add a new constant c to L ($\Omega \to \Omega \cup \{c\}$). Then $S \cup \{p[c/x]\}$ is consistent. If not, then $S \cup \{p[c/x]\} \vdash \perp$, so $S \vdash \neg p[c/x]$. Since c does not occur in S, we get $S \vdash \neg p$ (put x back in place of c in the proof). So by (Gen), $S \vdash (\forall x)\neg p$. By assumption $S \vdash \neg(\forall x)\neg p$. So $S \vdash \perp$ by MP, contradiction. Do this for every sentence $(\exists p)$ that is provable from S to get a new language $\overline{L} = L(\Omega \cup C, \Pi)$ and a consistent theory $\overline{S}$ in $\overline{L}$ such that if p is a formula in L with $\text{FV}(p) = \{x\}$ and $S \vdash (\exists x)p$, then there exists a closed term t in $\overline{L}$ such that $\overline{S} \vdash p[t/x]$.

Now start with a consistent theory S in $L = L(\Omega, \Pi)$, we inductively define languages $L_n = (\Omega \cup C_1 \cup \cdots \cup C_n, \Pi)$, each C_k is a new set of constants, and theories

$$S = S_0 \subset S_1 \subset T_1 \subset S_2 \subset T_2 \subset \cdots$$

such that $\forall n \in \mathbb{N}$, S_n is a complete consistent theory in L_{n-1} and T_n is a consistent theory in L_n which has witnesses for S_n. Let $L^* = \bigcup_n L_n$, $S^* = \bigcup_n S_n$.

It's straightforward to check that S^* is a consistent theory in L^* and S^* is complete and has witnesses.

A model for S^* in the language L^* will be a model of S when viewed as a structure in the language L. So without loss of generality, S is consistent in L and has witnesses and is complete.

Let A be the set of equivalence classes of closed terms in L where $s \sim t \iff S \vdash (s = t)$. For $\omega \in \Omega$ with $\alpha(\omega) = n$, define

$$\omega_A : A^n \to A, \omega_A([t_1], \ldots, [t_n]) = [\omega t_1 \ldots t_n].$$

For $\varphi \in \Pi$ with $\alpha(\varphi) = n$, define

$$\varphi_A : A^n \to \{0\}, \varphi_A([t_1], \ldots, [t_n]) = 1 \iff S \vdash \varphi t_1 \ldots t_n.$$

An easy induction shows that for a closed term s, $s_A = [s]$. Next, for a sentence p, $S \vdash p \iff p_A = 1$ (i.e. p holds in A). To prove this, use induction on the language. Then A is a model of S. □

Corollary 6 (Compactness)**.** Let S be a first-order theory. If every finite subset of S has a model, then S has a model.

Proof. If $S \models \perp$, then $S \vdash \perp$. Proofs are finite, so there exists finite $S' \subset S$ such that $S' \vdash \perp$. Hence $S' \models \perp$, contradiction. □

Applications

Can we axiomatise finite groups? In other words, does there exist a theory T whose models are the finite groups?

For $n \in \mathbb{N}$, let

$$t_n = (\exists x_1) \cdots (\exists x_n)(\forall x)(x = x_1 \vee x = x_2 \vee \cdots \vee x = x_n).$$

So t_n means "contains at most n elements". Want

$$T = \text{theory of groups} \cup \{t_1 \vee t_2 \vee t_3 \vee \cdots\}.$$

But $t_1 \vee t_2 \vee t_3 \vee \cdots$ is not a sentence (because it is not finite).

Corollary 7. Finite groups are not axiomatisable as a first-order theory.

Proof. Assume it is, and let T be such a theory. Consider $T' = T \cup \{\neg t_1, \neg t_2, \neg t_3, \dots\}$ where t_n are defined by

$$t_n = (\exists x_1) \cdots (\exists x_n)(\forall x)(x = x_1 \vee x = x_2 \vee \cdots \vee x = x_n).$$

Every finite subset of T' has a model: C_N for some large N (cyclic group of order N). By Corollary 6, T' has a model, but this model must be infinite, hence not a finite group. □

Corollary 8. If a first-order theory T has arbitrarilty large finite models, then it has infinite models.

Proof. Consider

$$T' = T \cup \{(\exists x_1)(\exists x_2)(x_1 \neq x_2), (\exists x_1)(\exists x_2)(\exists x_3)(x_1 \neq x_2 \wedge x_2 \wedge x_3 \wedge x_1 \neq x_2), \dots\}.$$

By assumption, every finite subset of T' has a model, so T' has a model. A model of T' is just an infinite model. □

Start of

lecture 16

Corollary 9 (Upward Löwenheim-Skolem Theorem)**.** Let S be a first-order theory. If S has an infinite model, then S has an uncountable model.

Proof. We introduce an uncountable set of new constants $\{c_i \mid i \in I\}$ to the language. We let

$$S' = S \cup \{\neg c_i = c_j \mid i, j \in I, i \neq j\}.$$

Let A be an infinite model of S. Then A is a model of any finite subset of S'. By Compactness, S' has a model.

A model of S' is a model B of S together with an injection $I \to B$. So B is uncountable. □

Remark. For any set X, can take $I = \gamma(X)$ (from Hartog's Lemma). The proof above shows that S has a model B with an injection $I \to B$. So then there will be no injection $B \to X$.

Corollary 10 (Downward Löwenheim-Skolem Theorem)**.** Let S be a consistent first-order theory in a countable language (Ω, Π are countable). Then if S has a model, then S has a countable model.

Proof. Since S is consistent (by Soundness Theorem), the proof of Theorem 3 builds a countable model (since the language is countable). □

4.1 Peano Arithmetic

We want to axiomatise $\mathbb{N}$ as a first-order theory. Language:

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

with arities $0, 1, 2, 2$. s means "successor", and the others are clear.

Axioms of Peano Arithmetic (PA):

$$(\forall x)(\neg sx = 0)$$
$$(\forall x)(\forall y)(sx = sy \Rightarrow x = y)$$
$$(\forall x)(x \times 0 = 0)$$
$$(\forall x)(\forall y)(x \times (sy) = (x \times y) + x)$$
$$(\forall t_1) \cdots (\forall t_n)[(p[0/x] \wedge (\forall x)(p \Rightarrow p[sx/x])) \Rightarrow (\forall x)p]$$

where the last sentence is for every formula p with $\mathrm{FV}(p) = \{x, t_1, \dots, t_n\}$. This is the axiom-scheme for induction.

Remark. Let p be the formula $x + (y + z) = (x + y) + z$. Then you can prove in PA that $(\forall x)(\forall y)(\forall z)p$ by induction on z with x, y parameters. You prove:

$$(\forall x)(\forall y)(p[0/z] \land (\forall z)(p \Rightarrow p[sz/z]))$$

Note. $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ is a model of PA. We can also interpret $\mathbb{N}$ as a model of PA by taking a bijection with $\mathbb{N}_0$ (but this would be rather unnatural to do).

By Upward Löwenheim-Skolem Theorem, there are uncountable models of PA. Didn't we lean $\mathbb{N}_0$ is uniquely determined by its properties? Yes, but *true* induction says:

$$(\forall A \subset \mathbb{N}_0)((0 \in A \land (\forall x)(x \in A \implies sx \in A)) \implies A = \mathbb{N}_0)$$

In first-order theory, we cannot quantify over subsets of structures. The axiom scheme for induction captures only countably many subsets of $\mathbb{N}_0$.

Definition (Definable set)**.** A subset A of $\mathbb{N}_0$ is *definable* if there's a formula p in language of PA with free variable x such that $p_{\mathbb{N}_0} = A$, i.e.

$$\{a \in \mathbb{N}_0 \mid a \text{ satisfies } p\} = A.$$

Example. Set of primes: use

$$p = (\forall y)((\exists z)(y \cdot z = x) \Rightarrow (y = \underbrace{1}_{=s0} \lor y = x)$$

Powers of 2: use

$$p = (\forall y)(((y \mid x) \land (y \text{ is a prime})) \Rightarrow y = \underbrace{2}_{=ss0})$$

A consequence of Gödel's Incompleteness Theorem: there exists a sentence p such that p holds in $\mathbb{N}_0$, but $\mathrm{PA} \nvdash p$.

