Logic and Set Theory

§ 1 Posets and Zorn's Lemma

Definition 1.1 Let A be a set

A partial order on A is a binary relation $\leq$

which is reflexive ie $x \leq x \quad \forall\, x \in A$

transitive $x \leq y$ and $y \leq z \Rightarrow x \leq z, \quad \forall\, x, y, z$

antisymmetric $x \leq y$ and $y \leq x \Rightarrow x = y \quad \forall\, x, y$

$\leq$ is a total order if in addition $x \leq y$ or $y \leq x \quad \forall\, x, y \in A$

A poset (or partially ordered set) is a pair $(A, \leq)$
where $\leq$ is a partial order on A

examples 1.2 a. The usual order $\leq$ on $\mathbb{R}$, defined by $x \leq y \Leftrightarrow \exists\, z \in \mathbb{R}$ such that $z^2 = y - x$

is a total order
Similarly the restrictions of this $\leq$ to $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N}$, etc

b. For any set A, $\mathcal{P}A$ is the set of all subsets of A (the power set of A)
$\mathcal{P}A$ is partially ordered by $B \leq C \Leftrightarrow \forall\, x,\ x \in B \Rightarrow x \in C$
Similarly the set $\mathrm{Sub}(G)$ of subgroups of a group G
the set $O(X)$ of open subsets of a topological space, etc

c. On $\mathbb{N}$, we have a relation $|$ defined by $m \mid n \Leftrightarrow \exists\, p \in \mathbb{N}$ such that $mp = n$

Note that $0 \in \mathbb{N}$ is the least element for the ordering in a.,
but the greatest element for $|$

d. Let Σ be a set of abstract symbols.
Σ^* is the set of all words over Σ,
ie finite strings of members of Σ
We can partially order Σ^* by setting $v \leq w \Leftrightarrow \exists\, u, x$ such that $w = uvx$
(the subword ordering)
or by $v \leq w \Leftrightarrow \exists\, x$ such that $w = vx$
(the prefix ordering)

e. A partial function $f : A \rightharpoonup B$ is a function from a subset of A to B
$[A \rightharpoonup B]$ is the set of all partial functions $A \rightharpoonup B$
The set is partially ordered by $f \leq g \Leftrightarrow$ g extends f
ie $g(x)$ is defined and equals $f(x)$
whenever $f(x)$ is defined

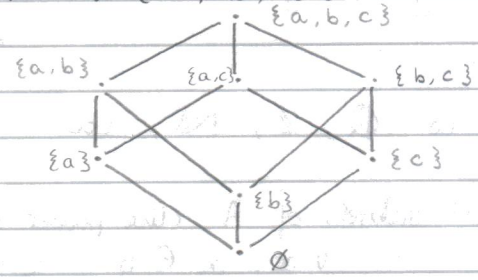
We can represent finite posets by diagrams in which each element is represented by a dot, and we draw an upward line from x to y for each 'irreducible' instance of $x \leq y$

Definition 1.3 y covers x in a Poset $(P, \leq)$ $(x \lhd y)$
if $x < y$ $(x \leq y$ and $x \neq y)$
but $\forall z \in P$, $x \leq z \leq y \Rightarrow z = x$ or $z = y$

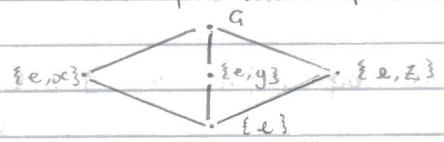
Note In a finite poset P, $x \leq y \Leftrightarrow \exists$ a chain
$$x = z_0 \lhd z_1 \lhd z_2 \ldots \lhd z_n = y$$
for some $n \geq 0$ (the case $n = 0$ means $x = y$)

The Hasse diagram of a finite poset P represents the elements of P by dots and the instances of $\lhd$ by lines

examples 1.4 a. $P = \mathcal{P}\{a, b, c\}$



b. $P = \mathrm{Sub}(G)$ where G is the non-cyclic group $\{e, x, y, z\}$ of order 4



Definition 1.5 Let S be a subset of a poset $(P, \leq)$

a. x is {the / a} greatest member of S
if $x \in S$ and $\forall y \in S$, $y \leq x$
(dually, least member)

b. x is an upper bound for S
if ~~$x \in S$ and~~ $\forall y \in S$, $y \leq x$
(dually, lower bound)

c. x is the least upper bound of S
if it is the least member of $\{y \in P : y$ is an upper bound for $S\}$
(we also say x is the supremum or join of S, and write $x = \bigvee S$)
(dually, greatest lower bound, infimum, or meet, denoted $x = \bigwedge S$)

d. $(P, \leq)$ is complete
if every $S \subseteq P$ has a least upper bound

Logic and Set Theory

Lemma 1.6 if $(P, \leq)$ is complete, then so is $(P, \geqslant)$

Proof Let $S \subseteq P$; required to prove S has a greatest lower bound (for $\leq$)
Let $T = \{ x \in P : x \text{ is a lower bound for } S \}$
and consider $y = \bigvee T$
If $z \in S$ and $x \in T$ then $x \leq z$,
so every $z \in S$ is an upper bound for T
and hence satisfies $z \geqslant y$
So y is a lower bound for S,
and hence the greatest member of T. □

examples 1.7 a. $\wp A$ is complete: if $S \subseteq \wp A$
then $\bigcup S = \{ x \in A : \exists B \in S, x \in B \}$
is the least ~~st~~ upper bound for S

b. $\mathrm{Sub}(G)$ is complete: if $S \subseteq \mathrm{sub}(G)$
then $\bigcap S$ is a subgroup and a greatest lower bound for S

c. The usual order $\leqslant$ on $\mathbb{R}$ is not complete but $\mathbb{R} \cup \{ \pm \infty \}$ is complete

Lecture 2

Definition 1.8 a. A chain in a poset $(P, \leq)$ is a non-empty subset $C \subseteq P$ which is totally ordered by $\leq$

b. $(P, \leq)$ is chain complete if every chain $C \subseteq P$ has a least upper bound in P

examples 1.9 a. Let G be a group
Let P be the set of abelian subgroups of G, ordered by $\leqslant$
If $H, K \in P$ contain elements x, y respectively with $xy \neq yx$
then $\{ H, K \}$ has no upper bound in P,
so P is not complete
However, if $\{ H_i \mid i \in I \}$ is a chain of abelian subgroups of G,
then $\bigcup_{i \in I} H_i$ is a subgroup
since if $x, y \in \bigcup_{i \in I} H_i$,
then $x \in H_i$ and $y \in H_j$ for some i, j
but either $H_i \leqslant H_j$ or $H_j \leqslant H_i$
so either $\{ x, y \} \subseteq H_j$ or $\{ x, y \} \subseteq H_i$
hence $x, y \in \bigcup_{i \in I} H_i$
and $\bigcup_{i \in I} H_i$ is abelian, by a similar argument
So $\{ H_i : i \in I \}$ has a least upper bound in P

b. Consider the set $[A \to B]$ of partial functions $A \to B$, ordered by extension.
If f, g are such that, for some $x \in A$, $f(x)$ and $g(x)$ are defined and $f(x) \neq g(x)$
then $\{f, g\}$ has no upper bound in $[A \to B]$
However, if $\{f_i \mid i \in I\}$ is a chain in $[A \to B]$,
then $\forall\, i, j$, $f_i(x) = f_j(x)$ whenever both are defined,
so $\exists$ a unique $g : A \to B$ with domain $\bigcup_{i \in I} \operatorname{dom} f_i$,
such that $g(x) = f_i(x)$ whenever $f_i(x)$ is defined
and this is the least upper bound of $\{f_i \mid i \in I\}$

Recursive Definitions

Consider the factorial function $f : \mathbb{N} \to \mathbb{N}$
defined by
$$f(n) = \begin{cases} 1 & \text{if } n = 0 \\ n f(n-1) & \text{if } n > 0 \end{cases}$$

To avoid the circularity in this, consider the recursion as defining a function
$$\Phi : [\mathbb{N} \to \mathbb{N}] \to [\mathbb{N} \to \mathbb{N}]$$
$$\Phi(g)(n) = \begin{cases} 1 & \text{if } n = 0 \\ n\,g(n-1) & \text{if } n > 0 \text{ and } g(n-1) \text{ is defined} \\ \text{undefined} & \text{otherwise} \end{cases}$$

Then the factorial function should be the unique f satisfying $\Phi(f) = f$.

Theorem 1.10 Knaster - Tarski Theorem
Let P be a complete poset
let $f : P \to P$ be an order preserving map (ie $x \leq y \Rightarrow f(x) \leq f(y)$)
Then f has a fixed point

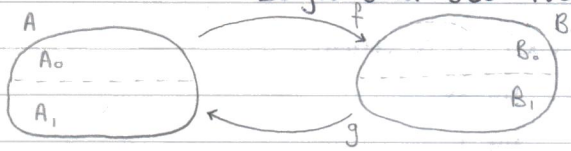
Proof exists, by completeness
Let $S = \{x \in P \mid x \leq f(x)\}$ be the set of prefixed points of f
and let $y = \bigvee S$
$\forall\, x \in S$, we have $x \leq y$
and hence $x \leq f(x) \leq f(y)$
so $f(y)$ is an upper bound for S
So $y \leq f(y)$, and hence $f(y) \leq f(f(y))$
ie $f(y) \in S$

$f(y) \in S$, y is a least upper bound
$\Rightarrow f(y) \leq y$
$\Rightarrow f(y) = y$ by ~~as~~ antisymmetry. $\square$

Corollary 1.11 Cantor - Bernstein Theorem
Suppose we have sets A and B
and injective functions $f : A \to B$, $g : B \to A$
Then $\exists$ a bijection $h : A \to B$

Logic and Set Theory

Proof



We seek a fixed point A_0 of the function $\overline{\Phi} : \wp A \to \wp A$
defined by $\overline{\Phi}(A') = A \setminus g(B \setminus f(A'))$
$\overline{\Phi}$ is order-preserving, since $f(-)$ and $g(-)$ preserve order
and $A \setminus (-)$ and $B \setminus (-)$ reverse it
So by the Knaster-Tarski Theorem $\overline{\Phi}$ has a fixed point $A_0 \subseteq A$,
and we can define h by

$$h(x) = \begin{cases} f(x) & \text{if } x \in A_0 \\ g^{-1}(x) & \text{if } x \in A \setminus A_0 \end{cases}$$

□

Note

The discrete ordering $x \leq y \Leftrightarrow x = y$ on any set A is chain-complete,
and any $f : A \to A$ is order-preserving

Given a poset P, f is inflationary if $x \leq f(x)$ $\forall\, x \in P$

Theorem 1.12 Bourbaki-Witt Theorem

N. Bourbaki, Sur le théorème de Zorn, Arch. Math. 2 (1950), 434-437

Let P be a chain-complete poset
and $f : P \to P$ an inflationary map
Then, for every $x \in P$, $\exists\, y \in P$ with $x \leq y = f(y)$

Non-proof

Set $x_0 = x$
$x_1 = f(x_0)$
$x_2 = f(x_1)$
⋮

set $x_\omega = \bigvee \{ x_0, x_1, x_2, \dots \}$
$x_{\omega+1} = f(x_\omega)$
$x_{\omega+2} = f(x_{\omega+1})$
⋮

set $x_{\omega+\omega} = \bigvee \{ x_\alpha \mid \alpha = 0, 1, 2, \dots, \omega, \omega+1, \omega+2, \dots \}$
Surely to goodness this must stop!

Proof

Define a subset $C \subseteq P$ to be closed if
i. $\forall\, y \in P$, $y \in C \Rightarrow f(y) \in C$
and ii. $\forall\, D \subseteq C$, D is a chain $\Rightarrow \bigvee D \in C$
Any intersection of closed sets is closed; in particular,
$C(x) = \bigcap \{ C \subseteq P \mid C \text{ is closed}, x \in C \}$
is the smallest closed set containing x

[Intuitively, $C(x)$ is the set of all elements x_α constructed in the non-proof]

Suppose we can show that $C(x)$ is a chain.
Then $y = \bigvee C(x) \in C(x)$ by ii,
and hence $f(y) \in C(x)$ by i
so $f(y) \leq y$
and hence $y = f(y)$, ie y is the fixed point we seek

Required to show that $C(x)$ is a chain

Step 1 $\forall\, y \in C(x),\ x \leq y$

Proof $\uparrow(x) = \{ y \in P \mid y \geq x \}$ is a closed set containing x,
so $C(x) \subseteq \uparrow(x)$

Definition $y \in C(x)$ is normal if $\forall\, z \in C(x),\ z < y \Rightarrow f(z) \leq y$

Step 2 If y is normal, then $\forall\, z \in C(x),\ z \leq y$ or $f(y) \leq z$

Proof Consider $D = \{ z \in C(x) \mid z \leq y \text{ or } f(y) \leq z \}$
Then $x \in D$ by Step 1
Suppose $z \in D$,
then either $z < y$, in which case $f(z) \leq y$ by normality
or $z = y$, in which case $f(y) = f(z)$
or $f(y) \leq z$, in which case $f(y) \leq f(z)$ since f is inflationary
so $f(z) \in D$
If $E \subseteq D$ is a chain,
then either $\forall\, z \in E,\ z \leq y$, in which case $\bigvee E \leq y$
or $\exists\, z \in E$, such that $f(y) \leq z$, in which case $f(y) \leq \bigvee E$
so $\bigvee E \in D$
So D is closed and contains x
hence $D = C(x)$

Step 3 $\forall\, y \in C(x)$, y is normal

Proof Consider $N = \{ y \in C(x) \mid y \text{ is normal} \}$
Then $x \in N$, since $z < x$ is never satisfied for $z \in C(x)$, by Step 1
Suppose $y \in N$, and $z \in C(x)$ satisfies $z < f(y)$
Then $z \neq f(y)$, so $z \leq y$ by Step 2
so either $z < y$, in which case $f(z) \leq y \leq f(y)$
or $z = y$, in which case $f(z) = f(y)$
Hence $f(y) \in N$
Finally, suppose $M \subseteq N$ is a chain, and $z < \bigvee M$.
Then $\exists\, y \in M$ such that $y \not\leq z$
and hence $\exists\, y \in M$ such that $z < y$ by Step 2
So $f(z) \leq y \leq \bigvee M$
So, as before, we must have $N = C(x)$
Hence $C(x)$ is a chain by Step 2, and $y = \bigvee C(x)$ is the required fixed point $\square$

Logic and Set Theory

Corollary 1-13 Suppose $(P, \leq)$ is chain-complete
and $f : P \to P$ is order-preserving
Then for any $x \in P$ with $x \leq f(x)$, $\exists$ a least $y \geq x$ with $y = f(y)$
In particular if P has a least element, then f has a least fixed point

Proof Let $Q = \{ y \in P \mid y \leq f(y) \}$
Then $y \in Q \Rightarrow f(y) \leq f(f(y))$
$\overset{\text{given}}{\Rightarrow} f(y) \in Q$ since f is order-preserving
and $C \subseteq Q$, C a chain
then $z = \bigvee C$ satisfies $\forall y \in C,\ y \leq z$
so $\forall y \in C,\ y \leq f(y) \leq f(z)$
so $f(z)$ is an upper bound for C, and $z \leq f(z)$
So Q is a chain-complete poset,
and $f|_Q$ is an inflationary map $Q \to Q$
Hence, for any $x \in Q$, $\exists\, y \in Q$ with $x \leq y = f(y)$

If $z \in Q$ is any other fixed point with $z \geq x$, then
$\downarrow z = \{ w \in Q \mid w \leq z \}$ is a closed set:
if $w \in \downarrow z$, then $f(w) \leq f(z) = z$
$\Rightarrow f(w) \in \downarrow z$
Hence z is an upper bound for the set $C(x)$ constructed in the proof of 1·12
and so the fixed point $y = \bigvee C(x)$ constructed in the proof
satisfies $y \leq z$
The final assertion follows from the fact that a least element 0 of P
necessarily satisfies $0 \leq f(0)$ □

Recall the function $\Phi : [\mathbb{N} \to \mathbb{N}] \to [\mathbb{N} \to \mathbb{N}]$
arising from the recursive definition of the factorial function.
$[\mathbb{N} \to \mathbb{N}]$ is chain-complete and has a least element
(the every-where undefined function)
and Φ is order-preserving: if $g_1 \leq g_2$, then $\Phi(g_1) \leq \Phi(g_2)$
So Φ has a least fixed point.

Note If $\Phi(f) = f$, then $0 \in \mathrm{dom}(f)$
and $\forall n,\ n \in \mathrm{dom}(f) \Rightarrow n+1 \in \mathrm{dom}(f)$
so $\mathrm{dom}\, f = \mathbb{N}$
ie f is total and hence maximal in $[\mathbb{N} \to \mathbb{N}]$
ie $\forall g \in [\mathbb{N} \to \mathbb{N}],\ f \not< g$
Hence in this case f is the unique fixed point of Φ.

