

## Introduction

The course is split into two parts:

- Logic: syntax and semantics.
- Set theory: what does the universe of sets look like?

Course structure

- (I) Propositional logic (logic)
- (II) Well-orderings & ordinals (set theory)
- (III) Posets & Zorn's lemma (set theory)
- (IV) Predicate logic (logic)
- (V) Set theory (set theory)
- (VI) Cardinals (set theory)

Books:

- 1. Johnstone, *Notes on Logic & Set Theory*
- 2. Van Dalen, *Logic & Structure* (Chapter 4 and what 'goes next')
- 3. Hajnal & Hamburger, *Set Theory* (Chapters 2 and 6)
- 4. Forster, *Logic, Induction & Sets*

## 1 Propositional Logic

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

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

e.g.  $((p_1 \Rightarrow p_2) \Rightarrow (p_1 \Rightarrow p_3)), (p_4 \Rightarrow \perp), (\perp \Rightarrow \perp)$ .

**Notes.**

- 1. Each proposition (member of  $L$ ) is a finite string of symbols from language:  $\vdash, \Rightarrow, \perp, p_1, p_2, \dots$  (for clarity often omit outer brackets, use other types of bracket, etc).
- 2. ' $L$  is defined inductively' means, more precisely, the following

- Put  $L_1 = P \cup (\perp)$ ;
- Having defined  $L_n$ , put  $L_{n+1} = L_n \cup \{(p \Rightarrow q) : p, q \in L_n\}$ ;
- Set  $L = \bigcup_{n \geq 1} L_n$ .

3. Every  $p \in L$  is uniquely built up from steps 1,2 using 3. For example,  $((p_1 \Rightarrow p_2) \Rightarrow (p_1 \Rightarrow p_3))$  can from  $(p_1 \Rightarrow p_2)$  and  $(p_1 \Rightarrow p_3)$ .

We can now introduce  $\neg p$  ('not  $p$ ') as an abbreviation for  $(p \Rightarrow \perp)$ ;  $p \vee q$  (' $p$  or  $q$ ') as an abbreviation for  $(\neg p) \Rightarrow q$ ;  $p \wedge q$  (' $p$  and  $q$ ') as an abbreviation for  $\neg(p \Rightarrow (\neg q))$ .

### 1.1 Semantic Implication

**Definition.** A *valuation* is a function  $v : L \rightarrow \{0, 1\}$  (thinking of 0 as ‘False’ and 1 as ‘True’) such that

$$(i) \quad v(\perp) = 0$$

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

**Remark.** On  $\{0, 1\}$ , could define a constant  $\perp = 0$  and an operation  $\Rightarrow$  by

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

Then a valuation is precisely a mapping  $L \rightarrow \{0, 1\}$  that preserves ( $\perp$  and  $\Rightarrow$ ).

**Proposition 1.1.**

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

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

*Proof.*

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

(ii) Set  $v(p) = w(p)$  for all  $p \in P$  and  $v(\perp) = 0$  to obtain  $v$  on  $L_1$ . Now put

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

to obtain  $v$  on  $L_2$ , then induction.

□

**Example.** Let  $v$  be the valuation with  $v(p_1) = v(p_3) = 1$ ,  $v(p_n) = 0$  for all  $n \neq 1, 3$ . Then  $v((p_1 \Rightarrow p_2) \Rightarrow p_3) = 0$ .

**Definition.** A *tautology* is an element  $t \in L$  such that  $v(t) = 1$  for any valuation  $v$ . We write  $\models t$ .

**Examples.**

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

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

So this is a tautology.

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

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

3.  $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$  ("how implication chains").  
 Suppose this is not a tautology. Then we have a  $v$  with  $v(p \Rightarrow (q \Rightarrow r)) = 1$  and  $v((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) = 0$ . Then  $v(p \Rightarrow q) = 1$  and  $v(p \Rightarrow r) = 0$ . Hence  $v(p) = 1$  and  $v(r) = 0$ , so  $v(q) = 1$ . Hence  $v(p \Rightarrow (q \Rightarrow r)) = 0$ , contradiction.

**Definition.** For  $S \subseteq L$ ,  $t \in L$ , we say  $S$  *entails* or *semantically implies*  $t$ , written  $S \models t$  if every valuation with  $v(s) = 1$  for all  $s \in S$  has  $v(t) = 1$ .

**Example.**  $\{p \Rightarrow q, q \Rightarrow r\}$  entails  $p \Rightarrow r$ . Indeed, suppose we have  $v$  with  $v(p \Rightarrow q), v(q \Rightarrow r) = 1$  but  $v(p \Rightarrow r) = 0$ . Then  $v(p) = 1, v(r) = 0$ . Hence  $v(q) = 1$ , contradicting  $v(q \Rightarrow r) = 1$ .

**Definition.** We say  $v$  is a *model* of  $S \subseteq L$  or  $S$  is *true* in  $v$ , if  $v(s) = 1$  for all  $s \in S$ . Thus  $S$  entails  $t$  means: every model of  $S$  is also a model of  $\{t\}$ .

**Remark.**  $\models t$  says  $\emptyset \models t$ .

## 1.2 Syntactic implication

For a notion of proof, we'll need axioms and deduction rules. As axioms, we'll take:

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

**Notes.**

1. Sometimes we call these 'axiom schemes' since each is actually a set of axioms.
2. Each of these are tautologies.

For deduction rules, we'll have only *modus ponens*: from each  $p$  and  $p \Rightarrow q$  we can deduce  $q$ .

**Definition.** For  $S \subseteq L$ , and  $t \in S$ , say  $S$  *proves* or *syntactically implies*  $t$ , written  $S \vdash t$  if there exists a sequence  $t_1, \dots, t_n$  in  $L$  with  $t_n = t$  such that every  $t_i$  is either

- (i) An axiom; or
- (ii) A member of  $S$ ; or
- (iii) Such that there exist  $j, k < i$  with  $t_k \Rightarrow (t_j \Rightarrow t_n)$  (modus ponens).

Say  $S$  consists of the *hypotheses* or *premises*, and  $t$  the *conclusion*.

**Example.**  $\{p \Rightarrow q, q \Rightarrow r\} \vdash p \Rightarrow r$ :

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

**Definition.** If  $\emptyset \vdash t$ , say  $t$  is a *theorem*, written  $\vdash t$ .

**Example.**  $\vdash (p \Rightarrow p)$ . We want to try to get to  $(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)$  using axiom 2.

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

Often, showing  $S \vdash p$  is made easier by:

**Proposition 1.2** (Deduction Theorem). *Let  $S \subseteq L$  and  $p, q \in L$ . Then  $S \vdash (p \Rightarrow q)$  if and only if  $S \cup \{p\} \vdash q$ . Informally: “provability corresponds to the connective ‘ $\Rightarrow$ ’ in  $L$ ”.*

*Proof.* First we show  $(\Rightarrow)$ : given a proof of  $p \Rightarrow q$  from  $S$ , write down:

1.  $p$  (hypothesis)
2.  $q$  (modus ponens)

Which is a proof of  $q$  from  $S \cup \{p\}$ .

Now we show  $(\Leftarrow)$ : we have a proof  $t_1, \dots, t_n$  of  $q$  from  $S \cup \{p\}$ . We’ll show that  $S \vdash (p \Rightarrow t_i)$  for all  $i$ .