5 Set Theory

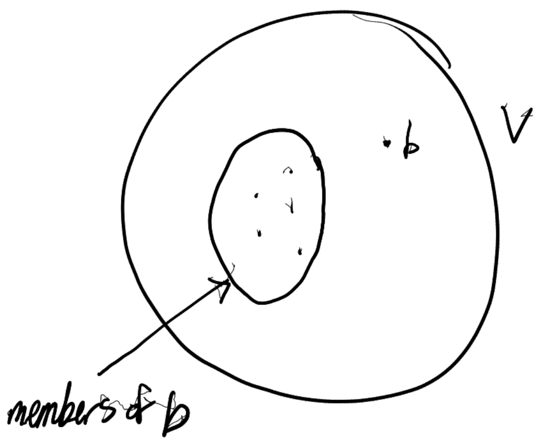
We will describe set theory as just another example of first-order theory. We want to understand what the "universe of sets" looks like.

Zermelo-Frankel Set Theory (ZF)

Language: $\Omega = \emptyset$, $\Pi = \{\in\}$, $\in$ has arity 2.

A structure is a set V together with $[\in]_V \subset V \times V$.

An element of V is called a "set". If $a, b \in V$ and $(a, b) \in [\in]_V$, we say "a belongs to b" or "a is an element of b". V will be the "universe of sets" (when V is a model of ZF).



There will be $2 + 4 + 3$ axioms of ZF.

(1) **Axiom of Extensionality (Ext):** "If two sets have the same members, then they are equal".

$$(\forall x)(\forall y)((\forall z)(z \in x \iff z \in y) \Rightarrow x = y)$$

(2) **Axiom of Separation (Sep):** "We can form subsets of a set."

$$(\forall t_1) \cdots (\forall t_n)[(\forall x)(\exists y)(\forall z)(z \in y \iff (z \in x \wedge p))],$$

where p is any formula with $\mathrm{FV}(p) = \{z, t_1, \ldots, t_n\}$. By (Ext), the set y whose existence is asserted is unique. We denote it by $\{z \in x \mid p\}$. (Formally, we introduce an $(n + 1)$-ary operation symbol to the language; informally, this is an abbreviation).

Example. Given t, x, we can form $\{z \in x \mid t \in z\}$.

Start of lecture 17

(3) **Empty set axiom (Emp):**

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

By (Ext), this set is unique which we denote by $\emptyset$. Formally, we add a constant $\emptyset$ to the language with the sentence $(\forall y)(\neg y \in \emptyset)$.

(4) **Pair set axiom (Pair):** "We can form unordered pairs".

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

Unique by (Ext). We denote this set z by $\{x, y\}$. Define singletons as $\{x, x\}$.

The following an be proved:

$$(\forall x)(\forall y)(\{x, y\} = \{y, x\}).$$

We can use (Pair) to define ordered pairs: for x, y the ordered pair $(x, y) = \{\{x\}, \{x, y\}\}$. One can then prove that:

$$(\forall x)(\forall y)(\forall t)(\forall z)((x, y) = (t, z) \iff (x = t \wedge y = z)).$$

We introduce abbreviations:

- "x is an ordered pair" for $(\exists y)(\exists z)(x = (y, z))$.
- "f is a function" for
$$(\forall x)(x \in f \Rightarrow x \text{ is an ordered pair})$$
$$\wedge(\forall x)(\forall y)(\forall z)(((x, y) \in f \wedge (x, z) \in f) \Rightarrow (y = z))$$
- "$x = \operatorname{dom} f$" for
$$\text{‘}f \text{ is a function'} \wedge (\forall y)(y \in x \iff (\exists z)((yz) \in f))$$
- "f is a function from x to y" for
$$(x = \operatorname{dom} f) \wedge (\forall t)((\exists z)(z, t) \in f \Rightarrow t \in y)$$

(5) **Union axiom (Un):**

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

Denote this set y by $\bigcup x$.

Example. For x, y, $t \in \{x, y\} \iff (t \in x \vee t \in y)$. We also write $\bigcup\{x, y\} = x \cup y$.

Remark. No new axiom eeded for intersection as this can be formed by (Sep). So the following line follows from axioms so far:

$$(\forall x)(\neg x = \emptyset \Rightarrow (\exists y)(\forall z)(z \in y \iff (\forall t)(t \in x \Rightarrow z \in t)).$$

Denote the set y by $\bigcap x$. To prove this, given x, form

$$y = \{z \in \bigcup x : (\forall t)(t \in x \Rightarrow z \in t)\}$$

by (Sep). Check that

$$(\forall z)(z \in y \iff (\forall t)(t \in x \Rightarrow z \in t)).$$

Given x, y, denote $\bigcap\{x, y\}$ by $x \cap y$.

(6) **Power set axiom (Pow):**

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

where $z \subset x$ is an abbreviation for $(\forall t)(t \in z \Rightarrow t \in x)$. We denote y by $\mathbb{P}x$.

We can now form Cartesian product $x \times y$ for sets x, y: an element of $x \times y$ is an ordered pair (s, t) where $s \in x$, $t \in y$. Note that

$$(s, t) = \{\{x\}, \{x, y\}\} \in \mathbb{P}\mathbb{P}(x \cup y),$$

so by (Sep) we can form

$$\{z \in \mathbb{P}\mathbb{P}(x \cup y) : (\exists s)(\exists t)(s \in x \wedge t \in y \wedge z = (s, t)\}.$$

We can also form, from sets x, y

$$y^x = \{f \in \mathbb{P}(x \times y) : (f : x \to y)\},$$

which is the set of all functions from x to y.

(7) **Axiom of infinity (Inf):** From axioms so far, any model V will be infinite, for example

$$\emptyset, \mathbb{P}\emptyset, \mathbb{P}\mathbb{P}\emptyset, \ldots$$

are all distincnt elements of V.

For a set x define the successor of x as $x^+ = x \cup \{x\}$. Then

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

are distinct elements of V:

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

We write $0 = \emptyset$, $1 = \emptyset^+$, $2 = \emptyset^{++}$, ... We have a copy of $\mathbb{N}_0$ in V. From the outside, V is infinite. From the inside, V is not a set: $\neg(\exists x)(\forall y)(y \in x)$ (Russell's paradox).

Abbreviate "x is a successor set":

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

Axiom (Inf) says:

$$\boxed{(\exists x)(x \text{ is a succcessor set})}.$$

The intersection of successor sets is a successor set. So we can construct "smallest" successor set, i.e. we can prove

$$(\exists x)((x \text{ is a successor set}) \wedge (\forall y)(y \text{ is a successor set} \Rightarrow x \subset y))$$

(Pick any sucessor set z, let $x = \bigcap\{y \in \mathbb{P}z \mid y \text{ is a sucessor set}\}$. x is then a successor set, and if y is any successor set then $x \subset (y \cap z)$.) We denote the smallest successor set by ω.

If $x \subset \omega$ is a successor set then $x = \omega$, i.e.

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

This is true induction.

We can prove by induction:

- $(\forall x)(x \in \omega \Rightarrow \neg x^+ = \emptyset)$
- $(\forall x)(\forall y)(((x \in \omega) \wedge (y \in \omega) \wedge (x^+ = y^+) \Rightarrow x = y))$.

We can define abbreviations:

- "x is finite" for $(\exists y)((y \in \omega) \wedge (\exists f)(f : x \to y \wedge f \text{ is a bijection}))$.
- "x is countable" for $(\exists f)(f : x \to \omega \wedge f \text{ is injective})$

Start of

lecture 18

(8) **Axiom of Replacement (Rep):** (Inf) says that there exist sets containing $0, 1, 2, 3, \ldots$. Are there sets containing $\emptyset, \mathbb{P}\emptyset, \mathbb{P}\mathbb{P}\emptyset, \ldots$? There's a function-like object that sends $0 \mapsto \emptyset$, $1 \mapsto \mathbb{P}\emptyset$, $2 \mapsto \mathbb{P}\mathbb{P}\emptyset, \ldots$. Need an axiom that says that the image of a set under a function-like object is a set. The axiom is:

$$(\forall t_1) \cdots (\forall t_n) \Big[(\forall x)(\forall y)(\forall z)((p \wedge p[z/y]) \Rightarrow y = z)$$

$$\Rightarrow (\forall x)(\exists y)(\forall z)(z \in y \iff (\exists u)(u \in x \wedge p[u/x, z/y])) \Big]$$

for any formula p with $\mathrm{FV}(p) = \{x, y, t_1, \ldots, t_n\}$.