The Axiom of Choice

Given a set $\{ A_i \mid i \in I \}$ of sets with $A_i \neq \emptyset$ for each i,
$\exists$ a choice function $f : I \to \bigcup_{i \in I} A_i$
such that $f(i) \in A_i \quad \forall\, i \in I$

Zorn's Lemma (1935)
Given a chain complete poset $(P, \leq)$,
every $x \in P$ lies below some maximal element

Corollary 1.14 Axiom of Choice $\Rightarrow$ Zorn's Lemma

Proof Let $(P, \leq)$ be chain-complete.
Define a family of sets $\{ A_x \mid x \in P \}$ as follows:
if x is not maximal, $A_x = \{ y \in P \mid x < y \}$
if x is maximal, $A_x = \{ x \}$
By the Axiom of Choice, $\exists\ f : P \to P$ with $f(x) \in A_x\ \forall x$
By construction, $f(x) = x \Leftrightarrow x$ is maximal in P
And such an element exists, by the Bourbaki-Witt Theorem. □

Lecture 4

1.5 Applications of Zorn's Lemma

a. Zorn's Lemma $\Rightarrow$ Axiom of Choice (so the two are logically equivalent)

Proof Given a family $\{ A_i \mid i \in I \}$ of nonempty sets,
consider the set $P \subseteq [I \to \bigcup_{i \in I} A_i]$ of partial choice functions,
ie partial functions $f : I \to \bigcup_{i \in I} A_i$
satisfying $f(i) \in A_i$ whenever $f(i)$ is defined
It is easy to see that if $\{ f_j \mid j \in J \}$ is a chain in P,
then the join $\bigvee \{ f_j \mid j \in J \}$ as constructed in $[I \to \bigcup_{i \in I} A_i]$
is a partial choice function
So P is chain complete
And P is non-empty, as the everywhere undefined function is in P
So by Zorn's Lemma P has a maximal element f_0
Suppose f_0 is not total
Pick $i_0 \in I \setminus \operatorname{dom} f_0$
$x_0 \in A_{i_0}$
Now define

$$f_1(i) = \begin{cases} f_0(i) & \text{if } f_0(i) \text{ is undefined} \\ x_0 & \text{if } i = i_0 \\ \text{undefined} & \text{otherwise} \end{cases}$$

then $f_1 \in P$ and $f_0 < f_1$ $\Rightarrow\Leftarrow$
So f_0 is a total choice function □

b. Hamel's Theorem
Every vector space has a basis

Proof Let V be a vector space
Consider the set P of all linearly independent subsets of V, ordered by inclusion
If $\{ S_i \mid i \in I \} \subseteq P$ is a chain, consider $\bigcup_{i \in I} S_i$
if we had a non-trivial linear relation $\sum_{j=1}^{n} \lambda_j x_j = 0$ on this set,
then for each j $\exists\ i_j$ such that $x_j \in S_{i_j}$

Logic and Set Theory

Since the S_i are totally ordered by $\subseteq$,
$\exists\ i$ such that $x_1, \ldots, x_n$ belong to S_i
so $\sum \lambda_j x_j = 0$ is a linear relation on S_i $\Rightarrow\Leftarrow$
So $\bigcup_{i \in I} S_i$ is linearly independent, and P is chain-complete
Hence any linearly independent set (in particular $\emptyset$)
is contained in a maximal linearly independent set S_0, say
Suppose S_0 does not span V
Pick $x \in V \setminus \langle S_0 \rangle$
and define $S_1 = S_0 \cup \{x\}$
Then S_1 is linearly independent,
since a non-trivial linear relation on it would be
either a linear relation on S_0
or an expression for x as a linear combination of members of S_0
But $S_0 \subset S_1$ $\Rightarrow\Leftarrow$
So S_0 spans V, and hence is a basis $\square$

c. Krull's Theorem
Every proper ideal in a ring R (with 1) is contained in a maximal ideal

Proof
This is exactly Zorn's Lemma applied to the poset P of proper ideals of R, ordered by $\subseteq$.
So we need to show P is chain-complete.
But if $\{ I_j \mid j \in J \}$ is a chain of proper ideals
then $\bigcup_{j \in J} I_j$ is an ideal (cf 1.9)
and it's proper since $I \lhd R$ is proper $\Leftrightarrow 1 \notin I$ $\square$

Definition 1.16 a. A lattice is a poset $(L, \leq)$ in which every finite subset (including $\emptyset$) has both a join and a meet
In particular, L contains $0 = \bigvee \emptyset$
and $1 = \bigwedge \emptyset$
and it has binary operations $\vee$, $\wedge$ defined by $a \vee b = \bigvee \{a, b\}$
$a \wedge b = \bigwedge \{a, b\}$
Note that the order relation is definable from either $\vee$ or $\wedge$, since
$a \leq b \Leftrightarrow a \vee b = b$
$\Leftrightarrow a \wedge b = a$
So a lattice homomorphism $f : L \to M$ (ie a function preserving $\wedge$, $\vee$, 0 and 1)
is automatically order-preserving

b. A lattice L is distributive if it satisfies the identity
$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ $\forall\ a, b, c \in L$

c. Elements a, b are complementary in a lattice L
if $a \vee b = 1$ and $a \wedge b = 0$

A Boolean algebra is a distributive lattice in which every element has a complement

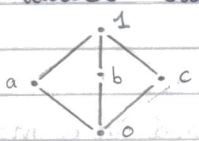
examples 1·17a For any A, $\mathcal{P}A = \{B \mid B \subseteq A\}$ is a Boolean algebra.
Given subsets B, C, D, we have

$$\begin{aligned} x \in B \cap (C \cup D) &\Leftrightarrow x \in B \text{ and } (x \in C \text{ or } x \in D) \\ &\Leftrightarrow (x \in B \text{ and } x \in C) \\ &\quad \text{or } (x \in B \text{ and } x \in D) \\ &\Leftrightarrow x \in (B \cap C) \cup (B \cap D) \end{aligned}$$

And any $B \subseteq A$ has a complement $A \setminus B = \{x \in A \mid x \notin B\}$

b. If L is a totally ordered set with greatest and least elements, then it's a lattice with $a \vee b = \max\{a, b\}$
$a \wedge b = \min\{a, b\}$
It is straightforward to verify that L is distributive, but no element other than 0 or 1 has a complement

c. The lattice $\mathrm{Sub}(G)$, where G is the non-cyclic group of order 4, is not distributive



$$\begin{aligned} a \wedge (b \vee c) &= a \wedge 1 \\ &= a \\ (a \wedge b) \vee (a \wedge c) &= 0 \vee 0 \\ &= 0 \end{aligned}$$

$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$

Lemma 1·18 i. If $(L, \leq)$ is a distributive lattice, so is $(L, \geqslant)$
ii. In a distributive lattice, any ~~be~~ element has at most one complement

Proof i. We have to show $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$

$$\begin{aligned} \text{But } (a \vee b) \wedge (a \vee c) &= ((a \vee b) \wedge a) \vee ((a \vee b) \wedge c) \\ &= \overset{a \leq}{a} \vee (a \wedge c) \vee (b \wedge c) \\ &= \overset{a \leq}{a} \vee (b \wedge c), \qquad \text{as required} \end{aligned}$$

ii. Suppose b, c are both complements of a

$$\begin{aligned} \text{Then } b \wedge (a \vee c) &= b \wedge 1 \\ &= b \\ \text{But } b \wedge (a \vee c) &= (b \wedge a) \vee (b \wedge c) \\ &= 0 \vee (b \wedge c) \\ &= b \wedge c \end{aligned}$$

So $b \leq c$
Similarly $c \leq b$
$\Rightarrow b = c$ □

Lecture 5

Lemma 1·19 Let L be a distribution lattice and a, b elements of L with $a \not\leq b$
Then there is a lattice homomorphism $f : L \to 2 = \{0, 1\}$ with $f(a) = 1$, $f(b) = 0$

Logic and Set Theory

Proof

Let P be the set of all pairs (A, B) of subsets of L satisfying

i. A is a filter: $x \in A,\ x \leq y \Rightarrow y \in A$
$x, y \in A \Rightarrow x \wedge y \in A$
$1 \in A$

ii. B is an ideal: $x \in B,\ y \leq x \Rightarrow y \in B$
$x, y \in B \Rightarrow x \vee y \in B$
$0 \in B$

iii. $A \cap B = \emptyset$

Partially order P by $(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow (A_1 \subset A_2$ and $B_1 \subset B_2)$

To show P is chain-complete, consider a chain $\{(A_i, B_i) \mid i \in I\}$ of elements in P

Then $\bigcup_{i \in I} A_i$ is a filter: if $x, y \in \bigcup_{i \in I} A_i$
then $\exists\ i$ such that $\{x, y\} \subset A_i$
and hence $x \wedge y \in A_i$

$\bigcup_{i \in I} B_i$ is an ideal: similarly

$(\bigcup_{i \in I} A_i) \cap (\bigcup_{i \in I} B_i) = \emptyset$: if $x \in (\bigcup_{i \in I} A_i) \cap (\bigcup_{i \in I} B_i)$
then $x \in A_i$ and $x \in B_j$, some $i, j \in I$
but either $A_i \subset A_j$, $x \in A_j \cap B_j$
or $B_j \subset B_i$, $x \in A_i \cap B_i$ $\Rightarrow\Leftarrow$

So $(\bigcup_{i \in I} A_i, \bigcup_{i \in I} B_i)$ is a least upper bound for the chain P

The pair $(\uparrow(a), \downarrow(b))$ is an element of P since $a \not\leq b$
so $\exists$ a maximal element (A_0, B_0) lying above it
(ie such that $a \in A_0$, $b \in B_0$)

Suppose $A_0 \cup B_0 \neq L$

Let $c \in L \setminus (A_0 \cup B_0)$

The set $A_1 = \{x \in L \mid x \geq (y \wedge c)$ for some $y \in A_0\}$
is a filter strictly containing A_0

So $(A_1, B_0) \notin P$ by maximality of (A_0, B_0) ($A_1 \cap B_0 \neq \emptyset$, ie)
and hence $\wedge\ \exists\ x \in A_0$ such that $x \wedge c \in B_0$

Similarly $B_1 = \{x \in L \mid x \leq (y \vee c)$ for some $y \in B_0\}$
is an ideal strictly containing B_0

So $A_0 \cap B_1 \neq \emptyset$
ie $\exists\ y \in B_0$ such that $y \vee c \in A_0$

Now $x \wedge (y \vee c) \in A_0$, since it's a meet of two elements of A_0

But $x \wedge (y \vee c) = (x \wedge y) \vee (x \wedge c) \in B_0$
since it's a join of two elements of B_0

So $A_0 \cap B_0 \neq \emptyset$ $\Rightarrow\Leftarrow$

Hence $A_0 \cup B_0 = L$ and we have a (total) function $f : L \to 2$
defined by $f(x) = \begin{cases} 1 & \text{if } x \in A_0 \\ 0 & \text{if } x \in B_0 \end{cases}$

And f preserves $\wedge$: $f(x \wedge y) = 1 \Leftrightarrow x \wedge y \in A_0$
$\Leftrightarrow \{x, y\} \in A_0$
$\Leftrightarrow f(x) = f(y) = 1$

Similarly f preserves $\vee$

So f is a lattice homomorphism, and maps a to 1, b to 0. $\square$

Theorem 1.20 Birkhoff - Stone
Any distributive lattice is isomorphic to a sublattice of a power set

Proof Given a distributive lattice L,
let F be the set of all homomorphisms $L \to 2$
define $\overline{\Phi} : L \to \mathcal{P}F$
by $\overline{\Phi}(x) = \{ f \in F \mid f(x) = 1 \}$
Then $\overline{\Phi}$ is a lattice homomorphism:

$f \in \overline{\Phi}(x \wedge y) \Leftrightarrow f(x \wedge y) = 1 \quad \Leftrightarrow f(x) \wedge f(y) = 1$, f is a homomorphism
$\Leftrightarrow f(x) = f(y) = 1$
$\Leftrightarrow f \in \overline{\Phi}(x)$ and $f \in \overline{\Phi}(y)$
ie $f \in \overline{\Phi}(x) \cap \overline{\Phi}(y)$

$f \in \overline{\Phi}(x \vee y) \Leftrightarrow f(x \vee y) = 1$
$\Leftrightarrow f(x) = 1$ or $f(y) = 1$
$\Leftrightarrow f \in \overline{\Phi}(x) \cup \overline{\Phi}(y)$

By 1.19, $\overline{\Phi}$ is injective,
since if $a \not\leq b$ then $\exists f \in \overline{\Phi}(a) \setminus \overline{\Phi}(b)$
So $\overline{\Phi}$ is an isomorphism from L to its image, $\overline{\Phi}(L)$
which is a sublattice of $\mathcal{P}F$ $\square$

Remark If L is a Boolean algebra
then im $\overline{\Phi}$ is a Boolean subalgebra of $\mathcal{P}F$
cc it's closed under taking complements

§2 Propositional Calculus

In the propositional calculus, we assume we are given a set P of 'primitive propositions', which are abstract symbols capable of being 'true' or 'false'

Definition 2.1 Given a set P of primitive propositions,
a valuation of P is a function $f : P \to 2 = \{0, 1\}$
The set $\mathcal{L}(P)$ of compound propositions or propositional formulae over P is defined recursively by
i. if $p \in P$, then $p \in \mathcal{L}(P)$
ii. if $s, t \in \mathcal{L}(P)$, so is $(s \Rightarrow t)$
iii. $\bot \in \mathcal{L}(P)$

$\bot$ is 'false' or 'bottom':
the statement that is always false

Formally, if $\Sigma = P \cup \{ (,), \Rightarrow, \bot \}$
then $\mathcal{L}(P)$ is the smallest subset of Σ^* closed under i – iii

Logic and Set Theory

Given a formulation of P, $v : P \to 2$,
we extend it to a function $\bar{v} : \mathcal{L}(P) \to 2$
by setting $\bar{v}(p) = v(p)$ if $p \in P$
$\bar{v}(s \Rightarrow t) = 1 \Leftrightarrow$ either $\bar{v}(t) = 1$
or $\bar{v}(s) = 0$

$\bar{v}(\bot) = 0$

We define the compound propositions $\neg p$, $\top$, $p \vee q$, $p \wedge q$, $p \Leftrightarrow q$
to be $\neg p : (p \Rightarrow \bot)$
$\top : \neg \bot$
$p \vee q : (\neg p) \Rightarrow q$
$p \wedge q : \neg (p \Rightarrow \neg q)$
$p \Leftrightarrow q : (p \Rightarrow q) \wedge (q \Rightarrow p)$

p	$\neg p$
0	1
1	0

p	q	$p \vee q$	$p \wedge q$	$p \Leftrightarrow q$
0	0	0	0	1
0	1	1	0	0
1	0	1	0	0
1	1	1	1	1

Lemma 2.2 Functional Completeness
For any function $f : 2^n \to 2$,
$\exists$ a compound proposition in n primitive propositions
whose truth table is f

Proof by induction on n
$n = 0$: the two maps $2^0 \xrightarrow{0} 2$
$\xrightarrow{1} 2$
are the truth tables of $\bot$ and $\top$ respectively
Suppose true for n
Given $f : 2^{n+1} \to 2$
define $f_0 : 2^n \to 2$
$f_0(x_1, \ldots, x_n) = f(x_1, \ldots, x_n, 0)$
$f_1 : 2^n \to 2$
$f_1(x_1, \ldots, x_n) = f(x_1, \ldots, x_n, 1)$
Let t_0, t_1 have truth tables f_0, f_1 respectively
consider $t = (\neg p_{n+1} \wedge t_0) \vee (p_{n+1} \wedge t_1)$
This has truth table f □

Note Two formulae s and t have the same truth-table
iff $(s \Leftrightarrow t)$ has a truth table consisting entirely of 1's

Definition 2.3 Suppose $S \subseteq \mathcal{L}(P)$, $t \in \mathcal{L}(P)$

S semantically entails t, denoted ~~$S \models$~~ $S \models t$

if every valuation $v : P \to 2$ satisfying $\bar{v}(s) = 1 \quad \forall s \in S$

also satisfies $\bar{v}(t) = 1$.

In the particular case $S = \emptyset$, we write $\models t$,

and call t a tautology

For S finite, it is easy to determine whether $S \models t$

Write down the truth-tables of all members of S and of t,

and cross out the rows in which $\bar{v}(s) = 0$ for some $s \in S$

examples 2.4 a. $\{p, (p \Rightarrow q)\} \models q$

p	q	$p \Rightarrow q$
0	0	1
0	1	1
1	0	0
1	1	1

b. $\models (p \Rightarrow (q \Rightarrow p))$

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

c. $\models ((p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r)))$

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

Logic and Set Theory

Definition 2.5 We construct a deduction-system for the Propositional Calculus as follows:
as axioms, we take all formulae of the form

$s \Rightarrow (t \Rightarrow s)$ (K)

$(s \Rightarrow (t \Rightarrow u)) \Rightarrow ((s \Rightarrow t) \Rightarrow (s \Rightarrow u))$ (S)

or $\neg\neg s \Rightarrow s$ (T)

where s, t, u are arbitrary elements of $\mathcal{L}(P)$
as our rule of inference, we take modus ponens:
from s and $s \Rightarrow t$, we can infer t

Given a set $S \subseteq \mathcal{L}(P)$ of hypotheses,
we define a deduction from S to be a finite list $t_1, t_2, t_3, \ldots, t_n$ of formulae,
such that for each $i \leq n$
either $t_i \in S$
or t_i is an axiom
or $\exists\, j, k < i$ such that t_k is $t_j \Rightarrow t_i$

S syntatically entails t, denoted $S \vdash t$,
if $\exists$ a deduction $(t_1, t_2, \ldots, t_n)$ from S with $t_n = t$
If $S = \emptyset$, we say t is a theorem, and write $\vdash t$

examples 2.6 a. We show $\vdash (p \Rightarrow p)$

$p \Rightarrow (p \Rightarrow p)$	by (K)
$p \Rightarrow ((p \Rightarrow p) \Rightarrow p)$	by (K)
$(p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p))$	by (S)
$(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)$	MP from lines 2 & 3
$p \Rightarrow p$	MP from lines 1 & 4

b. We show $\{(p \Rightarrow q), (q \Rightarrow r)\} \vdash (p \Rightarrow r)$

$q \Rightarrow r$	hypothesis
$(q \Rightarrow r) \Rightarrow (p \Rightarrow (q \Rightarrow r))$	by (K)
$p \Rightarrow (q \Rightarrow r)$	MP
$(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$	by (S)
$(p \Rightarrow q) \Rightarrow (p \Rightarrow r)$	MP
$p \Rightarrow q$	hypothesis
$p \Rightarrow r$	MP

Lemma 2.7 Soundness Theorem
If $S \vdash t$, then $S \models t$
In particular, every theorem is a tautology

Proof Let $(t_1, t_2, \ldots, t_n = t)$ be a deduction of t from S.
We show by induction on i that $S \models t_i \quad \forall i$
If $t_i \in S$ then trivially $S \models t_i$
If t_i is an axiom, then $\models t_i$
so $S \models t_i$
if $t_k = (t_j \Rightarrow t_i)$ for some $j, k < i$,
then by the induction hypothesis any v making all of S true
satisfies $\bar{v}(t_j) = 1$
and $\bar{v}(t_j \Rightarrow t_i) = 1$
This forces $\bar{v}(t_i) = 1$ □

Theorem 2.8 Deduction Theorem
Let $S \subseteq \mathcal{L}(P)$, $\sigma, t \in \mathcal{L}(P)$
Then $S \vdash (\sigma \Rightarrow t)$ iff $S \cup \{\sigma\} \vdash t$

Proof $\Rightarrow$ If we have a deduction of $(\sigma \Rightarrow t)$ from S, we can write it down and add

σ	hypothesis
t	MP

to obtain a deduction of t from $S \cup \{\sigma\}$

$\Leftarrow$ Suppose $(t_1, t_2, \ldots, t_n = t)$ is a deduction of t from $S \cup \{\sigma\}$
We show for each i that $S \vdash (\sigma \Rightarrow t_i)$
If $t_i \in S$, we write down

t_i	hypothesis
$t_i \Rightarrow (\sigma \Rightarrow t_i)$	by (K)
$\sigma \Rightarrow t_i$	MP

If t_i is an axiom we write down

t_i	(axiom)
$t_i \Rightarrow (\sigma \Rightarrow t_i)$	(K)
$\sigma \Rightarrow t_i$	(MP)

If $t_i = \sigma$, we write down the proof that $\sigma \Rightarrow \sigma$ from example 2.6a.
If $t_k = (t_j \Rightarrow t_i)$ for some $j, k < i$, we write down
our deductions of $\sigma \Rightarrow t_j$
$\sigma \Rightarrow (t_j \Rightarrow t_i)$ from S
and add $(\sigma \Rightarrow (t_j \Rightarrow t_i)) \Rightarrow ((\sigma \Rightarrow t_j) \Rightarrow (\sigma \Rightarrow t_i))$ by (S)
$(\sigma \Rightarrow t_j) \Rightarrow (\sigma \Rightarrow t_i)$ MP
$\sigma \Rightarrow t_i$ MP □

Lecture 7

Definition A proof in propositional calculus is a deduction from the empty set of hypotheses