If  $t_i$  is an axiom, write down

1.  $t_i$  (axiom)
2.  $t_i \Rightarrow (p \Rightarrow t_i)$  (axiom 1)
3.  $p \Rightarrow t_i$  (modus ponens)

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

If  $t_i \in S$ , do the same thing except step 1 will be “ $t_i$  (hypothesis)” instead of “ $t_i$  (axiom)”.

If  $t_i := p$ , we have  $S \vdash (p \Rightarrow p)$ , since  $\vdash (p \Rightarrow p)$ .

If  $t_i$  is obtained by modus ponens, we have  $t_j$  and  $t_k = (t_j \Rightarrow t_i)$  for some  $j, k < n$ . By induction, we can assume  $S \vdash (p \Rightarrow t_j)$  and  $S \vdash (p \Rightarrow (t_j \Rightarrow t_i))$ . So write down

1.  $[p \Rightarrow (t_j \Rightarrow t_i)] \Rightarrow [(p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)]$  (axiom 2)
2.  $(p \Rightarrow t_j) \Rightarrow (p \Rightarrow t_i)$  (modus ponens)

3.  $p \Rightarrow t_i$  (modus ponens)

So  $S \vdash p \Rightarrow t$ . □

**Example.** To show  $\{p \Rightarrow q, q \Rightarrow r\} \vdash (p \Rightarrow r)$ , it is sufficient to show  $\{p \Rightarrow q, q \Rightarrow r, p\} \vdash r$ , which is just modus ponens twice.

**Question:** how are  $\models$  and  $\vdash$  related?

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

This is made up of:

- $S \vdash t \Rightarrow S \models t$  (soundness) i.e. “our axioms and deduction rule are not silly”;
- $S \models t \Rightarrow S \vdash t$  (adequacy) “our axioms are strong enough to deduce from  $S$ , every semantic consequence of  $S$ ”.

**Proposition 1.3** (Soundness). *Let  $S \subseteq L$ ,  $t \in L$ . Then  $S \vdash t \Rightarrow S \models t$ .*

*Proof.* We have a proof  $t_1, \dots, t_n$  of  $t$  from  $S$ . So we must show that every model of  $S$  is a model of  $t$ , i.e if  $v$  is a valuation with  $v(s) = 1$  for all  $s \in S$ , then  $v(t) = 1$ . But  $v(p) = 1$  for each axiom  $p$  (each axiom is a tautology), and for each  $p \in S$  whenever  $v(p) = v(p \Rightarrow q) = 1$ , we have  $v(q)$ . So  $v(t_i) = 1$  for all  $i$  (induction). □

One case of adequacy is: if  $S \models \perp$ , then  $S \vdash \perp$ . We say  $S$  is *consistent* if  $S \not\models \perp$ . So our statement is:  $S$  has no model  $\Rightarrow S$  inconsistent, i.e  $S$  consistent  $\Rightarrow S$  has a model.

In fact, this implies adequacy in general. Indeed, if  $S \models t$  then  $S \cup \{\neg t\}$  has no model. Hence (by the special case)  $S \cup \{\neg t\} \vdash \perp$ . So  $S \vdash (\neg t \Rightarrow \perp)$ , i.e  $S \vdash (\neg \neg t)$ . But  $S \vdash (\neg \neg t) \Rightarrow t$  (axiom 3), so  $S \vdash t$ .

So our task is: given  $S$  consistent, find a model of  $S$ . Could try: define

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

But this fails, since  $S$  might not be *deductively closed*, meaning  $S \vdash p \Rightarrow p \in S$ . So we could first replace  $S$  with its deductive closure  $\{t \in L : S \vdash t\}$  (which is consistent, because  $S$  is). However, this still fails: if  $S$  does not ‘mention’  $p_3$ , then  $S \not\models p_3$  and  $S \not\models \neg p_3$ , so  $v(p_3) = v(\neg p_3) = 0$  which is impossible.

**Theorem 1.4** (Model Existence Theorem). *Let  $S \subseteq L$  be consistent. Then  $S$  has a model.*

Idea: extend  $S$  to ‘swallow up’, for each  $p$ , one of  $p$  and  $\neg p$ .

*Proof.* Claim: for any consistent  $S \subseteq L$  and  $p \in L$ ,  $S \cup \{p\}$  or  $S \cup \{\neg p\}$  is consistent.

Proof of claim: if not, then  $S \cup \{p\} \vdash \perp$  and  $S \cup \{\neg p\} \vdash \perp$ . So  $S \vdash (p \Rightarrow \perp)$  (deduction theorem), i.e.  $S \vdash (\neg p)$ . Hence from  $S \cup \{\neg p\} \vdash \perp$  we obtain  $S \vdash \perp$ .

Now,  $L$  is countable (as each  $L_n$  is countable) so we can list  $L$  as  $t_1, t_2, \dots$ . Let  $S_0 = S$ . Let  $S_1 = S_0 \cup \{t_1\}$  or  $S_1 = S_0 \cup \{\neg t_1\}$  with  $S_1$  consistent. In general, given  $S_{n-1}$  let  $S_n = S_{n-1} \cup \{t_n\}$  or  $S_n = S_{n-1} \cup \{\neg t_n\}$  so that  $S_n$  is consistent. Now set  $\bar{S} = S_0 \cup S_1 \cup S_2 \cup \dots$ . Thus for all  $t \in L$ , either  $t \in \bar{S}$  or  $(\neg t) \in \bar{S}$ .

Now  $\bar{S}$  is consistent: if  $\bar{S} \vdash \perp$  then, since proofs are finite, we’d have  $S_n \vdash \perp$  for some  $n$ , a contradiction.

Also,  $\bar{S}$  is deductively closed: if  $\bar{S} \vdash p$ , must have  $p \in \bar{S}$ , since otherwise  $(\neg p) \in \bar{S}$ , so  $\bar{S} \vdash (p \Rightarrow \perp)$  and  $\bar{S} \vdash \perp$ .

Now define  $v : L \rightarrow \{0, 1\}$  by

$$t \mapsto \begin{cases} 1 & t \in \bar{S} \\ 0 & \text{otherwise} \end{cases}.$$

We’ll show  $v$  is a valuation (then we’re done as  $v = 1$  on  $S$ ).

$v(\perp)$ : have  $\perp \notin \bar{S}$  (since  $\bar{S}$  is consistent), so  $v(\perp) = 0$ .