We will explain the reasoning below, by discussing function-classes.

Digression on classes

Definition (Class)**.** A *class* is a subset C of a structure V of the language of ZF such that there is a formula p with $\mathrm{FV}(p) = \{x\}$ such that $p_V = C$, i.e. $x \in C$ if and only if $p(x)$ holds in V.

Example. V is a class: for example take p to be $x = x$. The set of sets of size 1 is a class: for example, take p to be $(\exists y)(x = \{y\})$.

Definition (Proper class)**.** Say the class is a set if $(\exists y)(\forall x)(x \in y \iff p)$ holds in V. If C is not a set, we say C is a *proper class.*

Example. V is a proper class (Russell's paradox).

Definition (Function class)**.** A *function-class* is a subset G of $V \times V$ such that there's a formula p with free variables $\mathrm{FV}(p) = \{x, y\}$ such that

$$(\forall x)(\forall y)(\forall z)((p \wedge p[z/y]) \Rightarrow y = z)$$

holds in V and $G = p_V$, i.e. $(x, y) \in G$ if and only if $p(x, y)$ holds in V.

Example. $G = \{(x, \{x\}) \mid x \in V\}$ is the function-class mapping $x \mapsto \{x\}$, and is given by $p = (y = \{x\})$.

(9) **Axiom of Foundation (Fnd):** We want to avoid pathological behaviour like $x \in x$, i.e. $\{x\}$ has no $\in$-minimal member, or $x \in y \wedge y \in x$ (in which case $\{x, y\}$ has no $\in$-minimal member). (Fnd) says that every non-empty set has an $\in$-minimal member:

$$(\forall x)(\neg x = \emptyset \Rightarrow (\exists y)(y \in x \wedge (\forall z)(z \in x \Rightarrow \neg z \in y)))$$

The above axioms and axiom-schemes (1)-(9) form ZF. The axiom of choice (AC) is not included:

$$(\forall x)((\forall y)(y \in x \Rightarrow \neg y = \emptyset) \Rightarrow (\exists f)((f : X \to \bigcup x) \wedge (\forall y)(y \in x \Rightarrow f(y) \in y))).$$

We write ZFC for ZF + AC. For the rest of this chapter we work within ZF.

Aim: to describe the set-theoretic universe, i.e. any model V of ZF.

Definition (Transitive set)**.** We say a set x is *transitive* if every membet of x is a member of x. So "x is transitive" is shorthand for

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

Equivalently, $\bigcup x \subset x$.

Note. This is *not* the same as saying that $\in$ is a transitive relation on x.

Example. ω is transitive. We need to show that $x \subset \omega$ for all $x \in \omega$. Form the set $z = \{y \in \omega \mid y \subset \omega\}$. Check z is a successor set, so $z = \omega$. Similarly,

$$\{x \in \omega \mid \text{“}x \text{ is transitive”}\}$$

is a successor set ($\bigcup x^+ = x$) so it is ω. So every element of ω is a transitive set.

Lemma 1. Every set x is contained in a transitive set, i.e.

$$(\forall x)(\exists y)(\text{“}y \text{ is transitive”} \wedge x \subset y).$$

Remark. The intersection of transitive sets if transitive, so x is contained in a smallest transitive set, called the *transitive closure* of x, denoted by $\mathrm{TC}(x)$.

Idea: If $x \subset y$, y transitive, then $\bigcup x \subset y$ and so $\bigcup\bigcup x \subset y$, $\bigcup\bigcup\bigcup x \subset y$, ... Want to form

$$\bigcup \left\{x, \bigcup x, \bigcup\bigcup x, \dots\right\}$$

Is this a set? Yes, by (Rep). We need a function-class $0 \mapsto x$, $1 \mapsto \bigcup x$, $2 \mapsto \bigcup\bigcup x$, ...

Proof. Say "f is an attempt" to mean:

$$\text{“}f \text{ is a function”} \wedge \text{“}\mathrm{dom}\, f \in \omega\text{”} \wedge \text{“}f(0) = x\text{”}$$
$$\wedge (\forall m)(\forall n)[((m \in \mathrm{dom}\, f) \wedge (n \in \mathrm{dom}\, f) \wedge (n = m^+)) \Rightarrow (f(n) = \bigcup f(m))]$$

We prove by ω-induction that:

$$(\forall f)(\forall g)(\forall n)((\text{“}f \text{ is an attempt”} \wedge \text{“}g \text{ is an attempt”} \wedge (n \in \mathrm{dom}\, f \cap \mathrm{dom}\, g)) \Rightarrow (f(n) = g(n))) \tag{$*$}$$

and

$$(\forall n)(n \in \omega \Rightarrow (\exists f)(\text{“}f \text{ is an attempt”} \wedge n \in \mathrm{dom}\, f)) \tag{$**$}$$

Define a function-class via the formula $p(y, z)$:

$$(\exists f)(\text{“}f \text{ is an attempt”} \wedge f(y) = z).$$

By $(*)$ we do have

$$(\forall y)(\forall z)(\forall w)((p \wedge \subset p[w/z]) \Rightarrow w = z).$$

By (Rep) can form $w = \{z \mid (\exists y)(y \in \omega \wedge p(y, z))\}$ $(w = \{x, \bigcup x, \bigcup\bigcup x, \dots\})$ and by (Un) can form $t = \bigcup w$. Then $x \subset t$, since $x \in w$ ($\{(0, x)\}$ is an attempt). Given $a \in t$, we have $z \in w$, $a \in z$. There's an attempt f and $n \in w$ such that $z = f(n)$. By $(**)$, there's an attempt g with $n^+ \in \mathrm{dom}\, g$. Then $n \in \mathrm{dom}\, g$, so

$$\bigcup z = \bigcup f(n) \overset{(*)}{=} \bigcup g(n) = g(n^+) \in w$$

hence $a \subset t$. □

Start of lecture 19

Theorem 2 (Principle of $\in$-induction)**.** For any formula p with $\mathrm{FV}(p) = \{x, t_1, \dots, t_n\}$, we have

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

Proof. Fix $t_1, \dots, t_n$ and assume

$$(\forall x)(((\forall y \in x)p(y)) \Rightarrow p(x))$$

holds. We want to show that $(\forall x)p(x)$ holds. Assume not, so $\neg p(x)$ holds for some x. We'd like to pick an $\in$-minimal member of $\{y \mid \neg p(y)\}$, but this is not a set. Choose a transitive set t such that $x \in t$. For example can pick $t = \mathrm{TC}(\{x\})$. By (Sep) we can form the set $u = \{y \in t \mid \neg p(y)\}$. Note that $x \in u$ so $u \neq \emptyset$. Let z be an $\in$-minimal membet of u (exists by (Fnd)). If $y \in z$, then $y \in t$ (t is transitive) and $y \notin u$ (by minimality), so $p(y)$ holds. By assumption $p(z)$ holds, which contradicts $z \in u$. □

Remark. In the presence of axioms (1) - (8) of ZF, (Fnd) is equivalent to the principle of $\in$-induction.

Proof. Assume $\in$-induction (as well as axioms (1) - (8)). We deduce (Fnd). Clever idea: say "x is regular" to mean

$$(\forall y)(x \in y \Rightarrow \text{“}y \text{ has } \in\text{-minimal member”})$$

We prove by $\in$-induction that $(\forall x)$("x is regular"). This obviously implies (Fnd). Fix a set x and assume that y is regular for all $y \in x$. We want to deduce that x is regular.