Theorem 2.9 Completeness Theorem
If $S \models t$, then $S \vdash t$

Proof Using 2.8, we may reduce to the case $t = \bot$:
for if $S \models t$ then $S \cup \{\neg t\} \models \bot$
(ie $\nexists$ a valuation making all of $S \cup \{\neg t\}$ true)
and if $S \cup \{\neg t\} \vdash \bot$ then $S \vdash \neg\neg t$ by 2.8,
whence $S \vdash t$ by axiom (T) and MP

Logic and Set Theory

We prove the contrapositive:
if S is consistent (ie $S \nvdash \bot$)
then S has a model (ie a valuation v such that $\bar{v}(s) = 1 \quad \forall s \in S$)

To do this, we first use Zorn's Lemma to enlarge S to a maximal consistent set $\bar{S}$:
the set $\mathcal{C} \subseteq \mathcal{L}(P)$ of consistent subsets of $\mathcal{L}(P)$ is chain-complete under $\subseteq$,
since if $\{ S_i \mid i \in I \}$ is a chain of consistent subsets and $\bigcup_{i \in I} S_i \vdash \bot$,
then a deduction of $\bot$ from $\bigcup_{i \in I} S_i$ would use only finitely many hypotheses $s_1, s_2, \dots, s_n$ say
and then $\exists\, i \in I$ such that $\{ s_1, s_2, \dots, s_n \} \subseteq S_i$ $\Rightarrow\Leftarrow$

Note that
i. $\bar{S}$ is deductively closed, ie $\bar{S} \vdash t$ implies $t \in \bar{S}$
since if $\bar{S} \vdash t$
then $\bar{S} \cup \{t\}$ is consistent and hence equals $\bar{S}$
ii. for every t, either $t \in \bar{S}$
or $\neg t \in \bar{S}$
since if $t \notin \bar{S}$
then $\bar{S} \cup \{t\} \vdash \bot$
and so $\bar{S} \vdash \neg t$ by 2.8

We define $v : P \to 2$
by $v(p) = \begin{cases} 1 & \text{if } p \in \bar{S} \\ 0 & \text{if } \neg p \in \bar{S} \end{cases}$

Claim The canonical extension of v to $\bar{v} : \mathcal{L}(P) \to 2$
satisfies $\bar{v}(s) = 1$ iff $s \in \bar{S}$ $(*)$

Proof True for $s \in P$ by definition
True for $\bot$, since $\bot \notin \bar{S}$
Suppose true for s and t, and consider $(s \Rightarrow t)$
Case 1: if $\bar{v}(t) = 1$, then $t \in \bar{S}$
but $t \vdash (s \Rightarrow t)$
so $(s \Rightarrow t) \in \bar{S}$
so $(*)$ holds for $(s \Rightarrow t)$
Case 2: if $\bar{v}(s) = 0$, then $\neg s \in \bar{S}$
but $\neg s \vdash (s \Rightarrow t)$ (example sheet 2, question 4)
so $(s \Rightarrow t) \in \bar{S}$
so $(*)$ holds for $(s \Rightarrow t)$
Case 3: if $\bar{v}(t) = 0$ and $\bar{v}(s) = 1$, then $s \in \bar{S}$ and $t \notin \bar{S}$
but $\{s, (s \Rightarrow t)\} \vdash t$
so $(s \Rightarrow t) \notin \bar{S}$
so $(*)$ holds for $(s \Rightarrow t)$

Since $\bar{S} \supseteq S$, this in particular implies $\bar{v}(s) = 1 \quad \forall s \in S$
ie v is a model of S $\square$

Remarks 2·10 a. We have seen that if T is a maximal consistent subset of $\mathcal{L}(P)$,
then $\exists$ a valuation $v : P \to 2$
such that $\bar{v}(t) = 1$ iff $t \in T$

Conversely, given any valuation v,
the set $T = \{ t \in \mathcal{L}(P) \mid \bar{v}(t) = 1 \}$
is deductively closed (by Soundness)
contains exactly one of $\{ t, \neg t \}$ for each t
and does not contain $\bot$
And these properties imply that T is maximal consistent.

b. If P, and hence $\mathcal{L}(P)$, is countable,
then we can enlarge S to a maximal consistent set without using Zorn's lemma, as follows:

list the formula in $\mathcal{L}(P)$ as $(t_0, t_1, t_2, \ldots)$
set $S_0 = S$
define S_n recursively by $S_{n+1} = \begin{cases} S_n \cup \{ t_n \} & \text{if this set is consistent} \\ S_n \cup \{ \neg t_n \} & \text{if } S_n \cup \{ t_n \} \vdash \bot \end{cases}$

Then $\bar{S} = \bigcup_{n \in \omega} S_n$ is consistent,
since it is the union of a chain of consistent sets,
and $\forall n$ we have either $t_n \in \bar{S}$ or $\neg t_n \in \bar{S}$,
so $\bar{S}$ is maximal consistent

c. Given a set $S \subseteq \mathcal{L}(P)$,
consider the relation $\leq_S$ on $\mathcal{L}(P)$ defined by $r \leq_S t$ if $S \vdash (r \Rightarrow t)$.
or equivalently $S \cup \{ r \} \vdash t$

This is reflexive and transitive,
and if we form the ~~quo~~ quotient of $\mathcal{L}(P)$ by $\sim_S$,
where $r \sim_S t$ means $(r \leq_S t$ and $t \leq_S r)$
then we get a partial order on the quotient set $\mathcal{B}(S)$.
This poset is a Boolean algebra:
its top and bottom elements are $[T]$ and $[\bot]$
$[r] \vee [t] = [r \vee t]$
$[r] \wedge [t] = [r \wedge t]$
the complement of $[r]$ is $[\neg r]$
Also $[T] = \{ t \in \mathcal{L}(P) \mid S \vdash t \}$,
so if S is consistent then $[T] \neq [\bot]$
and so by 1·19 $\exists$ a homomorphism $f : \mathcal{B}(S) \to 2$
such that $f([T]) = 1$
$f([\bot]) = 0$

Then the function $\bar{v} : \mathcal{L}(P) \to 2$
$\bar{v}(t) = f([t])$ corresponds to a valuation of P.

Corollary 2·11 Decidability Theorem
$\exists$ an algorithm which, given a finite set S of propositions and a proposition t,
determines whether $S \vdash t$.

Logic and Set Theory

Proof This is obvious for $\models$
(write down truth tables for the members of S and for t)
and $\models$ coincides with $\vdash$ by 2·7 and 2·9 $\square$

Corollary 2·12 Compactness Theorem
If $S \vdash t$, then $\exists$ a finite $S' \subseteq S$ such that $S' \vdash t$.
In particular, if every finite subset of S has a model
then S has a model.

Proof This is obvious for $\vdash$,
since a deduction of t from S will use only finitely many hypotheses
from S $\square$

Remark 2·13 If we make the set V of all valuations $P \to 2$ into a topological space
by taking basic open sets to be $U_t = \{ v \in V \mid \bar{v}(t) = 1 \} \quad \forall t \in \mathfrak{L}(P)$
(note that $U_s \cap U_t = U_{s \wedge t}$, so these do form a basis for a
topology),
then the assertion that $S \models \bot$ is equivalent to saying that
$\{ U_{\neg t} \mid t \in S \}$ covers V
So 2·12 is equivalent to the assertion that V is a compact space.

Lecture 8

2·14 Applications of Compactness

a. A graph G is n-colourable $\Leftrightarrow$ all its finite subgraphs are

A graph is a pair (V, E)
where V is a set of vertices
E is a set of unordered pairs $\{v, w\}$ of distinct vertices
An n-colouring of (V, E) is a partition of V into n subsets $V_1, V_2, \ldots, V_n$
such that for each $\{v, w\} \in E$,
v and w belong to different subsets

Given G, consider the set $P = \{ p_{v,i} \mid v \in V, \ 1 \leq i \leq n \}$ of primitive propositions,
and the propositional theory [set of propositional formulae] S
whose members are
$$\{ \hat{\bigvee}_{i=1}^{n} p_{v,i} \mid v \in V \} \cup \{ (p_{v,i} \Rightarrow (p_{v,j} \Rightarrow \bot)) \mid v \in V, \ i \neq j \}$$
$$\cup \{ (p_{v,i} \Rightarrow (p_{w,i} \Rightarrow \bot)) \mid \{v, w\} \in E, \ 1 \leq i \leq n \}$$
Then a model for S 'is' an n-colouring of G.
Given a finite subset $S' \subseteq S$,
$\exists$ a finite $V' \subseteq V$ such that all $p_{v,i}$ occurring in members of S'
refer to vertices $v \in V'$.
Then an n-colouring of the subgraph $G' = (V', E \cap \mathcal{P}V')$
yields a model of S'.
Given such, S has a model by compactness.

b. If any partial order on a finite set can be extended to a total order, then the same is true for arbitrary partial orders (cf example sheet 1, question 9i)

Given a partial order $(X, \leq)$,
let P be the set $\{p_{x,y} \mid (x,y) \in X^2\}$
and consider the theory

$$S = \{p_{x,y} \mid x \leq y\} \cup \{(p_{x,y} \Rightarrow (p_{y,z} \Rightarrow p_{x,z})) \mid x,y,z \in X\}$$
$$\cup \{(p_{x,y} \Rightarrow (p_{y,x} \Rightarrow \bot)) \mid x \neq y\}$$
$$\cup \{((p_{x,y} \Rightarrow \bot) \Rightarrow p_{y,x}) \mid x,y \in X\}$$

Then models for S correspond to ~~tot~~ total orderings of X which extend $\leq$
For any finite $S' \subseteq S$,
$\exists$ a finite $X' \subseteq X$ such that S' refers only to $p_{x,y}$'s with $x,y \in X'$,
and a total ordering of X' extending $(\leq \cap X' \times X')$ will yield a model for S'.

Note that the result for finite posets can be proved without using Zorn's Lemma,
by an argument like that of 2.10 b

§ 3 Predicate Calculus

To formulate a language which enables us to talk about mathematical structures,
(eg groups $G \times G \xrightarrow{m} G$, $G \xrightarrow{i} G$, $G^0 \xrightarrow{e} G$
or posets $\leq \subseteq P \times P$)
we need to include ~~po~~ both operation symbols and predicate [relation] symbols in our language.

Definition 3.1 A (first-order) signature $\Sigma = (\Omega, \pi)$
consists of a set Ω of operation symbols,
equipped with a function $\alpha : \Omega \to \mathbb{N}$ assigning to each $\omega \in \Omega$ its arity,
together with a set π of predicate symbols,
again equipped with an arity function $\alpha : \pi \to \mathbb{N}$
(↑ number of inputs expected)
An operation symbol of arity 0 is a constant
A predicate symbol of arity 0 is a primitive proposition

Given Σ, a structure for Σ is a set A equipped with
a function $\omega_A : A^{\alpha(\omega)} \to A$ for each $\omega \in \Omega$
and a subset $[\![\pi_A]\!] \subseteq A^{\alpha(\pi)}$ for each $\pi \in \pi$
(equivalently, a function $\pi_A : A^{\alpha(\pi)} \to 2 = \{0,1\}$)

Logic and Set Theory

Definition 3.2 Given a signature $\Sigma = (\Omega, \Pi)$,
the terms over Σ (or over Ω) are defined recursively as follows

i. we have an infinite supply of variables $x, x', x'', x''', \ldots$
(in practice we denote variables by $x, y, z, \ldots$
or by $x_1, x_2, x_3, \ldots$)
which are terms

ii. If $w \in \Omega$, $\alpha(w) = n$, and $t_1, t_2, \ldots, t_n$ are terms,
then $w t_1 t_2 \ldots t_n$ is a term

(ie if $A = \Omega \cup \{x, '\}$
then the set of terms is the smallest subset of A^* with these closure properties)

Given a Σ-structure A, a term t, and a list of variables $(x_1, x_2, \ldots, x_n)$
including all the variables occurring in t,
we define the interpretation of t in A as a function $A^n \xrightarrow{t_A} A$, as follows:

if t is x_i for some $i \leq n$, then t_A is a projection on the i^{th} factor

if t is $w t_1 t_2 \ldots t_m$ where $\alpha(w) = m$
then t_A is the composite $A^n \xrightarrow{((t_1)_A, (t_2)_A, \ldots (t_m)_A)} A^m \xrightarrow{w_A} A$

Definition 3.3 Given Σ, the (first-order) formulae over Σ are defined as follows

i. If $\pi \in \Pi$, $\alpha(\pi) = n$, and $t_1, t_2, \ldots, t_n$ are terms
then $\pi(t_1, t_2, \ldots, t_n)$ is a formula
ii. If s and t are terms, then $(s = t)$ is a formula
} these are called atomic formulae

iii. $\bot$ is a formula,
and if φ and ψ are formulae then $(\varphi \Rightarrow \psi)$ is a formula
(as in § 2, we introduce $\neg\varphi = (\varphi \Rightarrow \bot)$, $\top = \neg\bot$,
$(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \Leftrightarrow \psi)$)

iv. If φ is a formula and x is a free variable of φ
then $(\forall x)\, \varphi$ is a formula in which all occurrences of x are bound,
but all other free variables of φ are free

Notation $\mathcal{L}(\Sigma)$ is the set of all first-order formulae over Σ

Lecture 9

Notation $(\exists x)\, \varphi$ is shorthand for $\neg(\forall x) \neg\varphi$

Definition For each $\varphi \in \mathcal{L}(\Sigma)$ with free variables in the set $\{x_1, \ldots, x_n\}$
and each Σ-structure A,
an interpretation of φ as a subset $[\![\varphi]\!]_A \leq A^n$,
or equivalently a mapping $\varphi_A : A^n \to 2$, is as follows:

if φ is $\pi(t_1, \ldots, t_m)$ where $\pi \in \Pi$
and $\alpha(\pi) = m$
then φ_A is the composite $A^n \xrightarrow{((t_1)_A, \ldots, (t_n)_A)} A^m \xrightarrow{\pi_A} 2$

if φ is $(s = t)$
then φ_A is the composite $A^n \xrightarrow{(s_A, t_A)} A^2 \xrightarrow{\delta} 2$
where $\delta(a, b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$

if φ is $\underline{\perp}$
then φ_A is the constant 0
if φ is $(\psi \Rightarrow \chi)$
then φ_A is the composite $A^n \xrightarrow{(\psi_A, \chi_A)} 2^2 \xrightarrow{\Rightarrow_2} 2$
where $\Rightarrow_2$ is the truth table of $\Rightarrow$
ie $\Rightarrow_2 (x, y) = 0$ iff $x = 1$, $y = 0$
if φ is $(\forall x_{n+1})\psi$
we interpret ψ relative to $\{x_1, \ldots, x_{n+1}\}$ as a subset $[\![\psi_A]\!]_A^{(n+1)}$ of A^{n+1}
and then put $[\![\varphi]\!]_A^{(n)} \subseteq A^n$ to be $\{(a_1, \ldots, a_n) \mid \forall\, a_{n+1} \in A, (a_1, \ldots, a_{n+1}) \in [\![\psi_A]\!]_A^{(n+1)}\}$

Definition
Two formulae φ, ψ are α-equivalent
if we can obtain one from the other by renaming (some of) its bound variables.

If φ and ψ are α-equivalent
then they have the same interpretation in any context $\{x_1, \ldots, x_n\}$

Given a formula φ, a variable x and a term t,
we define $\varphi[t \mid x]$ to be the formula obtained from φ
by substituting a copy of t for each free occurence of x in φ,
provided no variable occuring in t is bound in φ
(otherwise first replace φ by an α-equivalent formula).

Similarly, given a finite string $\vec{x} = (x_1, \ldots, x_n)$
and a string $\vec{t} = (t_1, \ldots, t_n)$ of terms,
we can define $\varphi[\vec{t} / \vec{x}]$ as the result of simultaneously substituting t_i
$\forall$ free occurences of x_i $\forall i$

φ is satisfied in a structure A, written $A \models \varphi$,
if φ_A is the constant 1
equivalently if $[\![\varphi]\!]_A^{(n)} = A^n$

Note
$A \models \varphi$ iff $A \models (\forall x)\varphi$ for any free variables x of φ
iff $A \models (\forall \vec{x})\varphi$, where $\vec{x} = (x_1, \ldots, x_n)$
is the string of all free variables of φ
and $(\forall \vec{x})$ is shorthand for $(\forall x_1) \ldots (\forall x_n)$

A closed formula or sentence is one with no free variable
$(\forall \vec{x})\varphi$ is the universal closure

Definition 3.4
A theory over a signature Σ is a set of sentences $\mathbb{T} \subseteq \mathcal{L}(\Sigma)$
A structure A is a model of $\mathbb{T}$, $A \models \mathbb{T}$
if it satisfies all members of $\mathbb{T}$, called the axioms of the theory.

Logic and Set Theory

Examples 3.5a The group theory has signature with $\Omega = \{ m, i, e \}$ (multiplication, inverse, identity)

arities $\alpha_m = 2$
$\alpha_i = 1$
$\alpha_e = 0$

$\Pi = \emptyset$

and contains the sentences

$\forall x, y, z,$	$mx\,my\,z = m\,m\,x\,y\,z$	Associativity
$\forall x,$	$m\,e\,x = x$	Identity
$\forall x,$	$m\,i\,x\,x = e$	Inverse

b. The theory of fields, $\Omega = \{ +, \times, -, 0, 1 \}$

arities $\alpha_+ = 2$
$\alpha_\times = 2$
$\alpha_- = 1$
$\alpha_0 = 0$
$\alpha_1 = 0$

$\Pi = \emptyset$

with axioms for rings, all of which are universal closures of atomic formulae,

plus $\neg(0 = 1)$ (not the trivial ring)

$\forall x, \quad \neg(x = 0) \Rightarrow (\exists y)(xy = 1)$

c. The theory of posets has signature with $\Omega = \emptyset$

$\Pi = \{ \leq \}$

and axioms

$\forall x,$	$x \leq x$
$\forall x, y, z,$	$(x \leq y) \Rightarrow ((y \leq z) \Rightarrow (x \leq z))$
$\forall x, y,$	$(x \leq y) \Rightarrow ((y \leq x) \Rightarrow (x = y))$

For tosets add $\forall x, y, \quad \neg(x \leq y) \Rightarrow (y \leq x)$

Remark 3.6 In most textbooks on logic,

a structure for Σ is defined to be a non-empty set A equipped with ω_A, $[\![\pi_A]\!]$

and the formula $(\forall x)\varphi$ is permitted even if x does not occur in φ.

Problems with $\emptyset$ as a structure occur because every formula with free variables is satisfied in it

in particular $(x = x)$, $((x = x) \Rightarrow \bot)$ are both true.

On the other hand $\emptyset^\circ$ is a singleton, so $\bot$ is not satisfied in $\emptyset$.

So in order to make the rule modus ponens sound,

we have to introduce a restriction on it

from φ and $\varphi \Rightarrow \psi$, we infer ψ, provided either ψ has a free variable or φ does not have a free variable.

With this restriction, the formulae $(\forall x, (x = x) \Rightarrow \bot)$ and $(\forall x)\bot$ are semantically equivalent, being true only in $\emptyset$,

but they're not syntactically equivalent.

