# Part II — Logic and Set Theory

Based on lectures by Dr Zsak and notes by thirdsgames.co.uk

Lent 2023

## Contents


**1 Propositional logic** **2**
- 1.1 Languages . . . 2
- 1.2 Semantic implication . . . 3
- 1.3 Syntactic implication . . . 5
- 1.4 Deduction theorem . . . 7
- 1.5 Soundness . . . 8
- 1.6 Adequacy . . . 8
- 1.7 Completeness . . . 10

**2 Well-Orderings** **12**
- 2.1 Definition . . . 12
- 2.2 Initial segments . . . 15
- 2.3 Relating well-orderings . . . 17
- 2.4 Constructing larger well-orderings . . . 17
- 2.5 Ordinals . . . 18
- 2.6 Some ordinals . . . 20
- 2.7 Uncountable ordinals . . . 20
- 2.8 Successors and limits . . . 21
- 2.9 Ordinal arithmetic . . . 22

# §1 Propositional logic

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

## §1.1 Languages

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

1. if $p \in P$, then $p \in L$;
2. $\perp \in L$, where the symbol $\perp$ is read 'false'/ 'bottom';
3. if $p, q \in L$, then $(p \Rightarrow q) \in L$.

**Example 1.1**
$((p_1 \Rightarrow p_2) \Rightarrow (p_1 \Rightarrow p_3)) \in L$. $(p_4 \Rightarrow \perp) \in L$.

If $p \in L$ then $((p \Rightarrow \perp) \Rightarrow \perp) \in L$.

*Remark* 1. Note that the phrase '$L$ is defined inductively' means more precisely the following. Let $L_1 = P \cup \{\perp\}$, and define $L_{n+1} = L_n \cup \{(p \Rightarrow q) : p, q \in L_n\}$. We set $L = \bigcup_{n=1}^{\infty} L_n$.

Note that the elements of $L$ are just finite strings of symbols from the alphabet $P \cup \{(,), \Rightarrow, \perp\}$. Brackets are only given for clarity; we omit those that are unnecessary, and may use other types of brackets such as square brackets.
We can prove that $L$ is the smallest (w.r.t. inclusion) subset of the set $\Sigma$ of all finite strings in $P \cup \{(,), \Rightarrow, \perp\}$ s.t. the properties of a language hold.
Note that $L \subsetneq \Sigma$. E.g. $\Rightarrow p_1 p_3 (\in \Sigma \setminus L$.

Note that the introduction rules for the language are injective and have disjoint ranges, so there is exactly one way in which any element of the language can be constructed using rules (i) to (iii).

Every $p \in L$ is uniquely determined by the properties of a language above, i.e. either $p \in P$ or $p = \perp$ or $\exists$ unique $q, r \in L$ s.t. $p = (q \Rightarrow r)$.

We can now introduce the abbreviations $\neg, \wedge, \vee, \top$, which are not, and, or and true/top respectively, defined by

**Notation.**

$$\neg p = (p \Rightarrow \perp); \quad p \vee q = \neg p \Rightarrow q; \quad p \wedge q = \neg(p \Rightarrow \neg q), \top = (\perp \Rightarrow \perp)$$

## §1.2 Semantic implication

**Definition 1.1** (Valuation)
A **valuation** is a function $v\colon L \to \{0,1\}$ s.t.

1. $v(\perp) = 0$;
2. If $p, q \in L$ then $v(p \Rightarrow q) = \begin{cases} 0 & v(p) = 1 \text{ and } v(q) = 0 \\ 1 & \text{else} \end{cases}$

**Example 1.2**
If $v(p_1) = 1$, $v(p_2) = 0$. Then

$$v\left(\underbrace{(\perp \Rightarrow p_1)}_{1} \Rightarrow \underbrace{(p_1 \Rightarrow p_2)}_{0}\right) = 0$$

*Remark* 2. On $\{0,1\}$, we can define the constant $\perp = 0$ and the operation $\Rightarrow$ in the obvious way. Then, a valuation is precisely a mapping $L \to \{0,1\}$ preserving all structure, so it can be considered a homomorphism.

**Proposition 1.1**
Let $v, v'\colon L \to \{0,1\}$ be valuations that agree on the primitives $p_i$. Then $v = v'$. Further, any function $w\colon P \to \{0,1\}$ extends to a valuation $v : L \to \{0,1\}$ s.t. $v :_P = w$.

*Remark* 3. This is analogous to the definition of a linear map by its action on the basis vectors.

*Proof.* Clearly, $v, v'$ agree on $L_1$ as $v(\perp) = v'(\perp) = 0$, the set of elements of the language of length 1. If $v, v'$ agree at $p, q \in L_n$, then they agree at $p \Rightarrow q$. So by induction, $v, v'$ agree on $L_{n+1}$ for all $n$, and hence on $L$.