Let z be a set such that $x \in z$. Then:

- either x is an $\in$-minimal membet of z
- or there's $y \in z$ such that $y \in x$. By induction hypothesis, y is regular, so z has an ε-minimal member.

□

Next step: $\in$-recursion. Want to define functions such that $f(x)$ depends on $f(y)$, $y \in x$, i.e. $f(x)$ depends on $f|_x$.

Theorem 3 ($\in$-recursion theorem)**.** For any function-class G (given by a formula p with two free variables such that $(x, y) \in G \iff p(x, y)$ holds) which is defined everywhere (so $(\forall x)(\exists y)p(x, y)$), then there is a function-class F (given by some formula q) defined everywhere such that

$$(\forall x)(F(x) = G(F|_x)).$$

Moreover, F is unique.

Note. $F|_x$ is a set by (Rep): $F|_x = \{(s,t) \mid s \in x, t = F(s)\}$ is the image of the set x under the function-class $s \mapsto (s, F(s))$.

Proof. **Uniqueness:** Assume F_1, F_2 both satisfy the theorem. Then we prove $(\forall x)(F_1(x) = F_2(x))$ by $\in$-induction. If $F_1(y) = F_2(y)\ \forall y \in x$, then $F_1|_x = F_2|_x$, so $F_1(x) = F_2(x)$.

Existence: Say "f is an attempt" to mean

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

Note that $f|_x$ makes sense as $\operatorname{dom} f$ is transitive. We prove by $\in$-induction that

$$(\forall f)(\forall g)(\forall x)((\text{“}f \text{ is an attempt”} \wedge \text{“}g \text{ is an attempt”} \wedge (x \in \operatorname{dom} f \cap \operatorname{dom} g)) \Rightarrow (f(x) = g(x)))$$

Call this property $(*)$. Then we show by $\in$-induction that

$$(\forall x)(\exists f)(\text{“}f \text{ is an attempt”} \wedge (x \in \operatorname{dom} f)).$$

Call this property $(**)$. Fix x. Assume every $y \in x$ is in the domain of some attempt, which is then defined on $\mathrm{TC}(\{y\})$ and is unique by $(*)$ – call this f_y. Then

$$f' = \bigcup \{f_y \mid y \in x\}$$

is an attempt by $(*)$, and is a set by (Rep). Finally $f = f' \cup \{(x, G(f'))\}$ is an attempt defined at x. Note that $f|_x = f'$. Let q be the formula:

$$(\exists f)(\text{“}f \text{ is an attempt”} \wedge (y = f(x))).$$

Then q defines the required function-class F. □

We can generalise induction and recursion to other relations. Let r be a relation (i.e. a formula with two free variables).

Definition (Well-founded)**.** We say a relation r is *well-founded* if

$$(\forall x)((\neg x = \emptyset) \Rightarrow (\exists y \in x)((\forall z \in x)(\neg zry)))$$

(i.e. every non-emptyer set has an r-minimal member).

Example. If r is $(x \in y)$ is the $\in$-relation, then r is well-founded by (Fnd).

Definition (Local)**.** We say a relation r is *local* if

$$(\forall x)(\exists y)(\forall z)(z \in y \iff zrx).$$

(i.e. the r-predecessors of x form a set).

Example. $\in$ is local: the $\in$ predecessors of x is precisely the set x.

"Local" is needed for r-closure. Then we can prove r-induction and r-recursion.

Can restrict r to a class or a set. Note that if r is a relation on a set a, then for any $x \in a$, $\{y \in a \mid yrx\}$ is a set by (Sep). So we only need well-foundedness to have r-induction and r-recursion on a.

Is this really more general than $\in$? No, provided we also assume that r is *extensional* on a:

Definition (Extensional)**.** We say a relation r is *extensional* if:

$$(\forall x, y \in a)((\forall z \in a)((zrx) \iff (zry)) \Rightarrow x = y).$$

Start of lecture 20

Theorem 4 (Mostowki's Collapsing Theorem)**.** Let r be a well-founded, extensional relation on a set a. Then there is a transitive b and a bijection $f : a \to b$ such that

$$(\forall x, y \in a)(xry \iff f(x) \in f(y)).$$

Moreover, (b, f) is unique.

Proof. By r-recursion on a, there's a function-class such that

$$\forall x \in a \quad f(x) = \{f(y) \mid y \in a \wedge yrx\}.$$

Note that f is a function, not just a function-class, since $\{(x, f(x)) \mid x \in a\}$ is a set by (Rep). Then

$$b = \{f(x) \mid x \in a\}$$

is a set by (Rep). Now we check:

- b is transitive: let $z \in b$ and $w \in z$. There's a $x \in a$ such that $z = f(x)$, and so there's $y \in a$ such that yrx and $w = f(y) \in b$.

- f is surjective (true by definition of b).

- $\forall x, y \in a$, $xry \Rightarrow f(x) \in f(y)$ is true by definition of f.

- It remains to show that f is injective. It will then follow that $\forall x, y \in a$, $f(x) \in f(y) \Rightarrow xry$. Indeed, if $f(x) \in f(y)$, then $f(x) = f(z)$ for some $z \in a$ with zry. Since f is injective, $x = z$, so xry. We will show

  $$(\forall x \in a)\underbrace{(\forall y \in a)(f(x) = f(y) \Rightarrow x = y)}_{\text{“}f\text{ is injective at }x\text{”}}$$

  by r-induction. Fix $x \in a$ and assume that f is injective at s whenever $s \in a$ and srx. Assume $f(x) = f(y)$ for some $y \in a$, i.e.

  $$\{f(s) \mid s \in \wedge srx\} = \{f(t) \mid t \in a \wedge try\}.$$

  Since f is injective at every $s \in a$ with srx, it follows that

  $$\{s \in a \mid srx\} = \{t \in a \mid try\}.$$

  By extensionality for r, it follows that $x = y$.

Now we check that (b, f) is unique. Assume that (b, f) and (b', f') both satisfy the theorem. We prove

$$(\forall x \in a)(f(x) = f'(x))$$

by r-induction. Fix $x \in a$ and assume $f(y) = f'(y)$ whenever $y \in a$ and yrx. If $z \in f(x)$, then $z \in b$ (b transitive), so $z = f(y)$ for some $y \in a$ with yrx. Then $z = f(y) = f'(y)$ (induction hypothesis). Then $z = f'(y) \in f'(x)$. Similarly, if $z \in f'(x)$ thne $z \in f(x)$. By (Ext), $f(x) = f'(x)$. □

Definition (Ordinal (set theoretic))**.** An *ordinal* is a transitive which is well-ordered by $\in$ (equivalently, linearly ordered since $\in$ is well-founded by (Fnd)).

Note. Let a be a set and r be a well-ordering on a. Then r is well-founded and extensional (if $x, y \in a$ and $\neg x = y$ then xry or ryx, but not both).

By Mostowki's Collapsing Theorem, there exists a transitive b and a bijection $f : a \to b$ such that $xry \iff f(x) \in f(y)$, i.e. $f(a, r) \to (b, \in)$ is an order-isomorphism. So b is an ordinal.

So by Mostowki's Collapsing Theorem, every well-ordered set is order-isomorphic to a unique ordinal, called the order-type of x.

We let ON denote the class of ordinals (given by the formula “x is an ordinal”). It is a proper class by Burati-Forti paradox.

Proposition 5. Let $\alpha, \beta \in \mathrm{ON}$, and let a be a set of ordinals. Then:

(i) Every member of α is an ordinal.

(ii) $\beta \in \alpha \iff \beta < \alpha$ (β is order-isomorphic to a proper initial segment of α)

(iii) $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$

(iv) $\alpha^+ = \alpha \cup \{\alpha\}$ (i.e. the set theoretic meaning and ordinal meanings for $^+$ agree).

(v) $\bigcup a$ is an ordinal and $\bigcup a = \sup a$.

Remark. (ii) says that α really *is* the set of ordinals $< \alpha$. (iii) says that $\in$ linearly orders the class ON. (iv) resolves the clash of notation x^+ in Section 2 and Section 5. (v) now shows that any set of well-ordered sets has an upper bound.