$v(p \Rightarrow q)$ : if  $v(p) = 1$ ,  $v(q) = 0$ , then have  $p \in \bar{S}$ ,  $q \notin \bar{S}$ . But if  $(p \Rightarrow q) \in \bar{S}$ , then since  $p \in \bar{S}$ ,  $q \in \bar{S}$  (since  $\bar{S}$  is deductively closed). Now if  $v(q) = 1$ ,  $q \in \bar{S}$ . But  $\bar{S} \vdash (q \Rightarrow (p \Rightarrow q))$  (axiom 1), so  $\bar{S} \vdash (p \Rightarrow q)$  hence  $(p \Rightarrow q) \in \bar{S}$  ( $\bar{S}$  is deductively closed). Finally, if  $v(p) = 0$ , we have  $p \notin \bar{S}$  and want to show  $(p \Rightarrow q) \in \bar{S}$ . Then  $(p \Rightarrow \perp) \in \bar{S}$ , so it is enough to show  $(p \Rightarrow \perp) \vdash (p \Rightarrow q)$ . So it’s enough to show  $(p, p \Rightarrow \perp) \vdash q$ , so enough to show  $\perp \vdash q$ . But  $\perp \vdash (\neg \neg q)$  (axiom 1), and  $(\neg \neg q) \vdash q$  (axiom 3), so  $\perp \vdash q$  as required.  $\square$

#### Remarks.

1. We used  $P = (p_1, p_2, \dots)$ , in saying  $L$  is countable. In fact, it also holds if  $P$  is uncountable (see later in course).
2. Sometimes this theorem is called ‘The Completeness Theorem’

By the remarks stated before this theorem, we have



**Corollary 1.5** (Adequacy). *Let  $S \subseteq L$ ,  $t \in L$ , with  $S \models t$ . Then  $S \vdash t$ .*

Hence we have

**Theorem 1.6** (Completeness Theorem). *Let  $S \subseteq L$ ,  $t \in L$ . Then  $S \vdash t \iff S \models t$ .*

**Corollary 1.7** (Compactness Theorem). *Let  $S \subseteq L$ ,  $t \in L$  with  $S \models t$ . Then some finite  $S' \subseteq S$  has  $S' \models t$ .*

*Proof.* This is trivial if we replace  $\models$  by  $\vdash$  (as all proofs are finite).  $\square$

For  $t = \perp$ , the theorem says: if  $S \models \perp$  then some finite  $S' \subseteq S$  has  $S' \models \perp$ , i.e. if every finite  $S' \subseteq S$  has a model then  $S$  has a model. In fact, this is equivalent to compactness in general:  $S \models t$  says  $S \cup \{t\}$  has no model, and  $S' \models t$  says  $S' \cup \{t\}$  has no model.

**Corollary 1.8** (Compactness Theorem equivalent form). *Let  $S \subseteq L$ . Then if every finite subset of  $S$  has a model, so does  $S$ .*

Another application:

**Corollary 1.9** (Decidability Theorem). *Let  $S \subseteq L$  be finite and  $t \in L$ . Then there is an algorithm to decide, in finite time, whether or not  $S \vdash t$ .*

**Remark.** This is a very surprising result.

*Proof.* Trivial if we replace  $\vdash$  with  $\models$ : to check if  $S \models t$  we just draw the truth table.  $\square$

## 2 Well-ordering & Ordinals

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

- (i) *irreflexive*: for all  $x \in X$ , not  $x < x$ ;
- (ii) *transitive*: for all  $x, y, z \in X$ , if  $x < y$ ,  $y < z$  then  $x < z$ ;
- (iii) *trichotomous*: for all  $x, y \in X$ , either  $x = y$  or  $x < y$  or  $y < x$ .

We sometimes write  $x > y$  if  $y < x$ , and  $x \leq y$  if  $x < y$  or  $x = y$ .

We can instead define a total order in terms of  $\leq$  as follows:

- (i) *reflexive*: for all  $x \in X$ ,  $x \leq x$ ;
- (ii) *transitive*: for all  $x, y, z \in X$ , if  $x \leq y$ ,  $y \leq z$  then  $x \leq z$ ;
- (iii) *antisymmetric*: for all  $x, y \in X$ , if  $x \leq y$ ,  $y \leq x$  then  $x = y$ ;
- (iv) *trichotomous*: for all  $x, y \in X$  either  $x \leq y$  or  $y \leq x$ .

**Examples.**

1.  $\mathbb{N}, <$ ;
2.  $\mathbb{Q}, \leq$ ;
3.  $\mathbb{R}, \leq$ ;
4.  $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$  under ‘divides’ is not a total order, e.g 2 and 3 are not related;
5.  $\mathcal{P}(S), \subseteq$  is not a total order - fails trichotomy.

**Definition.** A total order  $(X, <)$  is a *well-ordering* if every (non-empty) subset has a least element, i.e for all  $S \subseteq X$  if  $S \neq \emptyset$  then there exists  $x \in S$  such that  $x \leq y$  for all  $y \in S$ .

**Examples.**

1.  $\mathbb{N}, <$ ;
2.  $\mathbb{Z}, <$  is not a well ordering;
3.  $\mathbb{Q}, <$  is not a well ordering;
4.  $\mathbb{R}, <$  is not a well ordering;
5.  $[0, 1] \subseteq \mathbb{R}, <$  is not a well ordering, e.g  $(0, 1]$  has no least element;
6.  $\{1/2, 2/3, 3/4, \dots\} \subseteq \mathbb{R}$  is well ordered;
7.  $\{1/2, 2/4, 3/4, \dots\} \cup \{1\}$  is well ordered;

8.  $\{1/2, 2/4, 3/4, \dots\} \cup \{2\}$  is well ordered;
9.  $\{1/2, 2/3, 3/4, \dots\} \cup \{1 + 1/2, 1 + 2/3, 1 + 3/4, \dots\}$  is well ordered.

**Remark.**  $(X, <)$  is a well ordering if and only if there is no infinite strictly decreasing sequence.

We say total orders  $X, Y$  are *isomorphic* if there exists a bijection  $f : X \rightarrow Y$  such that  $x < y$  if and only if  $f(x) < f(y)$ . For example, Examples 1&6, 7&8 above are isomorphic. However examples 1&7 are not isomorphic, since in 7 there exists a greatest element, but not in 1.

**Proposition 2.1** (Proof by induction). *Let  $X$  be well ordered and let  $S \subseteq X$  be such that whenever  $y \in S$  for all  $y < x$ , then  $x \in S$ . Then  $S = X$ . Equivalently, if  $p(x)$  is a property such that  $p(y)$  for all  $y < x$  implies  $p(x)$ , then  $p(x)$  for all  $x \in X$ .*

*Proof.* Suppose  $S \neq X$  and let  $x$  be least in  $X \setminus S$ . Then  $y \in S$  for all  $y < x$  but  $x \notin S$ , a contradiction.  $\square$

**Proposition 2.2.** *Let  $X, Y$  be isomorphic well-orderings. Then there exists a unique isomorphism.*

**Note.** Note this is false for general total orders, for example  $\mathbb{Z} \rightarrow \mathbb{Z}$  could have  $x \mapsto x - t$  for any  $t$ , or  $\mathbb{R} \rightarrow \mathbb{R}$  could have  $x \mapsto x^3$ .

*Proof.* Let  $f, g : X \rightarrow Y$  be isomorphisms. We'll show  $f(x) = g(x)$  for all  $x$  by induction on  $X$ . Given  $f(y) = g(y)$  for all  $y < x$ , we want to show  $f(x) = g(x)$ . We must have  $f(x) = a$  where  $a$  is the least element of  $Y \setminus \{f(y) : y < x\}$  (non-empty since it contains  $f(x)$ ). Indeed, if not then  $f(x') = a$  for some  $x' > x$ , contradicting the fact  $f$  is order preserving. Similarly have  $g(x) = a$ .  $\square$

**Definition.** A subset  $I$  of a total order  $X$  is an *initial segment* if  $x \in I, y < x$  implies  $y \in I$  (i.e  $I$  is closed under  $<$ ). For example  $I_x = \{y \in X : y < x\}$  is an initial segment for any  $x \in X$ , however not every initial segment is of this form, e.g in  $\mathbb{Q}$   $\{x \in \mathbb{Q} : x \leq 0 \text{ or } x^2 < 2\}$ .