Let $v(p) = w(p)$ for all $p \in P$, and $v(\perp) = 0$ to obtain $v$ on the set $L_1$. Assuming $v$ is defined on $p, q \in L_n$ we can define it at $p \Rightarrow q$ in the obvious way. This defines $v$ on $L_{n+1}$, hence $v$ is defined on $\cup L_n = L$. By construction, $v$ is a valuation on $L$ and $v :_P = w$. $\square$

**Example 1.3**

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

**Definition 1.2** (Tautology)

A **tautology** is $t \in L$ s.t. $v(t) = 1$ $\forall$ valuations $v$. We write $\models t$.

**Example 1.4**

$p \Rightarrow (q \Rightarrow p)$ (a true statement is implied by any true statement).

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

Since the right-hand column is always 1, $\models p \Rightarrow (q \Rightarrow p)$.

**Example 1.5** (Law of Excluded Middle)

$\neg\neg p \Rightarrow p$, which expands to $((p \Rightarrow \perp) \Rightarrow \perp) \Rightarrow p$.

| $v(p)$ | $v(\neg p)$ | $v(\neg\neg p)$ | $v(\neg\neg p \Rightarrow p)$ |
|---|---|---|---|
| 0 | 1 | 0 | 1 |
| 1 | 0 | 1 | 1 |

Hence $\models \neg\neg p \Rightarrow p$.

**Example 1.6**

$\neg p \vee p$, which expands to $((p \Rightarrow \perp) \vee p)$.

| $v(p)$ | $v(\neg p)$ | $v(\neg p \vee p)$ |
|---|---|---|
| 0 | 1 | 1 |
| 1 | 0 | 1 |

Hence $\models \neg p \vee p$.

**Example 1.7**

$(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$. Suppose this is not a tautology. Then we have a valuation $v$ s.t. $v(p \Rightarrow (q \Rightarrow r)) = 1$ and $v((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) = 0$. Hence, $v(p \Rightarrow q) = 1, v(p \Rightarrow r) = 0$, so $v(p) = 1, v(r) = 0$, giving $v(q) = 1$, but then $v(p \Rightarrow (q \Rightarrow r)) = 0$ contradicting the assumption.

**Definition 1.3** (Semantic Implication)

Let $S \subseteq L$ and $t \in L$. We say $S$ **entails** or **semantically implies** $t$, written $S \models t$, if for every valuation $v$ on $L$, $v(s) = 1 \ \forall \ s \in S \Rightarrow v(t) = 1$.

**Example 1.8**

$\{p, p \Rightarrow q\} \models q$.

**Example 1.9**

Let $S = \{p \Rightarrow q, q \Rightarrow r\}$, and let $t = p \Rightarrow r$. Suppose $S \not\models t$, so there is a valuation $v$ s.t. $v(p \Rightarrow q) = 1, v(q \Rightarrow r) = 1, v(p \Rightarrow r) = 0$. Then $v(p) = 1, v(r) = 0$, so $v(q) = 1$ and $v(q) = 0$ ↯.

**Definition 1.4** (Model)

Given $t \in L$, say a valuation **$v$ is a model for $t$** (or **$t$ is true in $v$**) if $v(t) = 1$.

**Definition 1.5** (Model)

We say that **$v$ is a model of $S$** in $L$ if $v(s) = 1$ for all $s \in S$.

Thus, $S \models t$ is the statement that every model of $S$ is also a model of $t$/ $t$ is true in every model of $S$.

*Remark* 4. The notation $\models t$ is equivalent to $\varnothing \models t$.

## §1.3 Syntactic implication

For a notion of proof, we require a system of axioms and deduction rules. As axioms, we take (for any $p, q, r \in L$),

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

2. $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$;
3. $((p \Rightarrow \bot) \Rightarrow \bot) \Rightarrow p$.

*Remark* 5. Sometimes, these three axioms are considered axiom **schemes**, since they are really a different axiom for each $p, q, r \in L$.

These are all tautologies.

For deduction rules, we will have only the rule **modus ponens (MP)**, that from $p$ and $p \Rightarrow q$ one can deduce $q$.

**Definition 1.6** (Proof)
Let $S \subseteq L$, $t \in L$. A **proof of $t$ from $S$** is a finite sequence $t_1, \ldots, t_n$ of propositions in $L$ s.t. $t_n = t$ and every $t_i$ is either

1. an axiom;
2. an element of $S$ ($t_i$ is a premise or hypothesis); or
3. follows by MP, where $t_j = p$ and $t_k = p \Rightarrow q$ where $j, k < i$.

We say that $S$ is the set of **premises** or **hypotheses**, and $t$ is the **conclusion**.

We say $S$ **proves** or **syntactically implies** $t$, written $S \vdash t$, if there exists a proof of $t$ from $S$.