Proof.

(i) Let $\gamma \in \alpha$. Then $\gamma \subset \alpha$ (since α is transitive) and hence $\in$ linearly orders γ. Given $\eta \in \delta, \delta \in \gamma$, then $\delta \in \alpha$ and so $\eta \in \alpha$ (α transitive). Since $\in$ is transitive on α, we have $\eta \in \gamma$. So γ is a transitive, so γ is an ordinal.

(ii) If $\beta \in \alpha$, then $I_\beta = \{\gamma \in \alpha \mid \gamma \in \beta\} = \beta$, so $\beta < \alpha$. Any proper initial segment of α is of the form I_γ for some $\gamma \in \alpha$. So $\beta < \alpha \implies \beta \in \alpha$.

(iii) We know $\beta < \alpha$ or $\beta = \alpha$ or $\alpha < \beta$ is true. Then done by (ii).

(iv) Let $\beta = \alpha \cup \{\alpha\}$ (successor of α). If $\gamma \in \beta$ then either $\gamma = \alpha \subset \beta$ or $\gamma \in \alpha$, so $\gamma \subset \alpha \subset \beta$. Thus β is transitive, linearly ordered by $\in$ (by (iii)) and α is the greatest element. So $\beta = \alpha^+$ in the sense of Section 2.

(v) $\bigcup a$ is a union of transitive sets, hence transitive. Every member of $\bigcup a$ is an ordinal, so $\bigcup a$ is linearly ordered by $\in$ by (iii). If $\gamma \in a$, then $\gamma \subset \bigcup a$, so either $\gamma = \bigcup a$, or $\gamma \in \bigcup a$ (by (ii)), i.e. $\gamma \leq \bigcup \alpha$. If $\gamma \leq \delta$ for all $\gamma \in \alpha$, then $\gamma = \delta$ or $\gamma \in \delta$ for $\gamma \in a$, i.e. $\gamma \subset \delta$ (using (ii)). So $\bigcup a \subset \delta$, i.e. $\bigcup a \leq \delta$. $\square$

Example. $0 = \emptyset \in \mathrm{ON}$, hence $n \in \mathrm{ON}$ for all $n \in \omega$ (by ω-induction). ω is transitive, so $\bigcup \omega \subset \omega$. If $n \in \omega$, then $n \in n^+ \in \omega$, so $n \in \bigcup \omega$. So $\omega = \bigcup \omega$ is an ordinal and $\omega = \sup \omega$.

Start of

lecture 21

5.1 Picture of the Universe

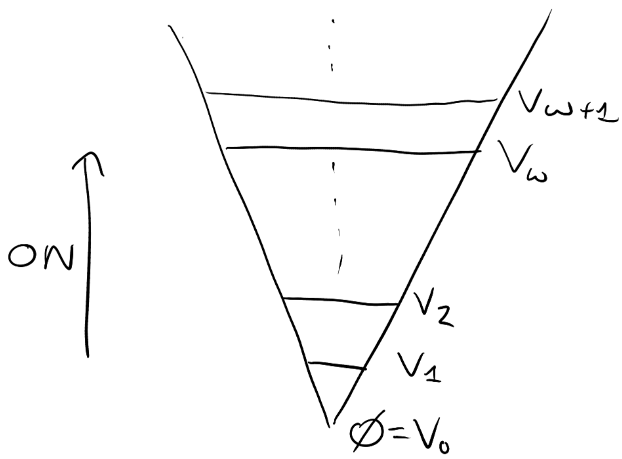
Idea: everything is built up from $\emptyset$ using $\mathbb{P}$ and $\cup$. Have

$$V_0 = \emptyset, V_1 = \mathbb{P}\emptyset = \{\emptyset\}, V_2 = \mathbb{P}\mathbb{P}\emptyset = \{\emptyset, \{\emptyset\}\}, \dots$$

and then will have

$$V_\omega = \bigcup\{V_0, V_1, V_2, \dots\}, V_{\omega+1} = \mathbb{P}V_\omega, \text{etc}$$

It will be (Fnd) that guarantees that every set appears in a V_α.



We define sets V_α, $\alpha \in \mathrm{ON}$ by $\in$-recursion:

- $\alpha = 0$: $V_0 = \emptyset$
- $\alpha = \beta^+$: $V_\alpha = \mathbb{P}V_\beta$
- $\alpha \neq 0$ limit: $V_\alpha = \bigcup\{V_\gamma \mid \gamma < \alpha\}$

The sets V_α form the *von Neumann hierarchy*.

Aim: Every set appears in this hierarchy.

Lemma 6. V_α is transitive for all $\alpha \in \mathrm{ON}$.

Proof. By induction on α.

- $\alpha = 0$: $V_0 = \emptyset$ is transitive.
- $\alpha = \beta^+$: Let $x \in V_\alpha = \mathbb{P}V_\beta$. Then $x \subset V_\beta$. If $y \in x$, then $y \in V_\beta$, so by induction hypothesis, $y \subset V_\beta$ (V_β is transitive). So every $y \in x$ has $y \in \mathbb{P}V_\beta = V_\alpha$. Thus V_α is transitive.
- $\alpha \neq 0$ limit: If $x \in V_\alpha$, then $\exists \gamma < \alpha$, $x \in V_\gamma$. By induction, V_γ is transitive, so $x \subset V_\gamma \subset V_\alpha$. So V_α is transitive □

Lemma 7. If $\alpha \leq \beta$, then $V_\alpha \subset V_\beta$.

Proof. By induction on β.

- $\beta = 0$: $\alpha \leq \beta$, so $\alpha = 0$, so $V_\alpha = V_\beta$.
- $\beta = \gamma^+$: If $\alpha = \beta$ then $V_\alpha = V_\beta$. If $\alpha < \beta$, then $\alpha \leq \gamma$, so by induction hypothesis, $V_\alpha \subset V_\gamma$. If $x \in V_\gamma$, then $x \subset V_\gamma$ (V_γ is transitive), so $x \in \mathbb{P}V_\gamma \subset V_\beta$. Thus $V_\gamma \subset V_{\gamma^+} = V_\beta$, and hence $V_\alpha \subset V_\beta$.
- If $\beta \neq 0$ limit: then if $\alpha < \beta$ then $V_\alpha \subset V_\beta$ by definition. □

Theorem 8. The von Neumann hierarchy exhausts the set theoretic universe V, i.e.

$$(\forall x)(\exists \alpha \in \mathrm{ON})(x \in V_\alpha)$$

or

$$V = \bigcup_{\alpha \in \mathrm{ON}} V_\alpha.$$

Note. If $x \in V_\alpha$ then $x \subset V_\alpha$ (by Lemma 6). If $x \subset V_\alpha$ then $x \in \mathbb{P}V_\alpha = V_{\alpha+1}$.

If $\exists \alpha \in \mathrm{ON}$, $x \subset V_\alpha$ then define the *rank of* x to be $\mathrm{rank}(x)$, the least $\alpha \in \mathrm{ON}$ such that $x \subset V_\alpha$.

Proof. We will show $(\forall x)(\exists \alpha \in \mathrm{ON})(x \subset V_\alpha)$ by $\in$-induction. Fix x and assume for each $y \in x$, $y \subset V_\alpha$ for some $\alpha \in \mathrm{ON}$, so for all $y \in x$, $y \subset V_{\mathrm{rank}(y)}$. Let

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

which is a set by (Rep). We'll show $x \subset V_\alpha$. If $y \in x$, then $y \subset V_{\text{rank}(y)}$, so $y \in \mathbb{P}V_{\text{rank}(y)} = V_{\text{rank}(y)^+} \subset V_\alpha$ (where the final $\subset$ is using Lemma 7). This shows $x \subset V_\alpha$. $\square$

Corollary 9. For every set x,

$$\text{rank}(x) = \sup\{\text{rank}(y)^+ \mid y \in x\}$$

Proof.

$\leq$: Follows from proof of Theorem 8.

$\geq$: We first show that $x \in V_\alpha \implies \text{rank}(x) < \alpha$.