Definition 3.7 If S is a set of sentences in $\mathcal{L}(\Sigma)$ and φ is a sentence,
we say S semantically entails φ, $S \models \varphi$
if for all Σ-structures A such that $A \models S$, we also have $A \models \varphi$.

For (sets of) formulae with free variables, we say $S \models \varphi$
if, given any Σ-structure A and an assignment of values in A to the free variables in φ making φ false,
we can extend to an assignment of values to the free variables in all members of S making at least one member of S false.

Definition 3.8 The (first-order) predicate calculus has the following axioms

(K) $\varphi \Rightarrow (\psi \Rightarrow \varphi)$
(S) $(\varphi \Rightarrow (\psi \Rightarrow \chi)) \Rightarrow ((\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \chi))$
} φ, ψ, χ any formulae of $\mathcal{L}(\Sigma)$

(T) $\neg\neg\varphi \Rightarrow \varphi$
(I) $(\forall x)\varphi \Rightarrow \varphi[t/x]$ (φ any formula, with x as a free variable; t any term)
(U) $(\forall x)(\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow (\forall x)\psi)$ x free in ψ, not in φ
(R) $(\forall x)(x = x)$
(E) $(\forall x, y)((x = y) \Rightarrow (\varphi \Rightarrow \varphi[y/x]))$ x free in φ

and the following rules of inference
(MP) from φ and $(\varphi \Rightarrow \psi)$, we may infer ψ,
provided either ψ has a free variable or φ does not
(Gen) from φ we may infer $(\forall x)\varphi$,
provided ~~either~~ x appears free in φ
but not in any hypothesis used in deducing φ

By a deduction from a set S of hypotheses,
we mean a finite list $\varphi_1, \varphi_2, \ldots, \varphi_n$ of formulae,
each of which is either an axiom
a member of S
or ~~an~~ obtainable from (an) earlier formula(e)
by either (MP) or (Gen)
We say $S \vdash \varphi$ if there is a deduction from S whose last member is φ

Proposition 3.9 Soundness Theorem
Suppose either that $S \cup \{\varphi\}$ is a set of sentences
or that φ has a free variable.
If $S \vdash \varphi$, then $S \models \varphi$

Proof Like 2.7
Check that each axiom is a tautology
and that the rules of inference are sound

Logic and Set Theory

Proposition 3.10 Deduction Theorem

Suppose either that Ψ has a free variable
or that φ does not
Then $S \cup \{\varphi\} \vdash \Psi$ iff $S \vdash (\varphi \Rightarrow \Psi)$

Proof $\Leftarrow$ Like 2.8

$\Rightarrow$ Like 2.8, except that we also have to consider the case when
$\Psi = (\forall x)\chi$ is obtained by an application of (Gen)

If x is not free in φ,
then $S \vdash (\varphi \Rightarrow \chi)$ by the induction hypothesis,
so $S \vdash (\forall x)(\varphi \Rightarrow \chi)$ by (Gen)
so $S \vdash (\varphi \Rightarrow (\forall x)\chi)$ by (U) and (MP)

x is not free in φ — why can we use (Gen)?

If x is free in φ,
then the hypothesis φ cannot have been used in the deduction of χ,
so we actually have $S \vdash \chi$ and $S \vdash (\forall x)\chi = \Psi$,
so $S \vdash (\varphi \Rightarrow \Psi)$ by (K) and (MP)

Theorem 3.11 Completeness Theorem

Suppose either that $S \cup \{\varphi\}$ consists of sentences
or that φ has a free variable.
Then $S \models \varphi$ implies $S \vdash \varphi$

Proof non-examinable?

First reduce to the case of sentences, by replacing all free variables in $S \cup \{\varphi\}$ by new constants.
If the resulting hypotheses and conclusion are S' and φ',
then from $S \models \varphi$ we can deduce $S' \models \varphi'$,
and from $S' \vdash \varphi'$ we can deduce $S \vdash \varphi$

Second, reduce to the case $\varphi = \bot$ using the deduction theorem,
just as in the propositional case.
We prove the contrapositive: if S is a consistent set of sentences (ie $S \not\vdash \bot$)
then S has a model (ie $S \not\models \bot$)

[Suppose we are given the axioms for the theory R of rings.
How would we construct a model of R?
A ring must contain $0, 1, 1+1, (1+1)+1, \ldots$
$-1, (-1)+(-1), \ldots$
$1+(-1), 0\times 1, \ldots$
Consider the set C of closed terms of the language,
modulo the equivalence relation $\sim$
where $s \sim t$ iff $R \vdash (s = t)$
This set has a ring structure, with $[s] + [t] = [s+t]$
$[s] \times [t] = [s \times t]$ etc
(and in fact it is the ring of integers)
(If we had a predicate π, we would say $([s], [t]) \in [\![\pi]\!]$
iff $R \vdash \pi(s,t)$)

Now consider the theory $\mathbb{F}$ of fields:
we have the same closed terms and the same equivalence relation,
but $\mathbb{Z}$ is not a field.
In $\mathbb{F}$ we can prove $\vdash (1+1=0) \vee (\exists x)((1+1)x = 1)$
but we need to decide which of these is going to be true in our field
If we add the axiom $(1+1=0)$ to get the theory $\mathbb{F}_2$,
then the term structure collapses to $\mathbb{Z}/2\mathbb{Z}$,
which is a model of $\mathbb{F}_2$
But if we add the axioms $(\exists x)((1+1+\dots+1)x = 1)$ $\forall$ primes p,
to get the theory $\mathbb{F}_0$, we still get $\mathbb{Z}$ as our term structure.
In $\mathbb{F}_0$ we can prove $(\exists x)(x+x=1)$,
but there is no closed term t such that $(t+t=1)$ is provable.]

Lecture 11

Suppose we are given a consistent theory [set of sentences] $\mathbb{T}_0$ in a language $\mathcal{L}(\Sigma_0)$
We define increasing sequences of signatures Σ_n and theories $\mathbb{T}_n \subseteq \mathcal{L}(\Sigma_n)$ as follows:
for even n, set $\Sigma_{n+1} = \Sigma_n$
let $\mathbb{T}_{n+1}$ be a maximal consistent extension of $\mathbb{T}_n$
(constructed either as in 2.9 or in 2.10b)
for odd n, let E_n be the set of formulae $\varphi \in \mathcal{L}(\Sigma_n)$ with one free variable x,
such that $\mathbb{T}_n \vdash (\exists x)\varphi$
and set $\Sigma_{n+1} = \Sigma_n \cup \{ c_\varphi \mid \varphi \in E_n \}$ where the c_φ's are constant symbols not in Σ_n

$$\mathbb{T}_{n+1} = \mathbb{T} \cup \{ \varphi[c_\varphi / x] \mid \varphi \in E_n \}$$

We need to know $\mathbb{T}_{n+1}$ is consistent in this case:
Consider adding a single new witness.
So suppose $\mathbb{T} \vdash (\exists x)\varphi$,
c is a constant not occurring in any member of $\mathbb{T}$
and $\mathbb{T} \cup \{\varphi[c_\varphi / x]\} \vdash \bot$
By the Deduction Theorem, we have $\mathbb{T} \vdash \neg\varphi[c_\varphi / x]$
We can rewrite this deduction, replacing all c_φ's by x's, to get
$\mathbb{T} \vdash \neg\varphi$
By (Gen), $\mathbb{T} \vdash (\forall x)\neg\varphi$
but we know $\mathbb{T} \vdash \neg(\forall x)\neg\varphi$ $\quad \mathbb{T} \vdash (\exists x)\varphi$
so $\mathbb{T} \vdash \bot$
Hence by induction we can add witnesses for any finite number of existential formulae without destroying consistency.
Since a deduction of $\bot$ from $\mathbb{T}_{n+1}$ would involve only finitely many of the new axioms, $\mathbb{T}_{n+1}$ is consistent.

Now define $\Sigma_\infty = \bigcup_{n \geq 0} \Sigma_n$
$\mathbb{T}_\infty = \bigcup_{n \geq 0} \mathbb{T}_n$
Then $\mathbb{T}_\infty$ is consistent, since it is the union of a chain of consistent sets.
$\mathbb{T}_\infty$ is maximal consistent
for any $\varphi \in \mathcal{L}(\Sigma_\infty)$, we have $\varphi \in \mathcal{L}(\Sigma_{2n})$ for some n,
so either φ or $\neg\varphi$ belongs to $\mathbb{T}_{2n+1}$

Logic and Set Theory

Similarly T_ω is deductively closed.
And T_ω has witnesses
if $T_\omega \vdash (\exists x)\varphi$
then $\exists\, n$ such that $\varphi \in \mathcal{L}(\Sigma_{2n+1})$ and $T_{2n+1} \vdash (\exists x)\varphi$
so $T_{2n+2} \vdash \varphi[c_\varphi / x]$

Now suppose $T \subseteq \mathcal{L}(\Sigma)$ is a maximal consistent set of sentences, and has witnesses.
Set $C = \{$closed terms over $\Sigma\}$
and factor C by $\sim$, where $s \sim t$ iff $T \vdash (s = t)$
On the set $A = C/\sim$, we interpret the operation and predicate symbols of Σ by
$w_A([t_1], [t_2], \ldots, [t_n]) = [wt_1t_2\ldots t_n]$
$([t_1], [t_2], \ldots [t_n]) \in [\![\pi]\!]_A$ iff $T \vdash \pi(t_1, \ldots, t_n)$
(check that these are well-defined)

Claim For any formula φ with free variables $x_1, \ldots, x_n$, we have
$([t_1], \ldots, [t_n]) \in [\![\varphi]\!]_A$ iff $T \vdash \varphi[t_1, t_2, \ldots, t_n / x_1, x_2, \ldots, x_n]$ (*)

Proof This is true for atomic formulae by construction
(and for $\bot$ since T is consistent)
if $\varphi = (\psi \Rightarrow \chi)$ and (*) holds for ψ and χ,
then it holds for φ by the argument in the proof of 2.9
if $\varphi = (\forall x)\psi$ and (*) holds for φ, then
if $T \vdash \varphi[t_1, \ldots t_n / x_1 \ldots x_n]$
then by (I) & (MP) we have $T \vdash \psi[t_1 \ldots t_{n+1} / x_1 \ldots x_{n+1}]$ for any t_{n+1}
so $([t_1], \ldots, [t_{n+1}]) \in [\![\psi]\!]_A$ for any t_{n+1}
so $([t_1], \ldots, [t_n]) \in [\![\varphi]\!]_A$
Conversely, if $T \nvdash \varphi[t_1, \ldots, t_n / x_1 \ldots x_n]$
then $T \vdash \neg\varphi[t_1 \ldots t_n / x_1 \ldots x_n]$
ie $T \vdash (\exists x_{n+1}) \neg\psi[t_1 \ldots t_n / x_1 \ldots x_n]$
so $\exists\, t_{n+1}$ such that $T \vdash \neg\psi[t_1 \ldots t_{n+1} / x_1 \ldots x_{n+1}]$
so $([t_1], \ldots, [t_n]) \notin [\![\varphi]\!]_A$

In particular, for any sentence φ, we have $[\![\varphi]\!]_A = 1$ iff $T \vdash \varphi$
iff $\varphi \in T$

In particular, A is a model of T. □

Corollary 3.12 Compactness Theorem
If T is a ~~sent~~ set of sentences and any finite subset of T has a model, then T has a model

Proof This is obvious if we replace 'has a model' by 'is consistent'. □

corollary 3.13 Upward Löwenheim - Skolem Theorem
If T has an infinite model,
or if T has finite models of arbitrarily large cardinality,
then for any set I there is a T-model A such that I injects into the underlying set of A.

Proof

Add new constants $\{c_i \mid i \in I\}$ to the language,
and let $T' = T \cup \{\neg(c_i = c_j) \mid i \neq j \text{ in } I\}$
Any finite subset of T' has a model, since we can assign distinct values to the members of a finite subset of $\{c_i \mid i \in I\}$ in some T-model
So T' has a model A;
but this is just a T-model equipped with an injection $I \longrightarrow A$
$i \longmapsto (c_i)_A$ □

Corollary 3.14 Downward Löwenheim-Skolem Theorem
Suppose Σ is a countable signature
and that T is a theory in $\mathcal{L}(\Sigma)$ which has an infinite model.
Then T has a countably infinite model.

Proof

Add constants $\{c_n \mid n \in \mathbb{N}\}$ to Σ
and let $T' = T \cup \{\neg(c_m = c_n) \mid m \neq n \text{ in } \mathbb{N}\}$
Then T' has a model;
but the language of T' is still ~~cont~~ countable,
and so the construction in the proof of 3.11 produces a countable model of T', which must be countably infinite. □

In fact the Downward Löwenheim-Skolem Theorem
says that any infinite model of a first-order theory T
has a countable structure which is still a model of T
The Löwenheim-Skolem Theorems
tell us that, for any infinite structure A,
we cannot have a first order theory whose only model (up to isomorphism) is A.
Lecture 12

Peano's Postulates for $\mathbb{N}$ (1899)
1. 0 is a natural number
2. Every natural number has a successor
3. 0 is not a successor
4. Distinct natural numbers have distinct successors
5. If P is a property of natural numbers which holds for 0
and holds for the successor of n whenever it holds for n,
then P holds $\forall\ \mathbb{N}$

In modern language
Σ contains a constant 0 and a ~~unary~~ arity! operation S
then 3. $\Leftrightarrow$ $(\forall x)\neg(sx = 0)$
4. $\Leftrightarrow$ $(\forall x, y)((sx = sy) \Rightarrow (x = y))$
We can replace 5. by a scheme of axioms
$(\forall y_1, \ldots, y_n)(\varphi[0/x] \Rightarrow ((\forall x)(\varphi \Rightarrow \varphi[sx/x])) \Rightarrow (\forall x)\varphi)$
for all $\varphi \in \mathcal{L}(\Sigma)$
with $\{x, y_1, \ldots, y_n\}$ the free variables of φ.

This is strictly weaker than 5.

Logic and Set Theory

To get first order Peano arithmetic, we add binary operation symbols a (addition), m (multiplication)
and axioms $(\forall x)(axc0 = x)$
$(\forall x, y)(axcsy = saxcy)$
$(\forall x)(mxc0 = 0)$
$(\forall x, y)(mxcsy = amxcyx)$

This theory has $\mathbb{N}$ as a model, but it also has an uncountable model.

Similarly, $\mathbb{R}$ is the unique (up to isomorphism) model of the theory of conditionally complete ordered fields
(ie every non-empty bounded set has a least upper bound)
We can replace this by a scheme of axioms in the language of ordered rings, but the resulting theory has countable models (eg the field of real algebraic numbers)

Definition
A first order theory T is countably categorical if any two countable models are isomorphic.
Similarly, a first order theory T is K-categorical for any infinite cardinal K if any two models of order K are isomorphic

The theory of dense tosets without top or bottom elements is countably categorical, but not (card $\mathbb{R}$)-categorical
- consider $\mathbb{R}$ and $\{x \in \mathbb{R} : x \leq 0 \text{ or } x \in \mathbb{Q}\}$

Similarly, the theory of conditionally complete ordered fields (real closed fields) is (card $\mathbb{R}$)-categorical
but has non-isomorphic countable models (eg $\overline{\mathbb{Q}}$ and $\overline{\mathbb{Q}(\pi)}$)

Definition
Two models of a first-order theory are elementarily equivalent, if they satisfy the same sentences.

Clearly any two T-models are elementarily equivalent
$\Leftrightarrow$ T is complete, ie $\forall \varphi$ either $T \vdash \varphi$ or $T \vdash \neg\varphi$.

§4 Zermelo - Fraenkel Set Theory

How should we 'axiomatise' 'the universe of all sets'?

G. Frege (1892) proposed an axiomatisation based on
extensionality $(\forall x, y)((\forall z)((z \in x) \Leftrightarrow (z \in y)) \Rightarrow (x = y))$
and comprehension $(\exists x)(\forall y)((y \in x) \Leftrightarrow \varphi)$ where the free variable of φ is $\{y\}$

B. Russell observed this is inconsistent
for $\varphi = \neg(y \in y)$, we have $(\exists x)(\forall y)((y \in x) \Leftrightarrow \neg(y \in y))$
and either $(x \in x)$ or $\neg(x \in x)$
leading to a contradiction.

Russell's response is to work in type theory rather than set theory,
where all entities have types indexed by a natural number
and $(s = t)$ is allowed only when s, t have the same type
$(s \in t)$ is allowed only when $\text{type}(t) = \text{type}(s) + 1$
but this is much weaker than set theory with global $\in$, = predicates

W. Quine's response (New Foundations) allows comprehension only for stratifiable formulae (ie can be interpreted in type theory).

R. Holmer (2012!) has claimed that New Foundations is consistent relative to Zermelo-Fraenkel (not yet checked)

E. Zermelo ~~rep~~ (1904) replaced comprehension by the scheme of separation $(\forall u)(\exists x)((\forall y)((y \in x) \Leftrightarrow ((y \in u) \wedge \varphi)))$

Lecture 13

Definition 4.1 Zermelo Set Theory is the first-order theory over a signature with one binary predicate $\in$ and the following axioms:

1. Extensionality $(\forall x, y)((\forall z)((z \in x) \Leftrightarrow (z \in y)) \Rightarrow (x = y))$
2. Separation scheme $(\forall w_1, \ldots, w_n)(\forall x)(\exists y)(\forall z)((z \in y) \Leftrightarrow ((z \in x) \wedge \varphi))$
   for any formula φ with free variables $z, w_1, \ldots, w_n$
3. Empty set $(\exists x)(\forall y) \neg (y \in x)$
4. Pair set $(\forall x, y)(\exists z)(\forall w)((w \in z) \Leftrightarrow ((w = x) \vee (w = y)))$
5. Union set $(\forall x)(\exists y)(\forall z)((z \in y) \Leftrightarrow (\exists w)((z \in w) \wedge (w \in x)))$
6. Power set $(\forall x)(\exists y)(\forall z)((z \in y) \Leftrightarrow (\forall w)((w \in z) \Rightarrow (w \in x)))$

These axioms allow us to introduce
a constant symbol $\emptyset$
a binary operation symbol $(x, y) \mapsto \{x, y\}$
and two unary operation symbols $\cup, \wp$
We abbreviate $\{x, x\}$ to $\{x\}$
and $\cup \{x, y\}$ to $(x \cup y)$

For any $x \neq \emptyset$, we can define $\cap x$ to be $\{z \in y \mid (\forall w)(w \in x \Rightarrow z \in w)\}$ for any $y \in x$

So we can define $x \cap y = \cap \{x, y\}$
and $x \setminus y = \{z \in x \mid \neg (z \in y)\}$

Definition The (Kuratowski-Wiener) ordered pair $\langle x, y \rangle$ is $\{\{x\}, \{x, y\}\}$

Note $(\forall x, y, z, w)((\langle x, y \rangle = \langle z, w \rangle) \Leftrightarrow ((x = z) \wedge (y = w)))$

Definition
$$\text{First}(t) = \begin{cases} \cup \cap t & \text{if } t \neq \emptyset \\ \emptyset & \text{if } t = \emptyset \end{cases}$$
$$\text{Second}(t) = \begin{cases} \cup (\cup t \setminus \cap t) & \text{if } \cup t \setminus \cap t \neq \emptyset \\ \text{first}(t) & \text{otherwise} \end{cases}$$

Logic and Set Theory

't is an ordered pair' means '$t = \langle \text{first}(t), \text{second}(t) \rangle$'

We can define

$$x \times y = \{ z \in \mathcal{P}\mathcal{P}(x \cup y) \mid (z \text{ is an ordered pair}) \wedge (\text{First}(z) \in x) \wedge (\text{Second}(z) \in y) \}$$

'$w : x \to y$' means $(w \subseteq x \times y) \wedge (\forall t)(t \in x) \Rightarrow (\exists! u)((\overset{u}{\underset{y}{}} \in y) \wedge \langle t, u \rangle \in w)$

where '$(\exists! u)\varphi$' means $(\exists u)(\varphi \wedge (\forall v)(\varphi[v/u] \Rightarrow (u = v)))$

And we can define the set y^x of all functions $x \to y$ as

$$\{ w \in \mathcal{P}(x \times y) \mid w : x \to y \}$$

To complete Definition 4.1, we need to add an axiom of infinity.
We can construct infinitely many sets which are distinct

eg $\emptyset, \mathcal{P}\emptyset, \mathcal{P}\mathcal{P}\emptyset, \mathcal{P}\mathcal{P}\mathcal{P}\emptyset, \ldots$

or $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \ldots$

or $\emptyset, \emptyset^+, \emptyset^{++}, \emptyset^{+++}, \ldots$ where $x^+ = x \cup \{x\}$

7. Infinity $\quad (\exists x)((\emptyset \in x) \wedge (\forall y)((y \in x) \Rightarrow (y^+ \in x)))$

Given this, there is a unique smallest x with this property, which we denote by ω

Classes

Given a first-order formula φ with one free variable x,
we think of it as defining the class of all x's satisfying φ.
If $(\forall x)(\varphi \Leftrightarrow \psi)$ holds, then we think of φ and ψ as defining the same class.