**Note.** In a well-ordering, every proper initial segment  $I$  is of the form  $I_x$ , for some  $x \in X$ . Indeed let  $x$  be the least element of  $X \setminus I$  (non-empty since  $I$  is proper). Then  $I = I_x$ , since if  $y < x$  then  $y \in I$  (by choice of  $x$ ), and conversely if  $y \in I$ , must have  $y < x$  or else  $y \geq x$  implying  $x \in I$  (as  $I$  is an initial segment).

Our aim is to show that every subset of a well-ordering  $X$  is isomorphic to an initial segment of  $X$ .

**Note.** This is false in general for total orders, e.g.  $\{1, 2, 3\}$  in  $\mathbb{Z}$ , or  $\mathbb{Q}$  in  $\mathbb{R}$ .

**Theorem 2.3** (Definition by recursion). *Let  $X$  be a well-ordering and let  $Y$  be any set. Take  $G : \mathcal{P}(X \times Y) \rightarrow Y$  (i.e. a ‘rule’). Then there exists a function  $f : X \rightarrow Y$  such that  $f(x) = G(f|_{I_x})$  for all  $x \in X$ . Moreover,  $f$  is unique.*

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

*Proof.* Say  $h$  is ‘an attempt’ if  $h : I \rightarrow Y$  for some initial segment  $I$  of  $X$ , and for all  $x \in I$  we have  $h(x) = G(h|_{I_x})$ . [This is the main idea].

Note that if  $h, h'$  are attempts both defined at  $x$ , then  $h(x) = h'(x)$ , by induction on  $x$  (if  $h(y) = h'(y)$  for all  $y < x$  then  $h(x) = h'(x)$ ).

Also, for every  $x$ , there exists an attempt defined at  $x$ , also by induction. Indeed, suppose that for all  $y < x$  there exists an attempt defined at  $y$ . So for all  $y < x$  there exists a unique (by above) attempt  $h_y$  with domain  $\{z : z \leq y\}$ . Now let  $h = \bigcup_{y < x} h_y$ , this is an attempt with domain  $I_x$  (single valued by uniqueness). Thus  $h \cup \{(x, G(h))\}$  is an attempt defined at  $x$ . Now define  $f : X \rightarrow Y$  by setting  $f(x) = y$  if there exists an attempt  $h$  defined at  $x$  such that  $h(x) = y$ .

Uniqueness of  $f$ : if  $f, f'$  are both such functions, then  $f(x) = f'(x)$  for all  $x$  by induction ( $f(y) = f'(y)$  for all  $y < x$  implies  $f(x) = f'(x)$ ).  $\square$

**Proposition 2.4** (Subset collapse). *Let  $X$  be a well-ordering and  $Y \subseteq X$ . Then  $Y$  is isomorphic to an initial segment of  $X$ . Moreover,  $I$  is unique.*

*Proof.* To have  $f : Y \rightarrow X$  an isomorphism with an initial segment of  $X$ , we need precisely that for every  $x \in Y$  we have that  $f(x)$  is the minimum element of  $X \setminus \{f(y) : y < x\}$ . So we’re done by the previous theorem.  $\square$

**Note.** We have  $X \setminus \{f(y) : y < x\} \neq \emptyset$ , since  $f(y) \leq y$  for all  $y$  (induction), so  $x \notin \{f(y) : y < x\}$ .

In particular,  $X$  itself cannot be isomorphic to a proper initial segment (uniqueness).

## How do different well-orderings relate to each other?

**Definition.** For well-orderings  $X, Y$  we write  $X \leq Y$  if  $X$  is isomorphic to an initial segment of  $Y$ .

**Example.** If  $X = \mathbb{N}$ ,  $Y = (\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots)$ , then  $X \leq Y$ .

**Proposition 2.5.** *Let  $X, Y$  be well-orderings. Then  $X \leq Y$  or  $Y \leq X$ .*

*Proof.* Suppose  $Y \not\leq X$ , we'll show  $X \leq Y$ . To obtain  $f : X \rightarrow Y$  an isomorphism with an initial segment of  $Y$ , we need precisely that for every  $x \in X$ ,  $f(x)$  is the least element in  $Y \setminus \{f(y) : y < x\}$  [note this can only be empty if  $Y$  is isomorphic to  $I_x$ ]. So we're done by recursion.  $\square$

**Proposition 2.6.** *Let  $X, Y$  be well-orderings with  $X \leq Y$  and  $Y \leq X$ . Then  $X$  and  $Y$  are isomorphic.*

**Note.** This proposition and the previous one are “the most we could ever hope for”.

*Proof.* We have isomorphisms  $f$  from  $X$  to some initial segment of  $Y$ , and  $g$  from  $Y$  to some initial segment of  $X$ . Then  $g \circ f : X \rightarrow X$  is an isomorphism from  $X$  to an initial segment of  $X$  (as initial segment of an initial segment of  $X$  is itself an initial segment). So by uniqueness  $g \circ f = \text{id}_X$ . Similarly  $f \circ g = \text{id}_Y$ . Hence  $f$  and  $g$  are inverses, thus bijections.  $\square$

## New well-ordering from old

For well-orderings  $X, Y$ , we say  $X < Y$  if  $X \leq Y$  and  $X$  is not isomorphic to  $Y$ . Equivalently,  $X < Y$  if and only if  $X$  is isomorphic to a proper initial segment of  $Y$ .

We can ‘make a bigger one’: given a well-ordering  $X$ , pick some  $x \notin X$  and well-order  $X \cup \{x\}$  by setting  $y < x$  for all  $y \in X$ . This is a well-ordering and is  $> X$ . Call this the *successor* of  $X$ , written  $X^+$ .

We can ‘put some together’: given  $\{X_i\}_{i \in I}$  well-orderings, seek  $X$  with  $X \geq X_i$  for all  $i$ . For well-orderings  $(X, <_X), (Y, <_Y)$  we say  $Y$  *extends*  $X$  if  $X \subseteq Y$ ,  $<_Y|_X = <_X$ , and  $X$  is an initial segment of  $(Y, <_Y)$ . Say well-orderings  $\{X_i\}_{i \in I}$  are *nested* if for all  $i, j$ ,  $X_i$  extends  $X_j$  or  $X_j$  extends  $X_i$ .

**Proposition 2.7.** *Let  $\{X_i\}_{i \in I}$  be a nested set of well-orderings. Then there exists a well-ordering  $X$  such that  $X \geq X_i$  for all  $i$ .*

*Proof.* Let  $X = \bigcup_{i \in I} X_i$ , with ordering  $<_X = \bigcup_{i \in I} <_i$ , i.e.  $x < y$  in  $X$  if there exists  $i$  such that  $x, y \in X_i$  and  $x <_i y$ . Given  $S \subseteq X$  non-empty, we have  $S \cap X_i$  non-empty for some  $i \in I$ . Let  $x$  be the least element of  $S \cap X_i$  (under  $<_i$ ). Then  $x$  is the least element of  $S$  in  $X$  since  $X_i$  is an initial segment of  $X$ , by nestedness. So  $X$  is a well-ordering, and  $X \geq X_i$  for all  $i$ .  $\square$

**Remark.** The above proposition also holds if we don’t know the  $X_i$  are nested.

## Ordinals

“Does the collection of all well-orderings itself form a well-ordering?”