- $\alpha = 0$ is true.
- $\alpha = \beta^+$: $x \in \mathbb{P}V_\beta$, so $x \subset V_\beta$, so $\text{rank}(x) \leq \beta < \alpha$
- $\alpha \neq 0$ limit: $x \in V_\alpha \implies \exists \gamma < \alpha$ with $x \in V_\gamma$, so $\text{rank}(x) < \gamma < \alpha$.

Now let $\alpha = \text{rank}(x)$. Then $x \subset V_\alpha$, so for $y \in x$, $y \in V_\alpha$ and so $\text{rank}(y) < \alpha$. Hence

$$\sup\{\text{rank}(y)^+ \mid y \in x\} \leq \alpha.$$ $\square$

Example. $\text{rank}(\alpha) = \alpha$ for all $\alpha \in \text{ON}$. By induction:

$$\begin{aligned} \text{rank}(\alpha) &= \sup\{\text{rank}(\beta)^+ \mid \beta < \alpha\} \\ &= \sup\{\beta^+ \mid \beta < \alpha\} && \text{(induction hypothesis)} \\ &= \alpha \end{aligned}$$

6 Cardinal Arithmetic

Look at the size of sets. We write $x \cong y$ to mean

$$(\exists f)(f : x \to y \land \text{“}f \text{ is a bijection”}).$$

This is an equivalence relation class. The equivalence classes are proper classes (except $\{\emptyset\}$).

How do we pick a representative from each equivalence class? We seek for each set x, a set $\operatorname{card} x$ such that

$$(\forall x)(\forall y)(\operatorname{card} x = \operatorname{card} y \iff x \cong y)$$

In ZFC this is easy: given a set x, x can be well-ordered, so $x \cong \mathrm{OT}(x)$, i.e. $x \cong \alpha$ for some $\alpha \in \mathrm{ON}$. Can define $\operatorname{card} x$ to be the least $\alpha \in \mathrm{ON}$ such that $x \cong \alpha$.

In ZF (due to D. S. Scott): define the *essential rank* as follows:

$$\operatorname{ess rank}(x) = \text{least } \alpha \text{ such that } \exists y \subset V_\alpha \text{ with } y \cong x.$$

Note $\operatorname{ess rank}(x) \le \operatorname{rank}(x)$. Define

$$\operatorname{card} x = \{y \subset V_{\operatorname{ess rank}(x)} \mid y \cong x\}.$$

Start of

lecture 22

TODO

Start of

lecture 23

In ZFC:

Definition (Cardinal sum and product)**.** Given a set I and cardinals m_i, $i \in I$, we define

$$\sum_{i \in I} m_i = \operatorname{card}\left(\bigsqcup_{i \in I} M_i\right)$$

(here M_i is a set of cardinalty m_i, $i \in I$, $\bigsqcup_{i \in I} M_i = \bigcup_{i \in I} M_i \times \{i\}$). We also define

$$\prod_{i \in I} m_i = \operatorname{card}\left(\prod_{i \in I} M_i\right)$$

($\prod_{i \in I} M_i = \{f : I \to \bigcup_{i \in I} M_i \mid f(i) \in M_i \ \forall i \in I\}$).

Need axiom of choice as we need to be able to choose M_i for each $i \in I$ and to prove these operations are well-defined, given $M_i \equiv M_i'$, $i \in I$, we need to choose for each $i \in I$, a bijection $f_i : M_i \to M_i'$, and show $\bigcup_i M_i \equiv \bigsqcup_i M_i'$, $\prod_i M_u \equiv \prod_i M_i'$.

Example. If $\operatorname{card} I \le \aleph_\alpha$, $m_i \le \aleph_\aleph$ for all $I \in I$, then $\sum_{i \in I} m_i \le \aleph_\alpha$.

Note. If $n = \operatorname{card} I$ and $m_i = m \; \forall i \in I$, $\prod_{i \in I} m_i = m^n$.

If $\alpha \le \beta$, then

$$2^{\aleph_\beta} \le \aleph_\alpha^{\aleph_\beta} = 2^{\aleph_\alpha \aleph_\beta} \le 2^{\aleph_\beta \aleph_\beta} = 2^{\aleph_\beta}.$$

So we've reduced to studying $2^{\aleph_\beta}$. Hard. $\aleph_\alpha < 2^{\aleph_\alpha}$, so $\aleph_0 < 2^{\aleph_0} = \operatorname{card}(\mathbb{R})$.

Continuum Hypothesis (CH): $2^{\aleph_0} = \aleph_1$.

Paul Cohen proved: if ZFC is consistent, then so are ZFC + Continuum Hypothesis and ZFC + $\neg$ Continuum Hypothesis.

THIS IS THE END OF ALL THE EXAMINABLE MATERIAL FOR THIS COURSE (THE NEXT SECTION IS COMPLETELY NON EXAMINABLE).

7 Classical Descriptive Set Theory (Non-examinable)

Study of "definable sets" in Polish spaces. Borel hierarchy, projective hierarchy.

Aim: Continuum Hypothesis holds for analytic sets.

We show that the analogous statement to $P \neq NP$ holds in this setting.

Definition (Polish space)**.** A topological space X is a *Polish space* if it is separable and complete metrizable.

Example. Baire space $\mathcal{N} = \mathbb{N}^{\mathbb{N}}$. Basic open sets are:

$$U_{m_1,\dots,m_k} = \{\mathbf{n} = (n_i)_{i=1}^{\infty} \in \mathcal{N} \mid n_i = m_i, 1 \le i \le k\}.$$

$d(\mathbf{m}, \mathbf{n}) = \sum_{k, m_k \neq n_k} 2^{-k}$.

$\{0,1\}^{\mathbb{N}} \subset \mathcal{N}$.

Lemma 1. Any Polish space is a continuous image of $\mathcal{N}$.

Proof. Let X be a Polish space with complete metric d. Let $X = \bigcup_{n\in\mathbb{N}} U_n$, U_n non-empty and open, $\operatorname{diam}(U_n) < 1$ (since X separable). Let $U_n = \bigcup_{p\in\mathbb{N}} U_{n,p}$, $U_{n,p}$ non-empty and open, $\operatorname{diam}(U_{n,p}) < \frac{1}{2}$.

Continue infinitely, by letting

$$U_{n_1,\dots,n_k} = \bigcup_{n_{k+1}\in\mathbb{N}} U_{n_1,\dots,n_{k+1}}$$

with $U_{n_1,\dots,n_{k+1}}$ always non-empty and open, and diam $< \frac{1}{k+1}$.

Now pick $x_{n_1,\dots,n_k} \in U_{n_1,\dots,n_k}$. Define

$$\phi : \mathcal{N} \to X$$
$$\varphi(\mathbf{n}) = \lim_{k\to\infty} x_{n_1,\dots,n_k}$$

□

Lemma 2. $\mathcal{N}$ is homeomorphic to the set of irrationals on $[0,1]$.

Proof. Continued fractions (for a definition and some properties, see Number Theory). □

Definition (Borel hierarchy)**.** Let X be a set. A *σ-field* on X is a subset $\mathcal{F} \subset \mathbb{P}X$ such that

(i) $\emptyset \in \mathcal{F}$

(ii) $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$

(iii) $A \in \mathcal{F} \implies X \setminus A \in \mathcal{F}$

If X is a Polish space, then the *Borel σ-field* $\mathcal{B}$ on X is the smallest σ-field on X that contains the open sets.

Remark. This is a field under the operations of symmetric difference and intersection, with identity $\emptyset$ (alternatively, it is also a field under the operations of symmetric difference and union, with identity X).

Definition (Σ_1^0, Π_1^0)**.** We let Σ_1^0 be the set of open subsets of X, and π_1^0 be the set of closed subsets of X. We define Σ_α^0, Π_α^1 for $1 \le \alpha < \omega_1$ b recursion:

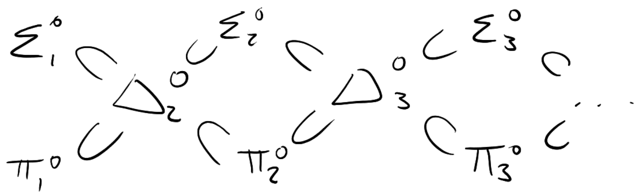
- $\Sigma_{\alpha+1}^0$ is the countable unions of members of Π_α^0 (for example, Σ_2^0 are the F_σ-sets).
- $\Pi_{\alpha+1}^0$ are the complemenets of membets of $\Sigma_{\alpha+1}^0$ (for example, Π_2^0 are the G_δ-sets).