Write M for a typical ~~set~~ class
$(t \in M)$ for $\varphi[t/x]$, where φ is a formula defining M.

M is a set if $(\exists y)(\forall x)((x \in M) \Leftrightarrow (x \in y))$
M is a proper class if $\neg(\exists y)(\forall x)((x \in M) \Leftrightarrow (x \in y))$

Similarly, we think of a formula with n free variables, viewed extensionally, as a class of n-tuples,
and write $\langle x_1, \ldots, x_n \rangle \in M$ iff φ holds,
where φ is a formula defining M.

A class of n-tuples is functional (or call it a function-class)
if $\quad (\forall x_1, \ldots, x_{n-1}, y, z)(((\langle x_1, \ldots, x_{n-1}, y \rangle \in M) \wedge (\langle x_1, \ldots, x_{n-1}, z \rangle \in M)) \Rightarrow (y = z))$

Think of such an M as defining a 'function'
from the class $D = \{ \langle x_1, \ldots, x_{n-1} \rangle \mid (\exists y)(\langle x_1, \ldots, x_{n-1}, y \rangle \in M) \}$
to the class $V = \{ x \mid (x = x) \}$, and write $M : D \to V$

Definition 4.2 Zermelo-Fraenkel set theory is obtained from Zermelo set theory by adding the axiom-scheme of Replacement and the axiom of Foundation.

Replacement: for any formula φ with free variables $x, y, w_1, \ldots, w_n$,
$(\forall w_1, \ldots, w_n)((\forall y, y')((\varphi \wedge \varphi[y'/y]) \Rightarrow (y = y')) \Rightarrow ((\forall u)(\exists v)((\forall y)(y \in v) \Leftrightarrow (\exists x)((x \in u) \wedge \varphi)))$

Sequences

$0, 1, 2, 3, \ldots$ where $0 = \emptyset$, $n+1 = n^+ = n \cup \{n\}$

$V_0, V_1, V_2, V_3, \ldots$ where $V_0 = \emptyset$, $V_{n+1} = \mathcal{P} V_n$

We can find a class of pairs whose members are precisely the pairs $\langle n, V_n \rangle$, $n \in \omega$.
So Replacement allows us to conclude that $\{V_0, V_1, V_2, \ldots\}$ is a set
and hence so is $\bigcup \{V_0, V_1, V_2, \ldots\} = V_\omega$

Now we can form $V_{\omega+1} = \mathcal{P} V_\omega$
$V_{\omega+2} = \mathcal{P} V_{\omega+1}$
$\vdots$

Does every set eventually appear as a member of some V_α?

To ensure this, we add the Axiom of Foundation (J. von Neumann 192?)
$(\forall x)((\underbrace{(\exists y)(y \in x)}_{x \text{ is not } \emptyset} \Rightarrow (\exists y)((y \in x) \wedge (\forall z)((z \in y) \Rightarrow \neg(z \in x))))$
$y \subset x, \ y \neq x$

Another sequence of sets

Starting from any set x, we can form $\bigcup x$, $\bigcup\bigcup x$, $\bigcup\bigcup\bigcup x$, ...
Replacement allows us to collect these into a set $\{x, \bigcup x, \bigcup\bigcup x, \ldots\}$
and then to form $\bigcup \{x, \bigcup x, \bigcup\bigcup x, \ldots\} = TC(x)$,
the transitive closure of x.

Definition 4.3 We say x is transitive if $(\forall y, z)(((y \in x) \wedge (z \in y)) \Rightarrow (z \in x))$

Equivalently $x \subseteq \mathcal{P} x$ or $\bigcup x \subseteq x$

Clearly, any intersection of transitive sets is transitive,
so if there is any transitive set containing x, there must be a subset one.
$TC(x)$ is the smallest transitive set containing x
for if y is transitive and $x \subseteq y$, then $\bigcup x \subseteq \bigcup y \subseteq y$
and similarly $\bigcup\bigcup x \subseteq y$, $\bigcup\bigcup\bigcup x \subseteq y$, ...
$\Rightarrow TC(x) \subseteq y$
But $\bigcup TC(x) = \bigcup \{\bigcup x, \bigcup\bigcup x, \ldots\} \subseteq TC(x)$

Theorem 4.4 In the presence of the other axioms of Zermelo-Fraenkel,
Foundation is equivalent to the scheme of $\in$-Induction:
$(\forall z_1, \ldots, z_n)((\forall x)((\forall y)((y \in x) \Rightarrow \varphi[y/x]) \Rightarrow \varphi) \Rightarrow (\forall x)\varphi)$
where φ is any formula with free variables $x, z_1, \ldots, z_n$

Logic and Set Theory

Proof Suppose $\in$-induction holds.

Consider the formula $(\forall y)((x \in y) \Rightarrow (\exists z)((z \in y) \wedge (z \cap y = \emptyset)))$

['x is a regular set']

Foundation is equivalent to the assertion $(\forall x)(x \text{ is regular})$

We can prove this by $\in$-induction. Consider x.

Suppose $(\forall y)((y \in \overset{x}{\not z}) \Rightarrow (y \text{ is regular}))$ and $x \in z$.

If $x \cap z = \emptyset$ then we are done.

otherwise $(\exists y)((y \in x) \wedge (y \in z))$

But $(y \in x)$ implies $(y \text{ regular})$

so $(y \in x)$ implies $(\exists y' \in z)(y' \cap z = \emptyset)$

Hence $(x \text{ is regular})$

Conversely, consider the no-parameter case for simplicity.

Suppose $\neg(\forall x)\varphi$

ie $(\exists x)\neg\varphi$

Consider $t = \{ y \in TC(\{x\}) \mid \neg\varphi[y/x] \}$ (a set by Separation)

? This set is non-empty, because $\varphi[x/x]$

So $(\exists y \in t)(\forall z)((z \in y) \Rightarrow \neg(z \in t))$

But $z \in y$ implies $z \in TC(\{x\})$ since $TC(\{x\})$ is transitive,

and $z \notin t$ implies $\varphi[z/x]$

So we have $(\forall z \in y)\varphi[z/x]$

but $\neg\varphi[y/x]$, contradicting the hypothesis that φ is inductive. □

Definition 4.5 Let R be a relation-class (ie a class of pairs)

let M be a class

R is well-founded on M if

$(\forall x \subseteq M)((x \neq \emptyset) \Rightarrow (\exists y \in x \cap M)((\forall z \in M)((\langle z, y \rangle \in R) \Rightarrow \neg(z \in x))$

R is local on M if

$(\forall x \in M)(\exists y)(\forall z \in M)((z \in y) \Leftrightarrow (\langle z, x \rangle \in R))$

If R is local, then for any $x \subseteq M$ we can construct the set

$\downarrow R(x)$ of all those $y \in M$ such that $\langle y, z \rangle \in R$ for some $z \in x$

and hence the set

$RC(x) = \bigcup \{ x, \downarrow R(x), \downarrow R(\downarrow R(x)), \ldots \}$

which is the smallest subset of M containing x which is R-closed,

ie such that $y, z \in M$, $\langle y, z \rangle \in R$, and $z \in RC(x)$ imply $y \in RC(x)$

Proposition 4.6 Suppose M is a class

and R is a relation-class which is well-founded and local on M.

Then

$(\forall z_1, \ldots, z_n)((\forall x \in M)((\forall y \in M)(\langle y, x \rangle \in R \Rightarrow \varphi[y/x]) \Rightarrow \varphi) \Rightarrow (\forall x \in M)\varphi)$

for any formula φ with free-variables $x, z_1, \ldots, z_n$.

Proof

As in 4·4, suppose $(\exists x \in M) \neg\varphi$
Construct $t = \{ y \in RC(\{x\}) \mid \neg\varphi[y/x] \}$
This is a non-empty subset of M,
so by well-foundedness of R
$(\exists y \in t)(\forall z \in M)(\langle z, y \rangle \in R \Rightarrow \neg(z \in t))$
Hence the induction hypothesis on φ fails at y. □

Lecture 15

Lemma 4·7 Suppose R is a class of pairs which is well-founded and local relative to a class M.
Then there is a class of pairs $\bar{R} \supseteq R \cap (M \times M)$
which is well-founded, local, and transitive on M.
[ie $(\forall x, y, z \in M)((\langle x, y \rangle \in \bar{R}) \Rightarrow (\langle y, z \rangle \in \bar{R}) \Rightarrow (\langle x, z \rangle \in \bar{R}))$]
? ∧

Proof

We already saw that, for any set $x \subseteq M$,
there is a set $RC_M(x)$ of all iterated R-predecessors of members of x lying in M
So we define $\bar{R}$ by $(\langle x, y \rangle \in \bar{R}) \Leftrightarrow ((y \in M) \wedge x \in RC_M(\{y\}))$
This is local by definition
and it is transitive since
$x \in RC_M(\{y\})$ implies $RC_M(\{x\}) \subseteq RC_M(\{y\})$
So we need to show it is well-founded.
Let $x \subseteq M$ be a non-empty set with no $\bar{R}$-minimal member.
Define $\bar{x} = x \cup \{ y \in RC_M(x) \mid (\exists z \in RC_M(\{y\}))(z \in x) \}$
Then $\bar{x}$ is non-empty
and if $\langle y, z \rangle \in R$ and $z \in \bar{x}$
then either $z \in x$, in which case it has an $\bar{R}$-predecessor in x
and hence an R-predecessor either in x
or in $\{ y \in RC_M(x) \mid (\exists z \in RC_M(\{y\}))(z \in x) \}$
or $z \in \{ y \in RC_M(x) \mid (\exists z \in RC_M(\{y\}))(z \in x) \}$
in which case it again has an $\bar{R}$-predecessor in x.
So $\bar{x}$ has no R-minimal number ⇒⇐. □

Theorem 4·8 R-Recursion Theorem
Let M be a class
let R be a well-founded local relation-class (relative to M)
let G be a class of triples which is functional and satisfies
$(\forall x, y)((x \in M) \Rightarrow (\exists! z)(\langle x, y, z \rangle \in G))$
[ie $G : M \times V \to V$
we write $G(x, y)$ for the unique z such that $\langle x, y, z \rangle \in G$]
So F is defined recursively?
Then there is a unique function class $F : M \to V$ satisfying
$(\forall x \in M)(F(x) = G(x, \{F(y) \mid (y \in M) \wedge (\langle y, x \rangle \in R)\}))$ (*)

Proof

Uniqueness
If F and F' both satisfy (*)
then we can prove
$(\forall x \in M)(F(x) = F'(x))$ by R-induction over M.

Logic and Set Theory

Existence

We consider attempts on F; ie sets f such that

$(f \text{ is a function}) \wedge (\text{dom } f \text{ is an } R\text{-closed subset of } M)$
$\wedge\ (\forall x \in \text{dom} f)(f(x) = G(x, \{f(y) \mid (y \in M) \wedge (\langle y, x\rangle \in R)\}))$

Note that if f and f' are attempts
then they agree on $\text{dom} f \cap \text{dom} f'$ by the uniqueness argument.

So $F = \{\langle x, y\rangle \mid (\exists f)(f \text{ is an attempt}) \wedge (\langle x, y\rangle \in f)\}$
is a function-class.

Suppose F is undefined at some $x \in M$
then there is an $\bar{R}$-minimal number x_0 of $\{x \in M \mid F(x) \text{ undefined}\}$
Define $f_0 = \{\langle x, y\rangle \mid (x \in RC_M(\{x_0\})) \wedge (\exists f)((f \text{ is an attempt}) \wedge (\langle x, y\rangle \in f))\}$
[Note: this is where we use replacement]
Then f_0 is an attempt, with domain $RC_M(\{x_0\})$, and we may extend it to
$f_1 = f_0 \cup \{\langle x_0, G(x_0, \{F(y) \mid (y \in M) \wedge (\langle y, x_0\rangle \in R)\})\rangle\}$
which is an attempt with $x_0 \in \text{dom} f_1$ $\Rightarrow\Leftarrow$
So F is defined on M. $\square$

Remark 4.9 If the class M in the statement of the Recursion Theorem is a set,
then the proof becomes significantly easier.
The reason is that F itself is (a set, and hence) an attempt,
so we don't have to construct f_0.
Also, we don't need Lemma 4.7 in this case:
if $\text{dom} F \neq M$, just pick x R-minimal in $M \setminus \text{dom} F$
and extend F by defining it at x.

example 4.10 The set $\{\langle m, m^+\rangle \mid m \in \omega\}$ is a well-founded and local relation on the set ω.
So we have the usual induction principle
$((\varphi[0/x] \wedge (\forall x \in \omega)(\varphi \Rightarrow \varphi[x^+/x])) \Rightarrow (\forall x \in \omega)\varphi)$
for any formula φ
But any subset of ω in V is definable by a formula
So ω is internally a model of higher-order Peano arithmetic
and externally, the members of ω within any model of Zermelo-Fraenkel
form a model of first-order Peano Arithmetic.

Definition 4.11 A relation-class R is extensional on a class M if
$(\forall x, y \in M)((\forall z \in M)((\langle z, x\rangle \in R) \Leftrightarrow (\langle z, y\rangle \in R)) \Rightarrow (x = y))$

Theorem 4.12 Mostowski's Isomorphism Theorem
Let a be a set
$r \subseteq a \times a$ an extensional, well-founded relation on a.
Then there is a unique pair (b, f)
where b is a transitive set
$f : a \to b$ is a bijection
and $(\forall x, y \in a)(\langle x, y\rangle \in r \Leftrightarrow (f(x) \in f(y)))$

Proof

Uniqueness

Suppose (b', f') also satisfy the conditions.

Let g be the composite $b \xrightarrow{f^{-1}} a \xrightarrow{f'} b'$

Then $(\forall x, y \in b)((x \in y) \Leftrightarrow (g(x) \in g(y)))$

so $(\forall x \in b)(g(x) = x)$ by $\in$-induction over b.

So $b' = b$, and $f' = f$.

Lecture 16

Existence

We define f by r-recursion over a

$$f(x) = \{ f(y) \mid (y \in a) \wedge (\langle y, x \rangle \in r) \}$$

(ie we take $G(x, y) = y$ in the statement of the Recursion Theorem)

and we define $b = \{ f(x) \mid x \in a \}$ (which is a set by Replacement)

Clearly, f is surjective and $((\langle x, y \rangle \in r) \Rightarrow (f(x) \in f(y)))$

To show $((f(x) \in f(y)) \Rightarrow (\langle x, y \rangle \in r))$, we need to show f is injective.

Consider the formula φ with one free variable x:

$$(\forall y \in a)((f(x) = f(y)) \Rightarrow (x = y))$$

We prove $(\forall x \in a)\varphi$ by r-induction

Assume $(\forall z \in a)((\langle z, x \rangle \in r) \Rightarrow \varphi[z/x])$ and $f(x) = f(y)$

Then $(\forall z \in a)((\langle z, x \rangle \in r) \Rightarrow (\exists t \in a)((\langle t, y \rangle \in r) \wedge (f(z) = f(t))))$

But for any such z we have $\varphi[z/x]$, so we can deduce $(z = t)$

ie $(\forall z \in a)((\langle z, x \rangle \in r) \Rightarrow (\langle z, y \rangle \in r))$

Similarly $(\forall t \in a)((\langle t, y \rangle \in r) \Rightarrow (\exists z \in a)((\langle z, x \rangle \in r) \wedge (f(t) = f(z))))$

and again we have $\varphi[z/x]$, so $(t = z)$

Hence by extensionality of r we have $(x = y)$ □

Definition 4.13 A binary relation $r \subseteq a \times a$ is trichotomous if

$$(\forall x, y \in a)((\langle x, y \rangle \in r) \vee (\langle y, x \rangle \in r) \vee (x = y))$$

Note that a well-founded relation is necessarily irreflexive

(ie $(\forall x) \neg (\langle x, x \rangle \in r)$,

since $(\langle x, x \rangle \in r)$ would imply that $\{x\}$ has no r-minimal number)

And a well-founded trichotomous relation is necessarily transitive,

since if we have $(\langle x, y \rangle \in r) \wedge (\langle y, z \rangle \in r)$ but not $(\langle x, z \rangle \in r)$,

then we have either $(\langle z, x \rangle \in r)$ or $(x = z)$,

and in either case $\{x, y, z\}$ has no r-minimal number.

In this case, we will normally write $(\langle x, y \rangle \in r)$ as $(x < y)$

and think of $((x < y) \vee (x = y))$ as a total ordering on a.

We call such a relation $<$ a well-ordering of a.

Equivalently, a well-ordering is a (strict) total ordering in which

every non-empty $b \subseteq a$ has a least member

(ie $(\exists x \in b)(\forall y \in b)((x < y) \vee (x = y))$)

Clearly a well-ordering is extensional, since if $x \neq y$ then one of x, y is a predecessor of the other but not of itself. Hence ...

Logic and Set Theory

Corollary 4.14 For any well-ordered set $(a, <)$, there is a unique transitive set b such that $\in \cap (b \times b)$ is trichotomous, together with an isomorphism of ordered sets $(a, <) \xrightarrow{f} (b, \in \cap (b \times b))$

§ 5 Ordinals

Definition 5.1 An ordinal is a transitive set α such that
$(\forall x, y \in \alpha)((x \in y) \vee (y \in x) \vee (x = y))$

examples $\emptyset$ is an ordinal
so is $\{\emptyset\} = 1$, $\{\emptyset, \{\emptyset\}\} = 2$, $\{0, 1, 2\} = 3$, ...