**Definition.** An *ordinal* is a well-ordered set, with two well-ordered sets regarded as the same if they are isomorphic.<sup>1</sup>

**Definition.** For a well-ordering  $X$ , corresponding to an ordinal  $\alpha$ , say  $X$  has *order-type*  $\alpha$ .

For any  $k \in \mathbb{N}$ , write  $k$  for the order-type of the (unique up to isomorphism) well-ordering on a set of size  $k$ . Write  $\omega$  for the order-type of  $\mathbb{N}$ .

**Example.** In  $\mathbb{R}$ :

- $\{-2, 3, \pi, 5\}$  has order-type 4;
- $\{1/2, 2/3, 3/4, \dots\}$  has order-type  $\omega$ .

<sup>1</sup>Just as a rational is an expression  $m/n$  with two regarded as the same if  $mn' = m'n$ . However, cannot formalise this using equivalence classes in the case of ordinals, see later chapter.

Write  $\alpha \leq \beta$  if  $X \leq Y$ , where  $X$  has order-type  $\alpha$  and  $Y$  has order-type  $\beta$  (note this is well defined since it doesn't depend on the choice of  $X, Y$ ). Similarly define  $\alpha < \beta$ ,  $\alpha^+$  etc.

Hence for all ordinals  $\alpha, \beta$ ,  $\alpha \leq \beta$  or  $\beta \leq \alpha$ . Also, if  $\alpha \leq \beta$  and  $\beta \leq \alpha$ ,  $\alpha = \beta$ .

**Proposition 2.8.** *For any ordinal  $\alpha$ , the ordinals  $< \alpha$  form a well-ordered set of order-type  $\alpha$ .*

*Proof.* Let  $X$  have order-type  $\alpha$ . Then the well-ordered sets  $< X$  are precisely (up to isomorphism) the proper initial segments of  $X$ , i.e they are  $I_x$  for  $x \in X$ . These order biject with  $X$  itself, via  $I_x \leftrightarrow x$ .  $\square$

So for any  $\alpha$ , have  $I_\alpha = \{\beta : \beta < \alpha\}$  a well-ordered set of order-type  $\alpha$ .

**Proposition 2.9.** *Every non-empty set  $S$  of ordinals has a least element.*

*Proof.* Choose  $\alpha \in S$ . If  $\alpha$  is minimal in  $S$ , we're done. Otherwise,  $S \cap I_\alpha$  is non-empty, so has a least element in  $I_\alpha$  since  $I_\alpha$  is well-ordered, and this element is least in all of  $S$ .  $\square$

However:

**Theorem 2.10** (Burali-Forti Paradox). *The ordinals do not form a set.*

*Proof.* Suppose  $X$  was the set of all ordinals. Then  $X$  is a well-ordered set, so has an order type, say  $\alpha$ . Thus  $X$  is order-isomorphic to  $I_\alpha$ , so  $X$  is order-isomorphic to a proper initial segment of itself, contradiction.  $\square$

**Note.** Given a set  $S = \{\alpha_i\}_{i \in I}$  of ordinals, there exists an upper bound  $\alpha$  for  $S$ , by applying proposition 2.7 to the nested family of the  $\{I_{\alpha_n}\}_{i \in I}$ . Hence by proposition 2.9 it has a least upper bound. We write  $\sup S$ .

**Example.**  $\sup\{2, 4, 6, \dots\} = \omega$ .

We'll give some examples of ordinals

**Examples.**

- $0, 1, \dots, \omega, \omega + 1, \omega + 2, \omega + 3, \dots, \omega + \omega$  (really  $\omega + 1$  is  $\omega^+$  and  $\omega + \omega = \omega \cdot 2 = \sup\{\omega, \omega + 1, \dots\}$ ).
- Continuing with  $\omega \cdot 2, \omega \cdot 2 + 1, \omega \cdot 2 + 2, \dots, \omega \cdot 3, \dots, \omega \cdot 4, \dots, \omega \cdot \omega^2$  where  $\omega^2 = \omega \cdot \omega = \sup\{\omega, \omega \cdot 2, \omega \cdot 3, \dots\}$ .
- Now  $\omega^2, \omega^2 + 1, \dots, \omega^2 + \omega$  and  $\omega^2 + \omega \cdot 2, \omega^2 + \omega \cdot 3, \dots, \omega^2 + \omega^2 = \omega^2 \cdot 2$ .
- $\omega^2 \cdot 2, \omega^2 \cdot 3, \dots, \omega^3$ .
- $\omega^3, \dots, \omega^3 + \omega^2 \cdot 7 + \omega \cdot 4 + 13$ .
- $\omega^\omega = \sup\{\omega, \omega^2, \omega^3, \dots\}$
- $\omega^{\omega+1} = \sup\{\omega^\omega + 1, \omega^\omega + 2, \dots\}$
- $\omega^{\omega \cdot 2}, \omega^{\omega \cdot 3}, \dots, \omega^{\omega^2}$ .
- $\omega^{\omega^\omega}$
- $\omega^{\omega^{\omega^2}}, \omega^{\omega^{\omega^4}}$
- $\omega^{\omega^\omega} = \varepsilon_0 = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\}$ .
- $\varepsilon_0, \varepsilon_0 + 1, \dots, \varepsilon_0 + \omega, \dots, \varepsilon_0 + \varepsilon_0$
- $\varepsilon_0 \cdot \omega, \dots, \varepsilon_0^2$
- $\varepsilon_0^{\varepsilon_0} = \sup\{\varepsilon_0^\omega, \varepsilon_0 \omega^\omega, \varepsilon_0^{\omega^\omega}\}$
- $\varepsilon_1 = \sup\{\varepsilon_0, \varepsilon_0^{\varepsilon_0}, \varepsilon_0^{\varepsilon_0^{\varepsilon_0}}\}$ .

All of the above are countable (e.g countable union of countable sets). Is there an uncountable ordinal? i.e is there is there an uncountable well-ordering. e.g can well-order  $\mathbb{N}$ , can well-order  $\mathbb{Q}$  (bijection with  $\mathbb{N}$ ), can we well-order  $\mathbb{R}$ ? Amazingly, we can prove we can.

**Theorem 2.11.** *There is an uncountable ordinal.*

*Proof.* Let  $A = \{R \in \mathcal{P}(\omega \times \omega) : R \text{ is a well-ordering of a subset of } \omega\}$ . Let  $B = \{\text{order-type}(R) : R \in A\}$ . So  $\alpha \in B$  if and only if  $\alpha$  is a countable ordinal. Let  $\omega_1 = \sup B$ . We must have  $\omega_1$  uncountable - if it was countable, then it would be the greatest element of  $B$ , contradicting  $\omega_1 < \omega_1^+$  since  $\omega_1^+$  is countable.  $\square$



**Remark.** Alternatively having the set  $B$ , could say that  $B$  isn't all ordinals since the set of ordinals is not a set (Burali-Forti), so there exists an uncountable ordinal.

**Note.**  $\omega_1$  is the least uncountable ordinal by definition of  $B$ .

The ordering  $\omega_1$  has some remarkable properties, e.g

1.  $\omega_1$  is uncountable but  $\{\beta : \beta < \alpha\}$  is countable for all  $\alpha < \omega_1$ .
2. Any sequence  $\alpha_1, \alpha_2, \dots$  in  $I_{\omega_1}$  is bounded. Namely, by  $\sup\{\alpha_1, \alpha_2, \dots\}$  which is countable as a countable union of countable sets.