For $\alpha \neq 0$ limit:

- Σ_α^0 consists of sets of the form $\bigcup_{n \in \mathbb{N}} A_n$, where $\forall n < \omega$, $\exists \beta < \alpha$ with $A_n \in \Pi_\beta^0$.
- Π_α^0 is the complements of members of Σ_α^0.

Definition (Δ_α^0)**.** We define $\Delta_\alpha^0 = \Sigma_\alpha^0 \cap \Pi_\alpha^0$.

We have:



Prove the $\subset$ property by induction, starting with $\Sigma_1^0 \subset \Sigma_2^0$, using the fact that we are a metric space (not just a topological space).

Proposition 3. $\bigcup_{\alpha<\omega_1} \Sigma_\alpha^0 = \bigcup_{\alpha<\omega_1} \Pi_\alpha^0 = \mathcal{B}$ (the set of Borel sets).

Proof. First notice:

$$\bigcup_{\alpha<\omega_1} \Sigma_\alpha^0 = \bigcup_{\alpha<\omega_1} \Pi_\alpha^0 \subset \mathcal{B}.$$

Need: $\mathcal{F} = \bigcup_{\alpha<\omega_1} \Sigma_\alpha^0$ is a σ-field. For example, if $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, then $A_n \in \Pi_{\alpha_n}^0$ for some $\alpha_n < \omega_1$. Let $\alpha = \sup(\alpha_n + 1)$. Then $\bigcup_n A_n \in \Sigma_\alpha^0$ etc. □

Definition (Universal subset)**.** A subset $A \subset \mathcal{N} \times \mathcal{N}$ is a *universal* Σ_α^0*-set* if:

(i) A is Σ_α^0

(ii) If $B \subset \mathcal{N}$ is Σ_α^0 then $\exists \mathbf{m} \in \mathcal{N}$, $B = \{\mathbf{n} \in \mathcal{N} \mid (\mathbf{m}, \mathbf{n}) \in A\}$.

Theorem 4. $\forall \alpha$, $1 \le \alpha < \omega_1$, there exists a universal Σ_α^0 set.

Proof.

$\alpha = 1$: Can enumerate the basic open set of $\mathcal{N}$ as $U_1, U_2, U_3, \ldots$. If $B \subset \mathcal{N}$ is open, then $B = \bigcup_{i \in \mathbb{N}} U_{m_i}$ for some $\mathbf{m} = (m_i) \in \mathcal{N}$. So $\mathbf{n} \in B \iff \exists i\ \mathbf{n} \in U_{m_i}$. So define

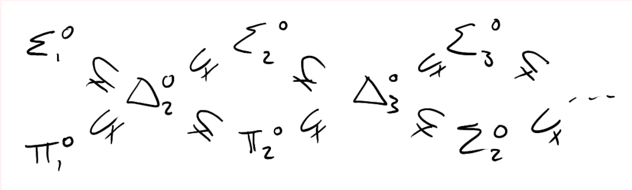
$$A = \{(\mathbf{m}, \mathbf{n}) \in \mathcal{N} \times \mathcal{N} \mid \exists i\ \mathbf{n} \in U_{m_i}\}.$$

This is open and universal by above.

$\alpha > 1$: use induction. □

Corollary 5. For every α, $1 \leq \alpha < \omega_1$, there exiss a set $A \in \Sigma^0_\alpha \setminus \Pi^0_\alpha$.

Note. It follows that



Proof. Let $A \subset \mathcal{N} \times \mathcal{N}$ be a universal Σ^0_α set.

$$B = \{\mathbf{n} \in \mathcal{N} \mid (\mathbf{n}, \mathbf{n}) \in A\}.$$

B is Σ^0_α ($\mathbf{n} \mapsto (\mathbf{n}, \mathbf{n})$ is continuous). If B is Π^0_α then $\exists \mathbf{m}$ with $B = \{\mathbf{n} \mid (\mathbf{n}, \mathbf{n}) \notin A\}$. $\mathbf{m} \in B$? contradiction. □

Start of lecture 24

Projective Hierarchy

Definition (Analytic st)**.** An *analytic set* (in a Polish space) is the continuous image of $\mathcal{N}$.

Example. Every Polish space (by Lemma 1). Every closed subset of Polish space.

Proposition 6. Let $A \subset X$, X Polish. Then the following are equivalent:

(i) A is analytic.

(ii) A is a continuous image of a Borel set.

(iii) A is the projection onto X of some Borel subset of $Y \times X$, Y Polish.

(iv) A is the projection onto X of some closed subset of $Y \times X$, Y Polish.

(v) A is the projection onto X of some Borel subset of $\mathcal{N} \times X$.

(vi) A is the projection onto X of some closed subset of $\mathcal{N} \times X$.

Note. $\mathcal{N} = \mathbb{N}^{\mathbb{N}}$, $\mathcal{N} \times \mathcal{N} = \mathbb{N}^{\mathbb{N} \sqcup \mathbb{N}}$ homeomorphic to $\mathcal{N}$, and $\mathcal{N}^{\mathbb{N}} = \mathbb{N}^{\mathbb{N} \times \mathbb{N}}$ is also homeomorphic to $\mathbb{N}$.

Proof. Enough to show (ii) $\Rightarrow$ (I) $\Rightarrow$ (vi).

(i) $\Rightarrow$ (vi) $A = f(\mathcal{N})$, f closed. A is the projection onto X of

$$\{(\mathbf{n}, f(\mathbf{n})) \mid \mathbf{n} \in \mathcal{N}\}$$

which is closed.

(ii) $\Rightarrow$ (i) Need: Borel $\implies$ analytic. Enough: every Borel set satisfies (vi). Π_1^0 is a subset of the sets satisfying (vi). Need that the set of sets satisfying (vi) is closed under countable union and intersection. Assume A_n is the projection of $F_n \subset \mathcal{N} \times X$, F_n closed. So $x \in A_n \iff \exists \mathbf{n} \in \mathcal{N}, (\mathbf{n}, x) \in F_n$. Then

$$x \in \bigcup_n A_n \iff \exists n \in \mathbb{N}\ \exists \mathbf{n} \in \mathcal{N} \quad (\mathbf{n}, x) \in F_n.$$

Let

$$F = \{(n, \mathbf{n}, x) \in \underbrace{\mathbb{N} \times \mathcal{N}}_{\mathcal{N}} \times X \mid (\mathbf{n}, x \in F_n)\}$$

which is closed and projects onto $\bigcup_n A_n$.

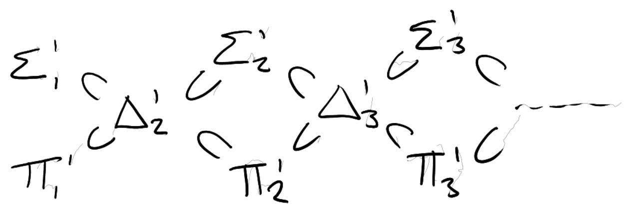
For intersection: $x \in \bigcap_n A_n$ if and only if $\forall n\ \exists \mathbf{n}, (\mathbf{n}, x) \in F_n$. Then

$$G = \{(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \dots, x) \in \mathcal{N}^{\mathbb{N}} \times X \mid (n_i, x) \in F_i\ \forall i\}$$

is closed and projects onto $\bigcap_n A_n$. □

Definition (Σ_n^1, Π_n^1)**.** Let Σ_1^1 be the set of analytic sets. Let Π_1^1 be the set of coanalytic sets, i.e. complements of analytic sets. For $1 \le n \le \omega$, let Σ_{n+1}^1 be the continuous image of Π_n^1 sets. Let Π_{n+1}^1 be the complements of Σ_{n+1}^1 sets.