Lemma 5.2 If α is an ordinal, then so is $\alpha^+ = \alpha \cup \{\alpha\}$

Proof Transitivity: if $x \in y$, $y \in x^+$
then either $y \in \alpha$, in which case $x \in \alpha \subseteq \alpha^+$ by transitivity of α
or $y = \alpha$, in which case $x \in \alpha \subseteq \alpha^+$

Trichotomy: if $x, y \in \alpha^+$
then either $x, y \in \alpha$, in which case $(x \in y) \vee (y \in x) \vee (x = y)$ by trichotomy of α
or $x \in \alpha$ and $y = \alpha$ in which case $(x \in y)$
or $x = \alpha$ and $y \in \alpha$ in which case $(y \in x)$
or $x = \alpha$ and $y = \alpha$ in which case $(x = y)$ □

Lemma 5.3 Every member of an ordinal is an ordinal

Proof Suppose α is an ordinal, $x \in \alpha$

Transitivity: if $y \in z \in x$, then $y, z \in \alpha$ by transitivity of α
Hence we have $(y \in x) \vee (x \in y) \vee (x = y)$ by trichotomy of α
but $(x \in y) \vee (x = y)$ would contradict Foundation, so $(y \in x)$

Trichotomy: if $y, z \in x$, then $y, z \in \alpha$ by transitivity of α,
so $(y \in z) \vee (z \in y) \vee (y = z)$ by trichotomy of α □

Notation $\underline{On}$ is the class of all ordinals (so 5.3 says $\underline{On}$ is a transitive class)

Lemma 5.4 If $\alpha, \beta \in \underline{On}$, then either $\alpha \subseteq \beta$
or $\beta \subseteq \alpha$

Proof Suppose $\alpha \not\subseteq \beta$.
Then $\alpha \setminus \beta$ is non-empty, so it has an $\in$-least member γ, say.

Now if $\delta \in \gamma$, then $\delta \in \alpha$ by transitivity, and $\delta \notin \alpha \setminus \beta$
so $\delta \in \beta$

But if $\delta \in \alpha \cap \beta$, then $(\delta \in \gamma) \vee (\gamma \in \delta) \vee (\gamma = \delta)$ by trichotomy of α
and either $(\gamma \in \delta)$ or $(\gamma = \delta)$ would imply $(\gamma \in \beta)$ $\Rightarrow\Leftarrow$
so $(\delta \in \gamma)$

? Hence by extensionality $\gamma = \alpha \cap \beta$
and in particular $\alpha \cap \beta \in \alpha$
Similarly if $\beta \notin \alpha$, then $\alpha \cap \beta \in \beta$.
So if $\alpha \notin \beta$ and $\beta \notin \alpha$,
? then we have $\alpha \cap \beta \in \alpha \cap \beta$, contradicting Foundation. $\Rightarrow\Leftarrow$ □

Corollary 5.5 i. For ordinals α, β, we have one of $(\alpha \in \beta)$, $(\beta \in \alpha)$ or $(\alpha = \beta)$.
ii. For ordinals α, β, $(\alpha \subseteq \beta)$ is equivalent to $(\alpha \in \beta) \vee (\alpha = \beta)$

Proof i. By 5.4, we have one of $\alpha = \beta$, $\alpha \subsetneq \beta$, or $\beta \subsetneq \alpha$
and $\alpha \subsetneq \beta$ implies $\alpha = \alpha \cap \beta \in \beta$ by the proof of 5.4.

ii. is similar
(note that $(\alpha \in \beta)$ implies $(\alpha \subseteq \beta)$ since β is transitive). □

Corollary 5.6 Burali-Forti paradox
On is a proper class

Proof If On were a set
then we would have $On \in On$ by 5.3 and 5.5,
contradicting Foundation. □

Lecture 17

Notation From now on, we may write $\begin{cases} \alpha < \beta & \text{for } \alpha \in \beta \\ \alpha \leq \beta & \text{for } \alpha \subseteq \beta \end{cases}$ if $\alpha, \beta \in On$

Lemma 5.7 If α is a subset of On, then $\bigcup \alpha \in On$

Proof $\bigcup \alpha$ is transitive, since it is a union of transitive sets.
The members of $\bigcup \alpha$ are ordinals by 5.3,
so by 5.5 they satisfy trichotomy. □

Theorem 5.8 Let M be any class satisfying
$(\forall x)((x \in M) \Rightarrow (x^+ \in M))$
and $(\forall x)((x \subseteq M) \Rightarrow (\bigcup x \in M))$
Then $On \subseteq M$

Proof By $\in$-induction
Suppose $(\alpha \in On)$
and $(\forall \beta \in \alpha)(\beta \in M)$

Logic and Set Theory

If α has an $\in$-greatest member, β, say,
then $(\forall \gamma)((\gamma \in \alpha) \Rightarrow ((\gamma \in \beta) \vee (\gamma = \beta)))$
But we also have $(\forall \gamma)(((\gamma \in \beta) \vee (\gamma = \beta)) \Rightarrow (\gamma \in \alpha))$ by transitivity of α.
So $\alpha = \beta^+$ by extensionality
but $\beta \in M$, so $\alpha \in M$.
If α has no greatest member
then $(\forall \beta)((\beta \in \alpha) \Rightarrow (\exists \gamma \in \alpha)(\beta \in \gamma))$,
ie $\alpha \subseteq \bigcup \alpha$
but we also have $\bigcup \alpha \subseteq \alpha$, since α is transitive.
Hence $\alpha = \bigcup \alpha$
? it is $\Rightarrow$, not $\Leftrightarrow$
but $\alpha \subseteq M$ by the hypothesis
so $\alpha \in M$ □

Definition If $\alpha = \beta^+$ for some β,
then α is a successor ordinal.
Otherwise, α is a limit ordinal

Note 0 is a limit

5.8 says that we can prove things by $\in$-induction over On,
or define them by $\in$-recursion,
by considering separately the cases of successor and limit ordinals.

For example, we define the function-class $(\alpha \mapsto V_\alpha) : On \to V$ by $\in$-recursion:
(If $\alpha = 0$, $V_\alpha = \emptyset$)
If $\alpha = \beta^+$, $V_\alpha = \mathcal{P}(V_\beta)$
If α is a limit, $V_\alpha = \bigcup \{ V_\beta \mid \beta < \alpha \}$
(We could combine the two classes into one by setting
$V_\alpha = \bigcup \{ \mathcal{P} V_\beta \mid \beta < \alpha \}$)

The sets V_α are called the von Neumann hierarchy of sets.
We define the function-class rank : $V \to On$ by $\in$-recursion.

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

Note We can prove $(\forall x)(\text{rank}(x) \in On)$ by $\in$-induction
and $(\forall \alpha \in On)(\text{rank}(\alpha) = \alpha)$

Theorem 5.9 For all sets x and ordinals α, we have
$(x \in V_\alpha) \Leftrightarrow (\text{rank}(x) < \alpha)$
$(x \subseteq V_\alpha) \Leftrightarrow (\text{rank}(x) \leq \alpha)$

Proof The second assertion follows from the first,
since $x \subseteq V_\alpha \Leftrightarrow x \in V_{\alpha^+}$
and $\text{rank}(x) \leq \alpha \Leftrightarrow \text{rank}(x) < \alpha^+$

For the first assertion

⇒ We use $\in$-induction on α

Suppose $(\forall x)(\forall \beta < \alpha)((x \in V_\beta) \Rightarrow \text{rank}(x) < \beta)$
and $(x \in V_\alpha)$

If α is a limit, then $(x \in V_\beta)$ for some $\beta < \alpha$
so $\text{rank}(x) < \beta < \alpha$

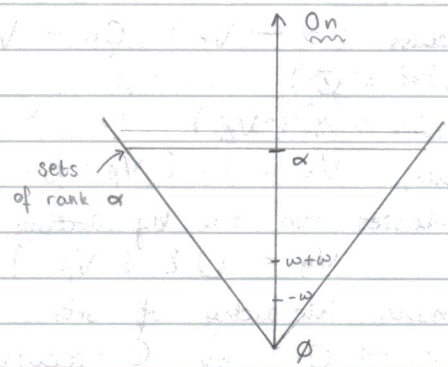
If $\alpha = \beta^+$ is a successor, then $(\forall y \in x)(y \in V_\beta)$
so $(\forall y \in x)(\text{rank}(y) < \beta)$
so $(\forall y \in x)(\text{rank}(y)^+ \leq \beta)$
so $\text{rank}(x) \leq \beta < \alpha$

⇐ We use $\in$-induction on x

Suppose $(\forall y \in x)(\forall \alpha \in On)((\text{rank}(y) < \alpha) \Rightarrow (y \in V_\alpha))$
and $(\text{rank}(x) \leq \beta < \alpha)$

Then $(\forall y \in x)(\text{rank}(y) < \beta)$
so $(\forall y \in x)(y \in V_\beta)$
ie $x \leq V_\beta$
so $x \in V_\beta{}^+ \leq V_\alpha$ □

The set-theorist's picture of the universe



Proposition 5·10 $(\alpha \in On)$ is equivalent to (α is a transitive set, all of whose members are transitive)

Proof ⇒ Immediate from 5·3

⇐ Proved by $\in$-induction:
Suppose x is a transitive set, all of whose members are transitive
and suppose $(\forall y \in x)((y \text{ is transitive}) \Rightarrow y \in On)$
Then $(\forall y \in x)(y \in On)$ since x ~~is transitive~~ all members of x are transitive
So x is a transitive set of ordinals
But then either $x = \beta^+$ for some $\beta \in x$
or $x = \bigcup x$
and so in either case x is an ordinal. □

Logic and Set Theory

To define ordered addition and multiplication synthetically,
we take $\alpha + \beta$ to be the order-type of the well-ordered set

$$\alpha \amalg \beta = \alpha \times \{0\} \cup \beta \times \{1\}$$

ordered by $\langle \gamma, i \rangle < \langle \delta, j \rangle \iff$ either $i < j$
or $i = j$ and $\gamma < \delta$

(easy to check this is a well-ordering)
and $\alpha . \beta$ to be the order-type of $\alpha \times \beta$,
ordered by the reverse lexicographical ordering
ie $\langle \gamma, \delta \rangle < \langle \gamma', \delta' \rangle \iff$ either $\delta < \delta'$
or $\delta = \delta'$, $\gamma < \gamma'$

Lemma 5.11 i. ordinal addition satisfies the recursive definition

$$\alpha + 0 = \alpha$$
$$\alpha + (\beta^+) = (\alpha + \beta)^+$$
$$\alpha + \lambda = \bigcup \{ \alpha + \gamma \mid \gamma < \lambda \} \quad \text{if } \lambda \text{ is a non-zero limit}$$

ii. ordinal multiplication satisfies the recursive definition

$$\alpha . 0 = 0$$
$$\alpha . (\beta^+) = \alpha\beta + \alpha$$
$$\alpha . \lambda = \bigcup \{ \alpha . \gamma \mid \gamma < \lambda \} \quad \text{if } \lambda \text{ is a non-zero limit}$$

Proof i. $\alpha + 0 = \text{otp}(\alpha \times \{0\})$ otp = order type
$= \alpha$

$$\alpha + (\beta^+) = \text{otp}(\alpha \times \{0\} \cup \beta \times \{1\} \cup \{\langle \beta, 1 \rangle\})$$
$$= \text{otp}(\alpha \times \{0\} \cup \beta \times \{1\})^+$$
$$= (\alpha + \beta)^+$$
$$\alpha + \lambda = \text{otp}(\bigcup \{ \alpha + \{0\} \cup \gamma \times \{1\} \mid \gamma < \lambda \})$$
$$= \text{otp}(\bigcup \{ \alpha + \gamma \mid \gamma < \lambda \})$$
$$= \bigcup \{ \alpha + \gamma \mid \gamma < \lambda \}$$

Lecture 18

ii. $\alpha \times 0 = \emptyset$
so $\text{otp}\{\alpha \times 0\} = 0$
$\alpha \times (\beta^+) = \alpha \times \beta \cup \alpha \times \{\beta\}$
so $\text{otp}(\alpha \times (\beta^+)) = \text{otp}((\alpha \times \beta) \amalg \alpha)$
$= \text{otp}(\alpha . \beta \amalg \alpha)$
$= \alpha . \beta + \alpha$

For a limit λ, we have $\alpha \times \lambda = \bigcup \{ \alpha \times \gamma \mid \gamma < \lambda \}$
By induction, we have order-isomorphisms $\alpha \times \gamma \xrightarrow{f_\gamma} \alpha . \gamma$ for each $\gamma < \lambda$
and if $\gamma < \delta$ then $\alpha \times \gamma$ is an initial segment of $\alpha \times \delta$,
so f_γ and f_δ agree where both are defined
So the f_γ can be patched together to produce an order-isomorphism
$\alpha \times \lambda \longrightarrow \bigcup \{ \alpha . \gamma \mid \gamma < \lambda \}$ □

Lemma 5·12

i. if $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$
ii. if $\alpha \leq \beta$, then $\alpha + \gamma \leq \beta + \gamma$
iii. if $\alpha > 0$ and $\beta < \gamma$, then $\alpha \cdot \beta < \alpha \cdot \gamma$
iv. if $\alpha \leq \beta$, then $\alpha \cdot \gamma \leq \beta \cdot \gamma$

Proof

i. & iii. We use the synthetic approach.

If $\beta < \gamma$, then $\alpha \amalg \beta$ is a proper initial segment of $\alpha \amalg \gamma$.
So $otp(\alpha \amalg \beta) < otp(\alpha \amalg \gamma)$

Similarly, if $\alpha \neq 0$, then $\alpha \times \beta$ is a proper initial segment of $\alpha \times \gamma$,
so $otp(\alpha \times \beta) < otp(\alpha \times \gamma)$

ii & iv. We use induction on γ.

If $\gamma = 0$, then $\alpha + \gamma = \alpha$
$\leq \beta$
$= \beta + \gamma$

If $\gamma = \delta^+$, then $\alpha + \delta \leq \beta + \delta$ by the induction hypothesis,
so $\alpha + \gamma = (\alpha + \delta)^+$
$\leq (\beta + \delta)^+$
$= \beta + \gamma$

If γ is a limit, then by the induction hypothesis we have $\alpha + \delta \leq \beta + \delta$ for all $\delta < \gamma$
so $\alpha + \gamma = \cup \{ \alpha + \delta \mid \delta < \gamma \}$
$\leq \cup \{ \beta + \delta \mid \delta < \gamma \}$
$= \beta + \gamma$

Similarly, if $\gamma = 0$, then $\alpha \cdot 0 = 0$
$= \beta \cdot 0$

If $\gamma = \delta^+$, then $\alpha \cdot \delta \leq \beta \cdot \delta$ by the induction hypothesis
and $\alpha \leq \beta$
so $\alpha \cdot \gamma = \alpha \cdot \delta + \alpha$
$\leq \beta \cdot \delta + \beta$ by i and ii
$= \beta \cdot \gamma$

If γ is a limit, then $\alpha \cdot \delta \leq \beta \cdot \delta$ for all $\delta < \gamma$ by the induction hypothesis,
so $\alpha \cdot \gamma = \cup \{ \alpha \cdot \delta \mid \delta < \gamma \}$
$\leq \cup \{ \beta \cdot \delta \mid \delta < \gamma \}$
$= \beta \cdot \gamma$

Lemma 5·13

i. $0 + \alpha = \alpha$ and $0 \cdot \alpha = 0$ for all $\alpha \in On$
ii. $1 \cdot \alpha = \alpha = \alpha \cdot 1$ for all $\alpha \in On$
iii. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for all $\alpha, \beta, \gamma \in On$
iv. $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$ for all $\alpha, \beta, \gamma \in On$
v. $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$ for all $\alpha, \beta, \gamma \in On$

i & ii are obvious synthetically (and can also be proved from the recursive definitions)
iii, iv, v can be proved either inductively or synthetically

Logic and Set Theory

iv. We use the synthetic approach

$\alpha . (\beta + \gamma) = \text{otp}(\alpha \times (\beta + \gamma))$
$= \text{otp}(\alpha \times (\beta \amalg \gamma))$

But $\alpha \times (\beta \amalg \gamma)$ contains elements of the form $\langle \delta, \langle \varepsilon, i \rangle \rangle$
with $\delta < \alpha$, $\varepsilon < (\beta \cup \gamma)$ and $i \in \{0, 1\}$

and $(\alpha \times \beta) \amalg (\alpha \times \gamma)$ contains elements of the form $\langle \langle \delta, \varepsilon \rangle, i \rangle$
with $\delta < \alpha$, $\varepsilon < (\beta \cup \gamma)$ and $i \in \{0, 1\}$

The obvious bijection $\langle \delta, \langle \varepsilon, i \rangle \rangle \longmapsto \langle \langle \delta, \varepsilon \rangle, i \rangle$
between these two sets is an order isomorphism,

so $\text{otp}(\alpha \times (\beta \amalg \gamma)) = \text{otp}((\alpha \times \beta) \amalg (\alpha \times \gamma))$
$= \text{otp}(\alpha \times \beta) + \text{otp}(\alpha \times \gamma)$
$= \alpha . \beta + \alpha . \gamma$

v. We use induction on γ

For $\gamma = 0$, we have $\alpha . (\beta . 0) = \alpha . 0$
$= 0$
$= (\alpha . \beta) . 0$

If $\gamma = \delta^+$, we have $(\alpha . \beta) . \gamma = (\alpha . \beta) . \delta + (\alpha . \beta)$ by the induction hypothesis
$= \alpha . (\beta . \delta) + \alpha . \beta$
$= \alpha . (\beta . \delta + \beta)$ by iv.
$= \alpha . (\beta . \gamma)$

If γ is a limit, then $(\alpha . \beta) . \gamma = \bigcup \{ (\alpha . \beta) . \delta \mid \delta < \gamma \}$
$= \bigcup \{ \alpha . (\beta . \delta) \mid \delta < \gamma \}$ by the induction hypothesis

but $\beta . \gamma = \bigcup \{ \beta . \delta \mid \delta < \gamma \}$
so for every $\varepsilon < \beta . \gamma$ we have $\varepsilon < \beta . \delta$ for some $\delta < \gamma$
and hence $\alpha . \bigcup \{ \beta . \delta \mid \delta < \gamma \} = \bigcup \{ \alpha . \varepsilon \mid \varepsilon < \beta . \gamma \}$
$= \alpha . (\beta . \gamma)$ □

Note

$\omega + 1 = \omega^+$
$> \omega$

but $1 + \omega = \bigcup \{ 1 + n \mid n < \omega \}$
$= \omega$

and $\omega . 2 = \omega + \omega$
$> \omega$

but $2 . \omega = \bigcup \{ 2 . n \mid n < \omega \}$
$= \omega$

also $(1 + 1) . \omega = \omega$
$\neq 1 . \omega + 1 . \omega$

Lemma 5.14 Division Algorithm

If $\alpha, \beta \in On$ and $\beta \neq 0$
then $\exists$ a unique $\gamma, \delta \in On$ with $\delta < \beta$
and $\alpha = \beta . \gamma + \delta$

Proof

First, since $\beta \geq 1$, we have $\beta . \mathcal{E} \geq \mathcal{E}$

so $\exists\ \mathcal{E}$ such that $\beta . \mathcal{E} > \alpha$

Hence there is a least such $\mathcal{E}$:

this must be a successor, since for a limit $\mathcal{E}$ we have $\beta . \mathcal{E} = \bigcup \{ \beta . \gamma \mid \gamma < \mathcal{E} \}$

Say $\mathcal{E} = \gamma^+$: then $\beta . \gamma \leq \alpha$

$$< \beta . \gamma^+$$
$$= \beta . \gamma + \beta$$

Set $\delta = \text{otp}(\alpha \setminus \beta . \gamma)$

then $\text{otp}(\beta . \gamma \amalg \delta) = \alpha$

ie $\alpha = \beta . \gamma + \delta$

and $\beta \geq \delta$ since $\beta . \gamma + \beta > \alpha$

$$= \beta . \gamma + \delta$$

Conversely, if $\alpha = \beta . \gamma + \delta$ and $\delta < \beta$

then $\beta . \gamma \leq \alpha$

$$< \beta . \gamma^+$$

so γ^+ is the least $\mathcal{E}$ such that $\beta . \mathcal{E} > \alpha$

and then δ is uniquely determined as $\text{otp}(\alpha \setminus \beta . \gamma)$ □

Can we define exponentiation α^β for ordinals α, β?