The same argument shows:

**Theorem** (Hartogs' Lemma). *For every set  $X$ , there exists an ordinal  $\alpha$  that does not inject into  $X$ .*

We call the least such ordinal as in Hartogs' Lemma  $\gamma(X)$ , e.g  $\gamma(\omega) = \omega_1$ .

**Definition.** Say  $\alpha$  is a *successor* if there exists  $\beta$  such that  $\alpha = \beta^+$ . Otherwise we say  $\alpha$  is a *limit*.

Note that  $\alpha$  has a greatest element if and only if it is a successor. So  $\alpha$  is a limit if and only if  $\alpha$  has no greatest element, i.e for all  $\beta < \alpha$  there exists  $\gamma < \alpha$  with  $\beta < \gamma$ .

**Example.** 5 is a successor:  $4^+$ .  $\omega+2$  is a successor:  $(\omega^+)^+$ .  $\omega$  is not a successor: no greatest element. 0 is also a limit.

## Ordinal Arithmetic

We define  $\alpha + \beta$  by induction on  $\beta$  ( $\alpha$  fixed) by:

- $\alpha + 0 = \alpha$ ;
- $\alpha + (\beta^+) = (\alpha + \beta)^+$ ;
- $\alpha + \lambda = \sup\{\alpha + \gamma : \gamma < \lambda\}$  for  $\lambda$  a non-zero limit.

### Examples.

- $\omega + 1 = \omega + 0^+ = (\omega + 0)^+ = \omega^+$ ;
- $\omega + 2 = \omega + 1^+ = (\omega + 1)^+ = \omega^{++}$ ;
- $1 + \omega = \sup\{1 + \gamma : \gamma < \omega\} = \omega$  - so  $+$  is not commutative.

**Remark.** Officially (as the ordinals do not form a set), this means: to define  $\alpha + \beta$  we actually define  $\alpha + \gamma$  on  $\{\gamma : \gamma \leq \beta\}$ , which is a set; plus uniqueness. Similarly, for proof by induction: if for some  $\alpha$  we have  $p(\alpha)$  false, then on  $\{\gamma : \gamma \leq \alpha\}$ ,  $p$  is not everywhere true.

**Proposition 2.12.** *We have  $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$  for all ordinals  $\alpha, \beta, \gamma$ .*

*Proof.* We proceed by induction on  $\gamma$  ( $\alpha, \beta$  fixed). If  $\gamma = 0$ :  $\alpha + (\beta + 0) = \alpha + \beta = (\alpha + \beta) + 0$ .

Successors:

$$\begin{aligned} \alpha + (\beta + \gamma^+) &= \alpha + (\beta + \gamma)^+ \\ &= (\alpha + (\beta + \gamma))^+ \\ &= ((\alpha + \beta) + \gamma)^+ \\ &= (\alpha + \beta) + \gamma^+ \end{aligned}$$

$\lambda$  a non-zero limit:

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

We claim that  $\beta + \lambda$  is a limit. Indeed, have  $\beta + \lambda = \sup\{\beta + \gamma : \gamma < \lambda\}$ . But for every  $\gamma < \lambda$ , there exists  $\gamma' < \lambda$  with  $\gamma < \gamma'$  ( $\lambda$  a limit), so  $\beta + \gamma < \beta + \gamma'$ . Thus there is no greatest element of  $\{\beta + \gamma : \gamma < \lambda\}$ , so  $\beta + \lambda = \sup\{\beta + \gamma : \gamma < \lambda\}$  is a limit.

Therefore  $\alpha + (\beta + \lambda) = \sup\{\alpha + \delta : \delta < \beta + \lambda\}$ . So need to show  $\sup\{\alpha + (\beta + \gamma) : \gamma < \lambda\} = \sup\{\alpha + \delta : \delta < \beta + \lambda\}$ . Indeed,  $\gamma < \lambda$  implies  $\beta + \gamma < \beta + \lambda$  so  $\{\alpha + (\beta + \gamma) : \gamma < \lambda\} \subseteq \{\alpha + \delta : \delta < \beta + \lambda\}$ . Conversely, for all  $\delta < \beta + \lambda$ , we have  $\delta \leq \beta + \gamma$  for some  $\gamma < \lambda$  (definition of  $\beta + \lambda$ ) so  $\alpha + \delta \leq \alpha + (\beta + \gamma)$ . So each member of right hand set is at most some member of the left hand set.  $\square$

**Notes.**

1. We used:  $\beta \leq \gamma \Rightarrow \alpha + \beta \leq \alpha + \gamma$  (trivial by induction on  $\gamma$ )
2.  $\beta < \gamma \Rightarrow \alpha + \beta < \alpha + \gamma$  since  $\beta < \gamma \Rightarrow \beta^+ \leq \gamma$  which implies  $\alpha + \beta^+ \leq \alpha + \gamma$  so  $\alpha + \beta < (\alpha + \beta)^+ = \alpha + \beta^+ \leq \alpha + \gamma$ .
3. However  $1 < 2$ , but  $1 + \omega = 2 + \omega = \omega$ . So “stuff on the right always works as expected”.

The above is the inductive definition of  $+$ . There is also a synthetic definition of  $+$ :  $\alpha + \beta$  is the order type of  $\alpha \cup \beta$  (disjoint union, e.g.  $(\alpha \times \{0\}) \cup (\beta \times \{1\})$ ), with all of  $\alpha$  coming before all of  $\beta$ .

**Example.**

- $\omega + 1$  is the order type of  $\underbrace{\omega}_{\text{sequence}} \underbrace{\bullet}_{\text{point}};$
- $1 + \omega$  is the order type of  $\underbrace{\bullet}_{\text{point}} \underbrace{\omega}_{\text{sequence}};$
- $\alpha + (\beta + \gamma)$  is the order type of  $\underbrace{\alpha}_{\text{sequence}} \underbrace{\beta}_{\text{sequence}} \underbrace{\gamma}_{\text{sequence}}.$

**Proposition 2.13.** *The two definitions of  $+$  are equivalent.*

*Proof.* We write  $+$  for the inductively defined one, and  $+'$  for the synthetic one. We'll show  $\alpha + \beta = \alpha +' \beta$  for all  $\alpha + \beta$  by induction on  $\beta$  ( $\alpha$  fixed).  
Zero:  $\alpha + 0 = \alpha +' 0 = 0 = \alpha$ .

Successors:  $\alpha + (\beta^+) = (\alpha + \beta)^+ = (\alpha +' \beta)^+$  which is the order type of  $\underbrace{\alpha}_{\text{sequence}} \underbrace{\beta}_{\text{sequence}} \underbrace{\bullet}_{\text{point}}$  which is  $\alpha +' \beta^+$ .

$\lambda$  a non-zero limit:  $\alpha + \lambda = \sup\{\alpha + \gamma : \gamma < \lambda\} = \sup\{\alpha +' \gamma : \gamma < \lambda\} = \alpha +' \lambda$  (since sup is a union as sets are nested)  $\square$

Moral: synthetic definition beats the inductive one, if we do have a synthetic definition.

**Definition.** Define  $\alpha\beta$  ( $\alpha$  fixed, recursion on  $\beta$ ) by:

- $\alpha 0 = 0$ ;
- $\alpha(\beta^+) = \alpha\beta + \alpha$ ;
- $\alpha\lambda = \sup\{\alpha\gamma : \gamma < \lambda\}$  for  $\lambda$  a non-zero limit.