As before:



projective hierarchy

$$P = \bigcup_{1 \le n < \omega} \Sigma_n^1 = \bigcup_{1 \le n < \omega} \Pi_n^1.$$

Theorem 7. There exists a universal analytic set $A \subset \mathcal{N} \times \mathcal{N}$.

Proof. Let U be a universal open set in $\mathcal{N} \times (\mathcal{N} \times \mathcal{N})$. So if $V \subset \mathcal{N} \times \mathcal{N}$ is open then there exists $\mathbf{p} \in \mathcal{N}$ such that

$$V = \{(\mathbf{m}, \mathbf{n}) \in \mathcal{N} \times \mathcal{N} \mid (\mathbf{p}, \mathbf{m}, \mathbf{n}) \in U\}.$$

Suppose $B \subset \mathcal{N}$ is analytic. So there exists closed $F \subset \mathcal{N} \times \mathcal{N}$ such that

$$B = \{\mathbf{n} \in \mathcal{N} \mid \exists \mathbf{m} \in \mathcal{N}, (\mathbf{m}, \mathbf{n}) \in F\}.$$

So $\exists \mathbf{p} \in \mathbb{N}$ such that

$$B = \{\mathbf{n} \in \mathcal{N} \mid \exists \mathbf{m} \in \mathcal{N}, (\mathbf{p}, \mathbf{m}, \mathbf{n}) \notin U\}.$$

Let

$$A = \{(\mathbf{r}, \mathbf{s}) \in \mathcal{N} \times \mathcal{N} \mid \exists \mathbf{m} \in \mathcal{N}, (\mathbf{r}, \mathbf{m}, \mathbf{s}) \notin U\}$$

This is a projection of a closed set, so analytic.

$$B = \{\mathbf{n} \in \mathcal{N} \mid (\mathbf{p}, \mathbf{n}) \in A\}. \qquad \square$$

Corollary 8. There exists an analytic, not coanalytic set in $\mathcal{N}$.

Proof. Let $A \subset \mathcal{N} \times \mathcal{N}$ be a universal analytic set, and $B = \{\mathbf{n} \in \mathcal{N} \mid (\mathbf{n}, \mathbf{n}) \in A\}$ analytic. If B is coanalytic, then

$$\exists \mathbf{m} \in \mathcal{N} \qquad B = \{\mathbf{n} \in \mathcal{N} \mid (\mathbf{m}, \mathbf{n}) \notin A\}.$$

Is $\mathbf{m} \in B$? No, contradiction. $\square$

Remark. So $B \in \Sigma_1^1 \setminus \Pi_1^1$, so B is not Borel (“$P \neq NP$”).

Aim: $\Sigma_1^1 \cap \Pi_1^1 = \mathcal{B}$. “$\supset$” is Proposition 6.

Theorem 9 (Lusin's Separation Theorem)**.** If A_1, A_2 are disjoint analytic sets, then there exists a orel set B, $A_1 \subset B$, $A_2 \subset X \setminus B$.

Proof. First: if $Y = \bigcup_n Y_n$, $Z = \bigcup_n Z_n$ and $\forall m, n$ Y_m, Z_n can be separated by Borel sets, then so can Y, Z. So for all m, n, find $Y_m \subset B_{m,n} \subset X \setminus Z_n$, $B_{m,n}$ Borel. Then

$$B = \bigcup_m \bigcap_n B_{m,n}$$

is Borel, and $Y \subset B \subset X \setminus Z$. Now suppose f, g are continuous and $f(\mathcal{N})$, $g(\mathcal{N})$ are disjoint, but cannot be separated. Recall

$$U_{m_1, m_2, \ldots, m_k} = \{\mathbf{n} \in \mathcal{N} \mid n_i = m_i, 1 \leq i \leq k\}$$

is our notation for the basic open sets in $\mathcal{N}$. $f(\mathcal{N}) = \bigcup_n f(U_n)$, $g(\mathcal{N}) = \bigcup_n g(U_n)$. There exists m_1, n_1 such that $f(U_{m_1})$, $g(U_{n_1})$ cannot be separated. Inductively, we get $\mathbf{m}, \mathbf{n} \in \mathcal{N}$ such that for all $\mathbf{m}, \mathbf{n} \in \mathcal{N}$, $f(U_{m_1, \ldots, m_k})$, $g(U_{n_1, \ldots, n_k})$ cannot be separated. But $\mathcal{N}$ is Hausdorff (and in fact we can separate points using the basic open sets U), which gives a contradiction. $\square$

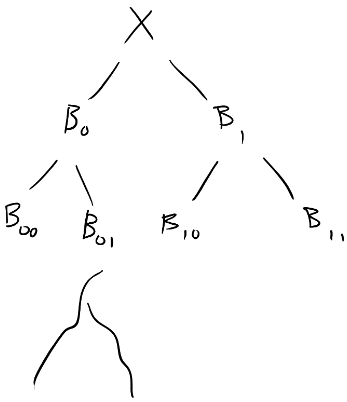
Corollary 10. $\Sigma_1^1 \cap \Pi_1^1 = \mathcal{B}$.

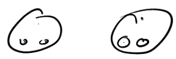
Example. Let $\Sigma = \bigcup_{k \in \mathbb{N}_0} \mathbb{N}^k$. $s, t \in \Sigma$, we write $s \prec t$ if $s = (n_1, \ldots, n_j)$, $t = (n_1, \ldots, n_i)$, $0 \leq j \leq i$. $s \in \Sigma$, $\mathbf{n} \in \mathcal{N}$, $s \prec \mathbf{n}$ if $s = (n_1, \ldots, n_j)$, $j \in \mathbb{N}_0$. $\mathbb{P}\Sigma = \{0, 1\}^\Sigma$ Polish space. $T \subset \Sigma$ is a *tree* if $s \prec t$, $t \in T \implies s \in T$. T is *well-founded* if $\nexists \mathbf{n} \in \mathcal{N}$ such that $\forall i$, $(n_1, \ldots, n_i) \in T$.

$$WFT = \{T \subset \Sigma \mid T \text{ is well-founded}\}$$

A subset A of a Polish space is *perfect* if A is closed and contains no isolated points. ($x \in A$ is *isolated* if $\exists r > 0$, $B_r(x) \cap A = \{x\}$).

Lemma 11. $A \neq \emptyset$ perfect set has cardinality $2^{\aleph_0}$.





Proof.

Closed balls, disjoint, diameter < 1, centres in A. $\{0,1\}^{\mathbb{N}} \hookrightarrow A$, so $\operatorname{card} A \geq 2^{\aleph_0}$. $\operatorname{card} A \leq \operatorname{card} \mathcal{N} = \aleph_0^{\aleph_0} = 2^{\aleph_0}$. □

Theorem 12. An analytic set is either countable or contains a non-empty perfect set. So Continuum Hypothesis holds for analytic sets.

$f(\mathcal{N})$. T tree

$$[T] = \{\mathbf{n} \in \mathcal{N} \mid (n_1, \ldots, n_i) \in T \ \forall i\}.$$

$[\Sigma] = \mathcal{N}$, $s \in \Sigma$,

$$T(s) = \{t \in \Sigma \mid t \prec s \text{ or } s \prec t\}.$$

$T^{(0)} = \Sigma$,

$$T^{(\alpha+1)} = (T^{(\alpha)})' = \{s \in T^{(\alpha)} \mid f([T^{(\alpha)}(s)]) \text{ is uncountable}\}.$$

$$T^{(\lambda)} = \bigcap_{\alpha<\lambda} T^{(\alpha)}.$$

$\exists \alpha < \omega_1$, $T^{(\alpha)} = T^{(\alpha+1)}$, ($\Sigma$ countable). Either $T^{(\alpha)} = \emptyset$ implies $f(\mathcal{N})$ countable or $T^{(\alpha)} \neq \emptyset$. Find a copy of $\{0,1\}^{\mathbb{N}} \subset [T^{(\alpha)}]$. The image of $\{0,1\}^{\mathbb{N}}$ is perfect.

Index