Consider the set $[\beta, \alpha]$ of all functions $\beta \to \alpha$:

is this well-orderable?

Not obviously: the lexicographic ordering $f < g$ if $f(\gamma) < g(\gamma)$ for the least γ such that $f(\gamma) \neq g(\gamma)$ is a total ordering, but not a well-ordering if $\beta \geq \omega$ and $\alpha \geq 2$:

define $f_n(\gamma) = \begin{cases} 1 & \text{if } \gamma = n \\ 0 & \text{otherwise} \end{cases}$

the set $\{ f_n \mid n \in \omega \}$ has no least member,

since $n < m \Rightarrow f_n > f_m$

Instead, we consider the set $[\beta, \alpha]_f$ of functions of finite support,

where the support of $f : \beta \to \alpha$ is $\{ \gamma < \beta \mid f(\gamma) \neq 0 \}$.

This is still not well-ordered by lexicographic ordering,

but it is well-ordered by reverse lexicographic ordering

ie $f < g \Leftrightarrow f(\gamma) < g(\gamma)$ for the largest γ such that $f(\gamma) \neq g(\gamma)$

(note that $\{ \gamma \mid f(\gamma) \neq g(\gamma) \}$ is finite,

so has a largest member if it is non-empty).

Lecture 19

$$[\beta, \alpha]_f = \{ f : \beta \to \alpha \mid \{ \gamma \overset{<}{\in} \beta \mid f(\gamma) \neq 0 \} \text{ is finite} \}$$

Lemma 5.15

i. $[\beta, \alpha]_f$ is well-ordered by reverse lexicographic ordering

ii. $\alpha^\beta = \text{otp}([\beta, \alpha]_f)$ satisfies the recursive definition

$$\alpha^0 = 1$$
$$\alpha^{(\beta^+)} = \alpha^\beta . \alpha$$
$$\alpha^\lambda = \bigcup \{ \alpha^\gamma \mid 0 < \gamma < \lambda \} \quad \text{if } \lambda \text{ is a non-zero limit}$$

Proof i. by induction on β.
Let $S \subseteq [\beta, \alpha]_f$ be a nonempty subset.
Pick $f \in S$.
If f is identically 0 then it is the least element of S.
Otherwise, there is a largest γ such that $f(\gamma) \neq 0$;
let $S_1 = \{ g \in S \mid g(\gamma') = 0 \quad \forall \gamma' > \gamma \}$
Then $S_1 \neq \emptyset$ since $f \in S_1$,
and it is an initial segment of S
so if it has a least member then that will be the least member of S.
Now consider $\{ g(\gamma) \mid g \in S_1 \}$
This is a nonempty subset of α, so has a least member δ, say.
Set $S_2 = \{ g \in S_1 \mid g(\gamma) = \delta \}$
Again, $S_2 \neq \emptyset$ and it is an initial segment of S_1.
Now the mapping $g \longmapsto g \restriction \gamma$ is an order-isomorphism
from S_2 to a (nonempty) subset of $[\gamma, \alpha]_f$,
so by the induction hypothesis S_2 has a least element,
which is the least element of S.

ii. There is a unique function $\emptyset \to \alpha$, and it has finite support.
So $\alpha^0 = \mathrm{otp}([\emptyset, \alpha]_f)$
$= 1$

We have an order-isomorphism $[\beta^+, \alpha]_f \xrightarrow{\sim} [\beta, \alpha]_f \times \alpha$
given by $(f \longmapsto \langle f \restriction \beta, f(\beta) \rangle)$.
So $\alpha^{(\beta^+)} = \mathrm{otp}([\beta^+, \alpha]_f)$
$= \mathrm{otp}([\beta, \alpha]_f \times \alpha)$
$= \alpha^\beta \cdot \alpha$

For a non-zero limit λ,
any function $\lambda \to \alpha$ of finite support has support $\subseteq \gamma$ for some $\gamma < \lambda$
So $[\lambda, \alpha]_f = \bigcup \{ S_\gamma \mid 0 < \gamma < \lambda \}$
where S_γ is order-isomorphic to $[\gamma, \alpha]_f$
and the S_γ are all initial segments of $[\lambda, \alpha]_f$
So $\alpha^\lambda = \bigcup \{ \alpha^\gamma \mid \gamma < \lambda \}$ by patching together the Mostowski isomorphism of the subsets S_γ. □

Lemma 5.16 Ordinal exponentiation satisfies the identities $\alpha^{(\beta + \gamma)} = \alpha^\beta \cdot \alpha^\gamma$
and $\alpha^{(\beta \cdot \gamma)} = (\alpha^\beta)^\gamma$

Proof by induction for the first
$\alpha^{(\beta + 0)} = \alpha^\beta$
$= \alpha^\beta \cdot 1$
$= \alpha^\beta \cdot \alpha^0$

$$\begin{aligned}
\alpha^{(\beta+\gamma^+)} &= \alpha^{(\beta+\gamma)^+} \\
&= \alpha^{(\beta+\gamma)} \cdot \alpha \\
&= (\alpha^\beta \cdot \alpha^\gamma) \cdot \alpha && \text{by the Induction hypothesis} \\
&= \alpha^\beta \cdot (\alpha^\gamma \cdot \alpha) && \text{by associativity of } \cdot \\
&= \alpha^\beta \cdot \alpha^{\gamma^+}
\end{aligned}$$

For λ a non-zero limit,

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

synthetically for the second

There is a bijection $[\gamma, [\beta, \alpha]] \to [\beta \times \alpha, \alpha]$

$$f \longmapsto (\langle \delta, \varepsilon \rangle \mapsto f(\varepsilon)(\delta)) = \hat{f}$$

which restricts to a bijection

$$[\gamma, [\beta, \alpha]_f]_f \to [\beta \times \gamma, \alpha]_f$$

since $\hat{f}$ has finite support $\Leftrightarrow$ only finitely many $f(\varepsilon)$ are not identically 0, and all $f(\varepsilon)$ have finite support

And this bijection is an order isomorphism when both sets are ordered by reverse lexicographic ordering,

so $(\alpha^\beta)^\gamma = \alpha^{(\beta \cdot \gamma)}$ □

N We do not have $(\alpha \cdot \beta)^\gamma = \alpha^\gamma \cdot \beta^\gamma$ in general

eg $$\begin{aligned}
(\omega \cdot 2)^2 &= \omega \cdot (2 \cdot \omega) \cdot 2 \\
&= \omega \cdot \omega \cdot 2 \\
&= \omega^2 \cdot 2 \\
&\neq \omega^2 \cdot 2^2
\end{aligned}$$

also $$\begin{aligned}
(2 \cdot 2)^\omega &= 4^\omega \\
&= \cup\{4^n \mid 0<n<\omega\} \\
&= \omega
\end{aligned}$$

but $$\begin{aligned}
2^\omega \cdot 2^\omega &= \omega \cdot \omega \\
&= \omega^2 \\
&> \omega
\end{aligned}$$

§6 Choice, Well-Ordering, and Cardinal Arithmetic

Among the cardinals, we have

$\omega,\ \omega+1,\ \omega+2,\ \ldots,\ \omega+\omega = \omega\cdot 2$
$\omega\cdot 3,\ \omega\cdot 4,\ \omega\cdot 5,\ \ldots,\ \omega\cdot\omega = \omega^2$
$\omega^3,\ \omega^4,\ \omega^5,\ \ldots,\ \omega^\omega$
$(\omega^\omega)^\omega = \omega^{\omega^2},\ \omega^{\omega^3},\ \ldots\ldots,\ \omega^{(\omega^\omega)}$
$(\omega^{\omega^{(\omega^\omega)}}),\ \ldots\ldots\ldots\ldots$ $\{$the least α such that $\alpha = \omega^\alpha\} = \varepsilon_0$
ε_1 = the 2nd least α such that $\alpha = \omega^\alpha$, ε_2, $\ldots$, ε_ω, $\ldots$
ζ_0 = the least α such that $\alpha = \varepsilon_\alpha$, $\ldots\ldots$

All these ordinals are countable.
How do we get uncountable ordinals?

$\omega_1 = \{\alpha \in \underline{On} \mid \alpha$ is countable$\}$ is an initial segment of $\underline{On}$
and it is a set by Replacement,
since Mostowski's Theorem gives us a surjective mapping
from the set of well-orderings of subsets of ω to ω_1.
Hence ω_1 is an ordinal,
and it is the first uncountable ordinal.

Similarly $\omega_2 = \{\alpha \in \underline{On} \mid \alpha$ can be injected into $\omega_1\}$ is an ordinal,
and so on.

More generally,

Lemma 6.1 **Hartogs' Lemma**
For any set a, there is an ordinal which cannot be mapped injectively to a.

Proof Consider the set $S = \{r \subseteq a \times a \mid r$ is a well-ordering of a subset of $a\}$
Mostowski yields a surjective function-class from S
to the class of ordinals which can be injected into a,
so the latter is a set $\gamma(a)$,
and it is an ordinal since $\alpha \leq \beta \in \gamma(a)$ implies $\alpha \in \gamma(a)$.
Since $\gamma(a) \notin \gamma(a)$, it cannot be injected into a. □

Lecture 20

Corollary 6.2 **Bourbaki-Witt Theorem** (Theorem 1.12)
Let P be a chain-complete poset
and $f : P \to P$ be an inflationary map
Then, for every $x \in P$, $\exists\ y \in P$ with $x \leq y = f(y)$.

Proof Let $\gamma(P)$ be the Hartogs ordinal of P
and define $g : \gamma(P) \to P$ by $\in$-recursion:

$$g(0) = x$$
$$g(\alpha^+) = f(g(\alpha))$$
$$g(\lambda) = \bigvee \{ g(\alpha) \mid \alpha < \lambda \} \text{ if } \lambda \text{ is a non-zero limit ordinal}$$

g is order-preserving by induction, so $\bigvee \{ g(\alpha) \mid \alpha < \lambda \}$ is a chain
By 6·1 g is not a chain,
so $\exists\ \alpha < \beta < \gamma(P)$ with $g(\alpha) = g(\beta)$.
Then $f(g(\alpha)) = g(\alpha^+)$
$\leq g(\beta)$
$= g(\alpha)$,
so $g(\alpha)$ is a fixed point of f. □

Note This proof uses replacement,
whereas the proof in 1·12 works in Zermelo set theory.

Notation For a set a, $\mathcal{P}^+ a$ is $\mathcal{P}a \setminus \{\emptyset\}$

Definition 6·3 A choice function for a is a function $g : \mathcal{P}^+ a \to a$
such that $g(b) \in b$ for all $b \in \mathcal{P}^+ a$.

Note If $\{ a_i \mid i \in I \}$ is a family of non-empty sets,
then each $a_i \in \mathcal{P}^+ \bigcup \{ a_i \mid i \in I \}$,
so a choice function (in this sense) for $\bigcup \{ a_i \mid i \in I \}$
yields a choice function (in the sense of §1) for $\{ a_i \mid i \in I \}$

So the assertion $(\forall x)(\exists g)(g \text{ is a choice function for } x)$
becomes our formulation of AC

Theorem 6·4 Zermelo (1904)
A set a can be well-ordered $\Leftrightarrow$ it has a choice function.

Proof Assume a is non-empty

$\Rightarrow$ Given a well-ordering $<$ of a,
we have a choice function $g : \mathcal{P}^+ a \to a$
sending each $b \in \mathcal{P}^+ a$ to its $<$-least element.

$\Leftarrow$ Suppose we are given a choice function $g : \mathcal{P}^+ a \to a$.
We define $f : \gamma(a) \to a$ by recursion

$$f(\alpha) = \begin{cases} g(a \setminus \{ f(\beta) \mid \beta < \alpha \}) & \text{if } \{ f(\beta) \mid \beta < \alpha \} \neq a \\ g(a) & \text{otherwise} \end{cases}$$

By 6·1, f is not injective.

But if we never use the second clause in the definition,
then f is injective,
so $\exists\ \alpha < \gamma(a)$ such that $\{ f(\beta) \mid \beta < \alpha \} = a$.
Consider the least such α.
Then $f \mid \alpha$ is defined using only the first clause of the definition,
so $f \mid \alpha$ is injective and hence bijective.
So we can well-order a by setting

$$x < y \iff (f \mid \alpha)^{-1}(x) < (f \mid \alpha)^{-1}(y)$$

Remarks 6·5 a. Zermelo's original proof of 6·4 did not use Hartog's Lemma.
He used an argument similar to the proof of Bourbaki-Witt in §1,
which works in Zermelo set theory.

b. We could alternatively deduce the Well-Ordering Theorem from Zorn's Lemma:
Given a, consider the set P of well-orderings of subsets of a,
ordered by $\langle b, <_b \rangle \leq \langle c, <_c \rangle$
$\iff$ b is a $<_c$-initial segment of c,
and $<_b = <_c \cap (b \times b)$.
With this ordering, P is chain-complete (joins of chains are unions).
So it has a maximal element $\langle b, <_b \rangle$ say.
If $b \neq a$, we can ~~set~~ pick $x \in a \setminus b$
and set $b' = b \cup \{x\}$
$<_{b'} = <_b \cup \{ \langle y, x \rangle \mid y \in b \}$
Then $\langle b, <_b \rangle < \langle b', <_{b'} \rangle$ $\Rightarrow\Leftarrow$
So $b = a$, and $<_b$ is a well-ordering of a. □

c. The applications of Zorn's Lemma that we saw in 1·15
can all be proved using the Well-Ordering Theorem instead.

For example, to prove Hamel's Theorem (every vector space has a basis),
we can argue as follows:

Given a vector space V,
well-order its underlying set as $\{ x_\alpha \mid \alpha < \gamma \}$ for some $\gamma \in \mathrm{On}$,
and define subsets S_α, $\alpha < \gamma$ of V ~~as follows~~ recursively:
$S_0 = \emptyset$
if $x_\alpha \in \langle S_\alpha \rangle$, set $S_{\alpha+} = S_\alpha$
otherwise $S_{\alpha+} = S_\alpha \cup \{ x_\alpha \}$
if λ is a limit ordinal, set $S_\lambda = \bigcup \{ S_\alpha \mid \alpha < \lambda \}$
Then S_α is linearly independent for all α,
and $S_\gamma = \bigcup \{ S_\alpha \mid \alpha < \gamma \}$ spans V
since $x_\alpha \in \langle S_{\alpha+} \rangle$ for all α.

Definition 6·6 An ordinal α is initial if there is no bijection $\alpha \to \beta$ for any $\beta < \alpha$.

According to this definition, every finite ordinal is initial;
ω is initial, and we can enumerate the infinite initial ordinals as $\{\omega_\alpha \mid \alpha \in On\}$
by the recursive definition

$$\omega_0 = \omega$$
$$\omega_{\alpha+} = \gamma(\omega_\alpha)$$
$$\omega_\lambda = \bigcup \{\omega_\alpha \mid \alpha < \lambda\} \text{ if } \lambda \text{ is a nonzero limit}$$

Lemma 6·7 The infinite initial ordinals are exactly the ω_α, $\alpha \in On$.

Proof

ω_0 is initial.
$\omega_{\alpha+}$ is initial since a bijection $\omega_{\alpha+} \to \gamma$ for $\gamma < \omega_{\alpha+}$
would yield an injection $\omega_{\alpha+} \to \omega_\alpha$
ω_λ is initial since a bijection $\omega_\lambda \to \gamma$ for some $\gamma < \omega_\lambda$
would yield an injection $\omega_{\beta+} \to \omega_\beta$, where $\gamma \leq \omega_\beta < \omega_\lambda$.

Conversely, suppose β is an infinite initial ordinal.
By example sheet 4, question 3, we have $\alpha \leq \omega_\alpha$ for all α
so there is a least α such that $\omega_\alpha \geq \beta$.
This α must be a successor δ^+, say;
so we have $\omega_\delta \leq \beta < \omega_{\delta+} = \gamma(\omega_\delta)$,
so we have injections $\omega_\delta \to \beta$
and $\beta \to \omega_\delta$.
Hence by Cantor-Bernstein (1·11) we have a bijection $\beta \to \omega_\delta$,
so by initiality of β we have $\beta = \omega_\delta$. □

Lecture 21

Informally, a cardinal is an equivalence class of sets under the relation
$(a \sim b \Leftrightarrow \exists \text{ a bijection } a \to b)$
Except for $\{\emptyset\}$, the equivalence classes of this relation are all proper classes,
so we seek a function-class $\text{card} : V \to V$
such that $(\forall x, y)((\text{card}\, x = \text{card}\, y) \Leftrightarrow (x \sim y))$

Definition 6·8 a. If we assume the Axiom of Choice,
then every $\sim$-equivalence class contains an ordinal by 6·4,
so it contains a unique initial ordinal,
and we ~~define~~ can define $\text{card}\, x$ to be the unique initial ordinal α
satisfying $x \sim \alpha$

b. If we do not assume the Axiom of Choice,
then for any x we define the essential rank of x to be the least ordinal α
such that $(\exists y)((\text{rank}\, y = \alpha) \wedge (y \sim x))$
Then we can define $\text{card}\, x = \{y \in V_{ess.rk(x)+} \mid x \sim y\}$ (a set by Separation)
Clearly $\text{card}\, x = \text{card}\, y$ implies $(\exists z)(z \in \text{card}\, x \cap \text{card}\, y)$
and hence $x \sim z \sim y$

We introduce a new name for card ω_α, even if we are using definition a: following Cantor, we denote it by $\aleph_\alpha$ ($\aleph$ is the first letter of the Hebrew alphabet)

We define a binary relation $\leq$ on cardinals by

$m \leq n \iff \exists$ an injection $x \to y$ where $\operatorname{card} x = m$ and $\operatorname{card} y = n$

This is obviously (well-defined), reflexive and transitive:
it's anti-symmetric (and hence a partial order) by $|\cdot|$.
By example sheet 4, question 7,
it's a total order on cardinals iff the Axiom of Choice holds.