**Examples.**

- $\omega 2 = \omega 1 + \omega = (\omega 0 + \omega) + \omega = \omega + \omega$ ;

- $\omega 3 = \omega + \omega + \omega$ ;
- $\omega\omega = \sup\{0, \omega, \omega + \omega, \dots\}$ ;
- $2\omega = \sup\{0, 2, 4, 6, 8, \dots\} = \omega$ , so again this is not commutative.

Can show that  $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ , etc.

We also have a synthetic definition (which can be shown to be equivalent):  $\alpha\beta$  is equal to the order type of

$$\underbrace{\underbrace{\alpha}_{\text{}} \underbrace{\alpha}_{\text{}} \underbrace{\alpha}_{\text{}} \dots \underbrace{\alpha}_{\text{}}}_{\beta \text{ times}},$$

ordered by:  $(x, y) < (z, w)$  if  $y < w$  or  $y = w$  and  $x < z$ .

**Example.**  $\omega 2$  is the order type of  $\underbrace{\omega}_{\text{}} \underbrace{\omega}_{\text{}}$  which is  $\omega + \omega$ . Also  $2\omega$  is the order type of

$$\underbrace{\underbrace{\bullet \bullet}_{\text{}} \underbrace{\bullet \bullet}_{\text{}} \underbrace{\bullet \bullet}_{\text{}} \dots \underbrace{\bullet \bullet}_{\text{}}}_{\omega \text{ times}},$$

which is  $\omega$ .

We can also do exponentiation, towers etc similarly. For example, define  $\alpha^\beta$  by

- $\alpha^0 = 1$ ;
- $\alpha^{(\beta^+)} = \alpha^\beta \alpha$ ;
- $\alpha^\lambda = \sup\{\alpha^\gamma : \gamma < \lambda\}$  for  $\lambda$  a non-zero limit.

For example,  $\omega^2 = \omega^1 \omega = (\omega^0 \omega) \omega = \omega \omega$ ;  $2^\omega = \sup\{2^0, 2^1, \dots\} = \omega$ .

### 3 Posets and Zorn's Lemma

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

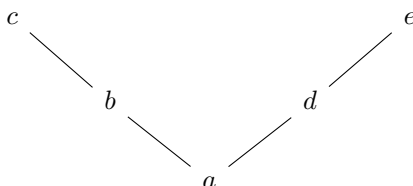
- (i) Reflexive:  $x \leq x$  for all  $x \in X$ ;
- (ii) Transitive:  $x \leq y, y \leq z$  implies  $x \leq z$  for all  $x, y, z \in X$ ;
- (iii) Antisymmetric:  $x \leq y, y \leq x$  implies  $x = y$  for all  $x, y$ .

We write  $x < y$  if  $x \leq y$  and  $x \neq y$ . In terms of  $<$ , a poset is:

- (i) Irreflexive:  $x \not< x$  for all  $x \in X$ ;
- (ii) Transitive:  $x < y, y < z$  implies  $x < z$  for all  $x, y, z \in X$ .

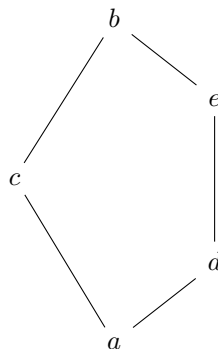
**Examples.**

1. Any total order.
2.  $\mathbb{N}^+$  with  $x \leq y$  if  $x|y$ .
3.  $(\mathcal{P}(S), \subseteq)$  for any set  $S$ .
4.  $X \subseteq \mathcal{P}(S)$  under  $\subseteq$ . For example,  $V$  a vector space,  $X$  the set of all subspaces.
- 5.

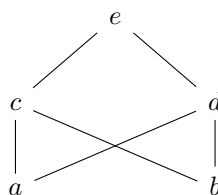


Consider a tree graph, with all edges pointing upwards. Then we say  $x \leq y$  for vertices  $x, y$  if the  $xy$  is a directed edge pointing upwards, and extend  $\leq$  by transitivity. In general the *Hasse diagram* of a poset is a drawing of its posets, with an upwards line from  $x$  to  $y$  if  $y$  covers  $x$  (meaning  $y > x$  and no  $z$  has  $y > z > x$ ). Hasse diagrams can be useful: e.g  $(\mathbb{N}, \leq)$ , or useless: e.g for  $(\mathbb{Q}, \leq)$  there are no covers.

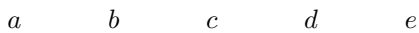
6.



7.



8.



A subset  $S$  of a poset  $X$  is a *chain* if it is totally ordered. E.g in example 2 above,  $\{1, 2, 4, 8, 16, \dots\}$ . Or in example 5,  $\{a, b, c\}$  or  $\{a, c\}$ .

A subset  $S$  is an *antichain* if no two elements are related. E.g in 2,  $\{n : n \text{ prime}\}$ , in 5,  $\{c, e\}$ , or in 8 take whole poset.

For  $S \subseteq X$ , an *upper bound* for  $S$  is an  $x \in X$  such that  $x \geq y$  for all  $y \in S$ . We say  $x$  is a *least upper bound* for  $S$  if  $x$  is an upper bound, and if  $y$  is an upper bound of  $S$ ,  $x \leq y$ .

### Examples.

- In  $\mathbb{R}$ : if  $S = \{x : x < \sqrt{2}\}$  then  $\sqrt{2}$  is an upper bound, and  $\sqrt{2}$  is the least upper bound. We write  $\sqrt{2} = \sup S$ , or  $\bigvee S$ .
- In  $\mathbb{Q}$ :  $\{x : x^2 < 2\}$  has  $\sqrt{2}$  as an upper bound, but there is no least upper bound.
- In example 5 from before,  $\{a, b\}$  has upper bounds  $b$  and  $c$ , so least upper bound  $b$ .  $\{b, d\}$  has no upper bound.
- From example 7 from before,  $\{a, b\}$  has upper bounds  $c, d, e$ , so does not have a least upper bound.

We say  $X$  is *complete* if every  $S \subseteq X$  has a least upper bound. For example,  $\mathbb{R}$  is not complete, e.g.  $\mathbb{Z}$  has no upper bound.  $(0, 1)$  is not complete since  $(0, 1)$  itself has no upper bound.

$X = \mathcal{P}(S)$  is always complete:  $\sup$  of  $\{A_i : i \in I\}$  is  $\bigcup_{i \in I} A_i$ .

**Note.** Every complete poset  $X$  has a greatest element  $x$ , namely  $\sup X$ , and also a least element  $y$ , namely  $\sup\{\emptyset\}$ .

For  $f : X \rightarrow Y$  where  $X, Y$  are posets, we say  $f$  is *order preserving* if  $x \leq y$  implies  $f(x) \leq f(y)$ .

**Examples.**

1.  $f : \mathbb{N} \rightarrow \mathbb{N}, x \mapsto x + 1$ .
2.  $f : [0, 1] \rightarrow [0, 1], x \mapsto \frac{1+x}{2}$ .
3.  $f : \mathcal{P}(S) \rightarrow \mathcal{P}(S), A \mapsto A \cup \{i\}$  for some fixed  $i \in S$ .