We define sum, product and exponentiation for cardinals:
if x and y are sets of cardinalities m and n respectively,
then $m + n = \operatorname{card}(x \amalg y)$
$m \cdot n = \operatorname{card}(x \times y)$
$m^n = \operatorname{card}(x^y)$ where x^y is the set of all functions $y \to x$

Again, we need to verify that these are well-defined;
but this is easy

Lemma 6.9
i. $+$ is associative and commutative, with $0 = \operatorname{card} \emptyset$ as the identity element
ii. $\cdot$ is associative and commutative, with $1 = \operatorname{card} \{\emptyset\}$ as the identity element
iii. $m \cdot (n + p) = m \cdot n + m \cdot p$
iv. $m^{(n+p)} = m^n \cdot m^p$
$m^{(n \cdot p)} = (m^n)^p$
$(m \cdot n)^p = m^p \cdot n^p$

Proof
iv. Let a, b, c be sets of cardinality m, n, p respectively.
We have a bijection $a^{(b \amalg c)} \to a^b \times a^c$
given by $f \longmapsto \langle g, h \rangle$
where $g(y) = f(\langle y, 0 \rangle)$
and $h(z) = f(\langle z, 1 \rangle)$
with inverse given by $\langle g, h \rangle \longmapsto f$
where $f(\langle x, j \rangle) = \begin{cases} g(x) & \text{if } j = 0 \\ h(x) & \text{if } j = 1 \end{cases}$

We have a bijection $a^{(b \times c)} \to (a^b)^c$
given by $f \longmapsto \underline{f}$
where $\underline{f}(z)(y) = f(\langle y, z \rangle)$

And we have a bijection $(a \times b)^c \to a^c \times b^c$
given by $f \longmapsto \langle \pi_1 \circ f, \pi_2 \circ f \rangle$
where $\pi_1 : a \times b \to a$
and $\pi_2 : a \times b \to b$
are the product projections. □

Lemma 6.10 For any ordinal α, we have $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$

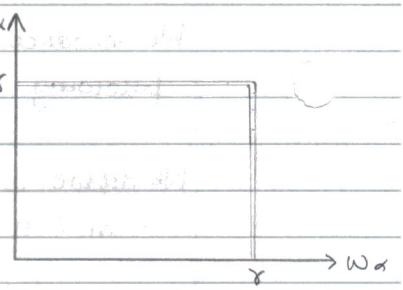
Proof We define a well-ordering of $\omega_\alpha \times \omega_\alpha$ as follows:

$\langle \gamma_1, \delta_1 \rangle < \langle \gamma_2, \delta_2 \rangle$

$\Leftrightarrow$ $(\gamma_1 \cup \delta_1) < (\gamma_2 \cup \delta_2)$

or $(\gamma_1 \cup \delta_1) = (\gamma_2 \cup \delta_2)$ and $\gamma_1 < \gamma_2$

or $\gamma_1 = \gamma_2 \geq (\delta_1 \cup \delta_2)$ and $\delta_1 < \delta_2$



We need to check this is a well-ordering:

given a non-empty $S \subseteq \omega_\alpha \times \omega_\alpha$

first set $S_1 = \{ \langle \gamma, \delta \rangle \in S \mid (\forall \langle \gamma', \delta' \rangle \in S)(\gamma \cup \delta \leq \gamma' \cup \delta') \}$

set $S_2 = \{ \langle \gamma, \delta \rangle \in S_1 \mid (\forall \langle \gamma', \delta' \rangle \in S_1)(\gamma \leq \gamma') \}$

If S_2 is not a singleton,

set $S_3 = \{ \langle \gamma, \delta \rangle \in S_2 \mid (\forall \langle \gamma', \delta' \rangle \in S_2)(\delta \leq \delta') \}$

Now, let β be the order-type of this well-ordering.

Then, for any $\theta < \beta$, there exists $\gamma < \omega_\alpha$ such that θ injects into $\gamma \times \gamma$.

Now either γ is finite, in which case $\gamma \times \gamma$ is finite,

so $\text{card}\, \theta < \aleph_0 \leq \aleph_\alpha$

or γ is infinite, in which case $\text{card}\, \gamma = \aleph_\delta$ for some $\delta < \alpha$,

so $\text{card}\,(\gamma \times \gamma) = \aleph_\delta \cdot \aleph_\delta$

$= \aleph_\delta$ (by the induction hypothesis)

so $\text{card}\, \theta \leq \aleph_\delta < \aleph_\alpha$

In either case, we deduce $\theta < \omega_\alpha$ for all such θ,

so $\beta \leq \omega_\alpha$.

But we also have $\omega_\alpha \leq \beta$, since ω_α injects into $\omega_\alpha \times \omega_\alpha$ and ω_α is initial.

So $\beta = \omega_\alpha$

and hence $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$

Corollary 6.11 For any ordinals α and β, we have $\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta$

$= \aleph_{(\alpha \cup \beta)}$

Proof Assume $\beta \leq \alpha$

We have injections $\omega_\alpha \longrightarrow \omega_\alpha \amalg \omega_\beta \xrightarrow{f} \omega_\alpha \times \omega_\beta \longrightarrow \omega_\alpha \times \omega_\alpha \cong \omega_\alpha$

$(\gamma \longmapsto \langle \gamma, 0 \rangle)$

where $f(\langle \gamma, 0 \rangle) = \langle \gamma, 0 \rangle$

$f(\langle \delta, 1 \rangle) = \langle 0, \delta \rangle$ if $\delta \neq 0$

$f(\langle 0, 1 \rangle) = \langle 1, 1 \rangle$

Hence $\aleph_\alpha \leq \aleph_\alpha + \aleph_\beta$

$\leq \aleph_\alpha \cdot \aleph_\beta$

$\leq \aleph_\alpha^2$

$= \aleph_\alpha$, and the result follows from 1.11.

Lecture 22

Lemma 6.12 Suppose $m = \text{card}(a)$ and $n = \text{card}(b)$ satisfy $m + n = m \cdot n$.

Then $\exists$ either an injection $a \to b$

or a surjection $a \to b$

Proof By assumption, we have a bijection $a \amalg b \xrightarrow{f} a \times b$

Consider the composite $g : a \longrightarrow a \amalg b \longrightarrow a \times b \longrightarrow b$
$(x \longmapsto \langle x, 0\rangle) \quad (\langle x, y\rangle \longmapsto y)$

If this is surjective, then we are done.
If not, pick $y_0 \in b \setminus \text{img}$,
and consider the map $a \longrightarrow a \times b \xrightarrow{f^{-1}} a \amalg b$
$(x \longmapsto \langle x, y_0\rangle)$

This has image contained in $b \times \{1\}$,
so its composite with the mapping $(\langle y, 1\rangle \longmapsto y)$ is an injection $a \to b$ □

Corollary 6·13 Suppose every cardinal $m \geq \aleph_0$ satisfies $m^2 = m$.
Then the Axiom of Choice holds.

Proof We show that any set a can be well-ordered.
Given a, consider $\gamma(a)$;
if $\gamma(a) < \omega$ then there is a bijection $(\gamma(a) - 1) \longrightarrow a$.
Otherwise, $\gamma(a) \geq \omega$.
Let $m = \text{card}(\gamma(a))$
$n = \text{card}(a)$
Then $m + n \geq \aleph_0$
But we have $m + n \leq m \cdot n$ by the proof of 6·11
$\leq (m+n)^2$ since $(m+n)^2 = m^2 + m\cdot n + n\cdot m + n^2$
$= m + n$
So by Cantor-Bernstein we have $m + n = m \cdot n$.
Hence, since there is no injection $\gamma(a) \to a$,
there must be a surjection $\gamma(a) \xrightarrow{f} a$ by 6·12.
Now define $g : a \longrightarrow \gamma(a)$
by $g(x) =$ least element of $f^{-1}(x)$.
Then g is injective, and we can well-order a by
$x < y \iff g(x) < g(y)$. □

What do we know about cardinal exponentiation?

If $m = \text{card}(a)$, then $2^m = \text{card}(\mathcal{P}a)$,
and Cantor's diagonal argument tells us that $m < 2^m \quad \forall\, m$.

From now on, we assume the Axiom of Choice,
so that all infinite cardinals are $\aleph$'s

Lemma 6·14 If $\beta \leq \alpha^+$, then $\aleph_\beta^{\aleph_\alpha} = 2^{\aleph_\alpha}$

Proof Since $2 \leq \aleph_\beta$, we have $2^{\aleph_\alpha} \leq \aleph_\beta^{\aleph_\alpha}$
But $2^{\aleph_\alpha} \geq \aleph_{\alpha^+}$, so $\aleph_\beta^{\aleph_\alpha} \leq (2^{\aleph_\alpha})^{\aleph_\alpha}$
$= 2^{\aleph_\alpha \cdot \aleph_\alpha}$
$= 2^{\aleph_\alpha}$
The result follows from Cantor-Bernstein. □

The Continuum Hypothesis
$2^{\aleph_0} = \aleph_1$

Cantor tried to prove this for many years.

The Generalised Continuum Hypothesis
$2^{\aleph_\alpha} = \aleph_{\alpha^+} \quad \forall\, \alpha \in \underset{\sim}{On}$

In ~ 1930 K. Gödel showed that if ZF is consistent
then $\exists$ models of ZF
in which GCH holds.
(Informally, he did this by 'slowing down' the von Neumann hierarchy V_α
to get a sequence of sets L_α
where L_{α^+} is the set of subsets of L_α
definable by formulae with parameters in L_α)

In 1964, P. Cohen showed that if ZF is consistent
then so is $ZF + AC + \neg(CH)$
(Informally, this is done by 'freely adding' a new subset of $\mathbb{R}$
and ensuring that we don't add bijections from this set to $\mathbb{R}$ or to $\mathbb{N}$).

Following Cohen's work, we know that there are very few restrictions on the function class $F : \underset{\sim}{On} \to \underset{\sim}{On}$
defined by $\aleph_{F(\alpha)} = 2^{\aleph_\alpha}$
beyond the obvious ones $(\alpha < F(\alpha))$
and $(\alpha \leq \beta) \Rightarrow (F(\alpha) \leq F(\beta))$

Given an infinite family of sets $(a_i \mid i \in I)$ with $\text{card}(a_i) = m_i$,
we write $\sum_{i\in I} m_i = \text{card}\left(\coprod_{i\in I} a_i\right)$
where $\coprod_{i\in I} a_i = \bigcup \{ a_i \times \{i\} \mid i \in I \}$
$\prod_{i\in I} m_i = \text{card}\left(\prod_{i\in I} a_i\right)$
where $\prod_{i\in I} a_i$ is the set of all choice functions for $(a_i \mid i \in I)$

Lemma 6.15 König's Lemma
If $m_i < n_i \ \forall\, i$, then $\sum_{i\in I} m_i < \prod_{i\in I} n_i$

Proof We will only show $\sum_{i\in I} m_i \neq \prod_{i\in I} n_i$
Suppose $\text{card}(a_i) = m_i$ and $\text{card}(b_i) = n_i$ for each i
Given a function $f : \coprod_{i\in I} a_i \to \prod_{i\in I} b_i$, we show f cannot be surjective.
For each i, the composite
$f_i : a_i \to \coprod_{i\in I} a_i \xrightarrow{f} \prod_{i\in I} b_i \to b_i$ is not surjective,
so we can choose $y_i \in b_i \setminus \text{im}\, f_i$.
Then the function $(i \mapsto y_i)$ is an element of $\prod_{i\in I} b_i$
not in the image of f. $\square$

Remark If we set $m_i = 0 \quad \forall i$
then König's Lemma becomes the statement of the Axiom of Choice.
If we set $m_i = 1$ and $n_i = 2 \quad \forall i$
then König's Lemma reduces to Cantor's Theorem, $k < 2^k$

Corollary 6.16 $2^{\aleph_0} \neq \aleph_\omega$

Proof Take $I = \omega$ and set $a_0 = \omega$
$a_i = \omega_i \setminus \omega_{i-1}$ for $i > 0$
$b_i = \omega_\omega \quad \forall i$

Then $\operatorname{card} a_i = \aleph_i$
$< \aleph_\omega$
$= \operatorname{card} b_i \quad \forall i$

So $\aleph_\omega = \operatorname{card} \omega_\omega$
$= \operatorname{card} \left(\coprod_{i \in I} a_i \right)$
$< \operatorname{card} \left(\prod_{i \in I} b_i \right)$
$= \aleph_\omega^{\aleph_0}$

But $2^{\aleph_0} = 2^{(\aleph_0 \cdot \aleph_0)}$
$= (2^{\aleph_0})^{\aleph_0}$

so $2^{\aleph_0} \neq \aleph_\omega$

Similarly, if λ is any limit ordinal of cardinality ω (cf example sheet 4, q 12)
then $2^{\aleph_0} \neq \aleph_\lambda$ □

Non-examinable

§8 Consistency and independence

Are the axioms of ZF independent?

As we presented it, three axioms are derivable from others:
Empty set (follows from Infinity)
Pair set (follows from Replacement)
Separation (follows from Replacement)

To form $\{ y \in x \mid \varphi \}$ apply Replacement to the function class
$(F(y) = z) \Leftrightarrow (y = z) \wedge \varphi$

(We could reformulate Replacement as Collection,
so that it doesn't imply Separation)

The others are all necessary,
in that leaving them out allows models which aren't models of ZF:
Infinity, Power-Set, Union — see exercise sheet 3, question 9
Replacement — see exercise sheet 3, question 13
Foundation — see exercise sheet 3, question 5

Are they complete?

No, because we know set-theoretic statements which we can neither prove nor disprove in ZF.

eg AC is consistent with ZF
since it holds in Gödel's model $L = \bigcup \{ L_\alpha \mid \alpha \in On \}$.
the idea is that, given a well-ordering of L_α,
we can well-order the formulae with parameters in L_α,
and hence we can well-order $L_{\alpha+1}$
And we can do this in such a way so that
$L_\lambda = \bigcup \{ L_\alpha \mid \alpha < \lambda \}$ can be well-ordered for limit ordinals λ.

Fraenkel (1920s) showed that AC is independent of 'ZF with atoms',
ie ZF with a collection of atoms which can be members of sets,
but have no members.

Mosstowski (1930s) modified Fraenkel's method using sets x
satisfying $x = \{ x \}$ instead of atoms (see exercise sheet 3, question 12)

Cohen (1963) showed that AC is independent of ZF,
by 'forcing' the counterexample in a F-M model
into the well-founded part of the universe.

We could still hope that ZF is completable
by adding some list of axioms or axiom-schemes.
(Just as the theory of fields can be extended to the complete theory
of algebraically closed fields of a given characteristic).

Gödel showed that this is not possible.

Recall that we have an interpretation of Peano Arithmetic in ZF,
by means of the set ω.

Also, ZF is recursively presented,
in the sense that there is an algorithm for determining whether or not
a given formula is an axiom.
(And the language of ZF is countable).
We can enumerate the formulae of the language of ZF:
write $\ulcorner \varphi \urcorner$ for the number coding the formula φ.
We can also enumerate finite strings of formulae.
Write $\mathrm{Der}_T(x, y)$ for the binary relation on $\mathbb{N}$
which holds if x is the code for a derivation of the formula
coded by y from the theory T.

If T is a recursively-presented extension of ZF,
then $\mathrm{Der}_T(-,-)$ is definable in PA.

We write $\mathrm{Thm}_T(y)$ for $(\exists x)\,\mathrm{Der}_T(x,y)$

If $T \vdash \varphi$ then $PA \vdash \mathrm{Thm}(\ulcorner \varphi \urcorner)$.
The converse may fail, since PA lacks witnesses.

But the construction of the derivation of $\mathrm{Thm}_T(\ulcorner\varphi\urcorner)$ from that of φ is algorithmic.

So $PA \vdash (\forall x)(\mathrm{Thm}_T(x) \Rightarrow \mathrm{Thm}_T(\ulcorner \mathrm{Thm}_T(x)\urcorner))$

If we write $\mathrm{imp}(x,y)$ for the function (definable in PA)
such that $\mathrm{imp}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) = \ulcorner(\varphi \Rightarrow \psi)\urcorner$
then $PA \vdash (\forall x, y)((\mathrm{Thm}_T(x) \wedge \mathrm{Thm}_T(\mathrm{imp}(x,y))) \Rightarrow \mathrm{Thm}_T(y))$

Now let $\varphi(x) = \neg\,\mathrm{Thm}_T(x)$
let c be the function such that $c(n) = \ulcorner n \urcorner$
let sub_x be the function such that $\mathrm{sub}_x(\ulcorner\varphi\urcorner, \ulcorner t\urcorner) = \ulcorner \varphi[t/x]\urcorner$
let $\psi(y) = \varphi(\mathrm{sub}_x(y, c(y)))$
let $m = \ulcorner \psi(x)\urcorner$

Then $\psi(m) = \varphi(\mathrm{sub}_x(m, c(m)))$
$= \varphi(\mathrm{sub}_x(\ulcorner\psi(x)\urcorner, \ulcorner m\urcorner))$
$= \varphi(\ulcorner\psi(m)\urcorner)$
$= \neg\,\mathrm{Thm}_T(\ulcorner\psi(m)\urcorner)$

Clearly $T \nvdash \psi(m)$ unless T is inconsistent.

(We could have $T \vdash \neg\,\psi(m)$,
but not if T satisfies the stronger property of ω-consistency)

Write Con_T for $\neg\,\mathrm{Thm}_T(\ulcorner\bot\urcorner)$

Claim: $PA \vdash (\psi(m) \Leftrightarrow \mathrm{Con}_T)$
so if T is consistent then $T \vdash \mathrm{Con}_T$

Clearly $\vdash (\bot \Rightarrow \psi(m))$
so $PA \vdash \mathrm{Thm}_T(\ulcorner(\bot \Rightarrow \psi(m))\urcorner)$
so by formalised modus ponens we have
$PA \vdash \mathrm{Thm}_T(\ulcorner\bot\urcorner) \Rightarrow \mathrm{Thm}_T(\ulcorner\psi(m)\urcorner)$
so $PA \vdash \psi(m) \Rightarrow \neg\,\mathrm{Thm}_T(\ulcorner\bot\urcorner)$

Conversely $PA \vdash \mathrm{Thm}_T(\ulcorner \Psi(m) \urcorner) \Rightarrow \mathrm{Thm}_T(\ulcorner \mathrm{Thm}_T(\ulcorner \Psi(m) \urcorner) \urcorner)$

ie $PA \vdash \mathrm{Thm}_T(\ulcorner \Psi(m) \urcorner) \Rightarrow \mathrm{Thm}_T(\ulcorner \neg \Psi(m) \urcorner)$

so by formalised modus ponens

$PA \vdash \mathrm{Thm}_T(\ulcorner \Psi(m) \urcorner) \Rightarrow \mathrm{Thm}_T(\ulcorner \bot \urcorner)$

hence $PA \vdash \neg \mathrm{Thm}_T(\ulcorner \bot \urcorner) \Rightarrow \Psi(m)$