Not every order preserving  $f : X \rightarrow X$  has a fixed point, e.g. example 1 above. However:

**Theorem 3.1** (Knaster-Tarski Fixed Point Theorem). *Let  $X$  be a complete poset. Then any order preserving  $f : X \rightarrow X$  has a fixed point.*

*Proof.* Let  $E = \{x \in X : x \leq f(x)\}$ , and let  $s = \sup E$ . We'll show that  $f(s) = s$ .

We'll first show  $s \leq f(s)$ . It is enough to show  $f(s)$  is an upper bound for  $E$ , then done since  $s$  is a least upper bound for  $E$ . Indeed, if  $x \in E$ , then  $x \leq s$  so  $f(x) \leq f(s)$ . Now since  $x \in E$ ,  $x \leq f(x) \leq f(s)$ .

Now we show  $f(s) \leq s$ . It is enough to show  $f(s) \in E$ , then done since  $s$  is an upper bound for  $E$ . We have  $s \leq f(s)$ , so  $f(s) \leq f(f(s))$ . i.e.  $f(s) \in E$ .  $\square$

**Remark.** We need to show  $s \leq f(s)$  before  $f(s) \leq s$  since  $s \leq f(s)$  says  $s \in E$ .

An application of this is:

**Corollary 3.2** (Schröder-Bernstein). *Let  $f : A \rightarrow B$  and  $g : B \rightarrow A$  be injections. Then there exists a bijection  $h : A \rightarrow B$ .*

*Proof.* We seek partitions  $A = P \cup Q, P \cap Q = \emptyset, B = R \cup S, R \cap S = \emptyset$  such that  $f(P) = R$  and  $g(S) = Q$  (then set  $h = f$  on  $P$  and  $h = g^{-1}$  on  $R$ ). Thus we seek exactly a fixed point of  $\Theta : \mathcal{P}(A) \rightarrow \mathcal{P}(A)$  given by  $P \mapsto A \setminus g(B \setminus f(P))$ . But  $\mathcal{A}$  is complete, and  $\Theta$  is order preserving, so done by Knaster-Tarski.  $\square$

## Zorn's Lemma

For a poset  $X$ ,  $x \in X$  is said to be *maximal* if no  $y \in X$  has  $y > x$ . For example, in  $[0, 1]$ , 1 is maximal. We've seen many posets without any maximal element, for example  $(\mathbb{R}, \leq)$  or  $(\mathbb{N}^+, |)$ . In each case, there exists a chain with no upper bound (e.g in  $(\mathbb{N}^+, |)$  take powers of 2).

**Theorem 3.3** (Zorn's Lemma). *Let  $X$  be a (non-empty) poset in which every chain has an upper bound. Then there exists a maximal element of  $X$ .*

*Proof.* Suppose not. So for each  $x \in X$  have  $x' \in X$  with  $x' > x$ , and for each chain we  $C$  we have an upper bound  $u(C)$ . Fix some  $x \in X$  and define  $x_\alpha$  for each  $\alpha < \gamma(X)$  by recursion:  $x_0 = x$ ,  $x_{\alpha+1} = x'_\alpha$  and  $x_\lambda = U(\{x_\beta : \beta < \lambda\})$  for  $\lambda$  a non-zero limit [note that the  $x_\beta$ ,  $\beta < \lambda$  do form a chain, by induction]. Then we have injected  $\gamma(X)$  into  $X$ , a contradiction.  $\square$

**Remark.** The proof was easy, given well-orderings; recursion and Harthogs' from Chapter 2.

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

### Examples.

- $\mathbb{R}^3$  has a basis  $\{e_1, e_2, e_3\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ .
- Space of all real polynomials has basis  $\{1, X, X^2, X^3, \dots\}$ .
- Space  $S$  of all real sequences -  $\{e_i\}_{i \in \mathbb{N}}$  is not a basis - span doesn't contain  $(1, 1, 1, \dots)$ . In fact  $S$  has no countable basis (and can actually show there's no explicit basis).
- $\mathbb{R}$  as a vector space over  $\mathbb{Q}$  - no explicit basis. A basis in this case is called a *Hamel basis*.

**Theorem 3.4.** *Every vector space  $V$  has a basis.*

*Proof.* Let  $X = \{A \subseteq V : A \text{ is linearly independent}\}$ , ordered by  $\subseteq$ . We seek a maximal element of  $X$  [then done since if a maximal element doesn't span, could extend by adding something not in span]. Have  $X \neq \emptyset$  since  $\emptyset \in X$ . Given a chain  $\{A_i : i \in \mathbb{I}\}$ , let  $A = \bigcup_{i \in \mathbb{I}} A_i$ . Certainly  $A \supseteq A_i$  for every  $i$ , so we just need to show  $A \in X$ , i.e  $A$  is linearly independent. Suppose  $A$  is not linearly independent. Then we have  $x_1, \dots, x_n \in A$  which are linearly dependent. We have  $x_1 \in A_{i_1}, \dots, x_n \in A_{i_n}$  for some  $i_1, \dots, i_n \in \mathbb{I}$ , but some  $A_{i_k}$  contains all of  $A_{i_1}, \dots, A_{i_n}$  (since the  $A_i$  form a chain), so  $x_1, \dots, x_n \in A_{i_k}$ , contradicting the fact  $A_{i_k}$  is linearly independent.

Hence by Zorn's Lemma, there exists a maximal element of  $X$ .  $\square$

### Notes.

1. The only actual linear algebra we did was in the 'then done' part.



2. In the statement of Zorn's Lemma, the hypothesis  $X \neq \emptyset$  is not strictly needed since  $\emptyset$  is not a chain so has an upper bound.

Another application: completeness theorem for propositional logic with no restriction on  $P$ .

**Theorem 3.5.** *Let  $S \subseteq L = L(P)$  (for any set  $P$ ) be a set that is consistent. Then  $S$  has a model.*

*Proof.* We'll extend  $S$  to  $\bar{S}$  (consistent) such that for all  $t \in L$ ,  $t \in \bar{S}$  or  $(\neg t) \in \bar{S}$  - then done as in Chapter 1 by setting  $v = 1$  on  $\bar{S}$ ,  $v = 0$  elsewhere.

Let  $X = \{T \supseteq S : T \text{ consistent}\}$ , ordered by  $\subseteq$ . We seek a maximal element of  $X$  [then done: let  $\bar{S}$  be maximal, then if  $t \notin \bar{S}$ , we must have  $\bar{S} \cup \{t\} \vdash \perp$  (maximality of  $\bar{S}$ ), so  $\bar{S} \vdash (\neg t)$  (deduction theorem), so  $(\neg t) \in \bar{S}$  by maximality of  $\bar{S}$ ].

Now,  $X \neq \emptyset$  since  $S \in X$ . Given a non-empty chain  $\{T_i : i \in I\}$ , put  $T = \bigcup_{i \in I} T_i$ . Have  $T \supseteq T_i$  for all  $i$ , so we just need to show  $T \in X$ . Indeed  $S \subseteq T$  (chain is non-empty). Also, we claim  $T$  is consistent. Suppose not, i.e.  $T \vdash \perp$ . Then some  $\{t_1, \dots, t_n\} \vdash \perp$  for some  $t_1, \dots, t_n \in T$  (as proofs are finite). Now,  $t_1 \in T_{i_1}, \dots, t_n \in T_{i_n}$  for some  $i_1, \dots, i_n \in I$ . Then  $t_1, \dots, t_n \in T_{i_k}$  for some  $k$  since the  $T_i$  form a chain. But this contradicts the fact  $T_{i_k}$  is consistent.

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