**Example 1.10**
We will show $\{p \Rightarrow q, q \Rightarrow r\} \vdash (p \Rightarrow r)$.

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

**Definition 1.7** (Theorem)
If $\varnothing \vdash t$, we say $t$ is a **theorem**, written $\vdash t$.

**Example 1.11**

$\vdash (p \Rightarrow p)$.

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

## §1.4 Deduction theorem

**Theorem 1.1** (Deduction Theorem)

Let $S \subseteq L$, and $p, q \in L$. Then $S \vdash (p \Rightarrow q)$ iff $S \cup \{p\} \vdash q$.

*Remark* 6. This show '$\Rightarrow$' really does behave like implication in formal proofs.

*Proof.* ($\Rightarrow$): Given a proof of $p \Rightarrow q$ from $S$, add the line $p$ to the hypothesis and deduce $q$ from modus ponens, to obtain a proof of $q$ from $S \cup \{p\}$.

($\Leftarrow$): Suppose we have a proof of $q$ from $S \cup \{p\}$. Let $t_1, \ldots, t_n$ be the lines of the proof. We will prove that $S \vdash (p \Rightarrow t_i)$ for all $i$ by induction.

- If $t_i$ is an axiom, we write $t_i$ (axiom); $t_i \Rightarrow (p \Rightarrow t_i)$ (axiom 1); $p \Rightarrow t_i$ (modus ponens).
- If $t_i \in S$, we write $t_i$ (hypothesis); $t_i \Rightarrow (p \Rightarrow t_i)$ (axiom 1); $p \Rightarrow t_i$ (modus ponens).
- If $t_i = p$, we write the proof of $\vdash p \Rightarrow p$ given above.
- Suppose $t_i$ is obtained by modus ponens from $t_j$ and $t_k = t_j \Rightarrow t_i$ where $j, k < i$. We may assume by induction that $S \vdash p \Rightarrow t_j$ and $S \vdash p \Rightarrow (t_j \Rightarrow t_i)$. We write

  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)

  giving $S \vdash p \Rightarrow t_i$.

$\square$

**Example 1.12**

Consider $\{p \Rightarrow q, q \Rightarrow r\} \vdash p \Rightarrow r$. By the Deduction Theorem, it suffices to prove $\{p \Rightarrow q, q \Rightarrow r, p\} \vdash r$, which is obtained easily from modus ponens.

## §1.5 Soundness

We aim to show $S \models t$ iff $S \vdash t$. The direction $S \vdash t$ implies $S \models t$ is called **soundness**, which is a way of verifying that our axioms and deduction rule make sense. The direction $S \models t$ implies $S \vdash t$ is called **adequacy**, which states that our axioms are powerful enough to deduce everything that is (semantically) true.

**Proposition 1.2** (Soundness Theorem)

Let $S \subseteq L$ and $t \in L$. Then $S \vdash t$ implies $S \models t$.

*Proof.* We have a proof $t_1, \dots, t_n$ of $t$ from $S$. We aim to show that any model of $S$ is also a model of $t$, so if $v$ is a valuation that maps every element of $S$ to 1, then $v(t) = 1$.

We show this by induction on the length of the proof. $v(p) = 1$ for each axiom $p$ (as axioms are tautologies) and for each $p \in S$. Further, $v(t_i) = 1, v(t_i \Rightarrow t_j) = 1$, then $v(t_j) = 1$. Therefore, $v(t_i) = 1$ for all $i$. □

## §1.6 Adequacy

Consider the case of adequacy where $t = \perp$. If our axioms are adequate, $S \models \perp$ implies $S \vdash \perp$. We say $S$ is **consistent** if $S \nvdash \perp$ and **inconsistent** if $S \vdash \perp$. Therefore, in an adequate system, if $S$ has no models then $S$ is inconsistent; equivalently, if $S$ is consistent then it has a model.

In fact, the statement that consistent axiom sets have a model implies adequacy in general. Indeed, if $S \models t$, then $S \cup \{\neg t\}$ has no models, and so it is inconsistent by assumption. Then $S \cup \{\neg t\} \vdash \perp$, so $S \vdash \neg t \Rightarrow \perp$ by the deduction theorem, giving $S \vdash t$ by axiom 3.

We aim to construct a model of $S$ given that $S$ is consistent. Intuitively, we want to write

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

but this does not work on the set $S = \{p_1, p_1 \Rightarrow p_2\}$ as it would evaluate $p_2$ to false.

We say a set $S \subseteq L$ is **deductively closed** if $p \in S$ whenever $S \vdash p$. Any set $S$ has a **deductive closure**, which is the (deductively closed) set of statements $\{t \in L : S \vdash t\}$ that $S$ proves. If $S$ is consistent, then the deductive closure is also consistent. Computing the deductive closure before the valuation solves the problem for $S = \{p_1, p_1 \Rightarrow p_2\}$. However, if a primitive proposition $p$ is not in $S$, but $\neg p$ is also not in $S$, this technique still does not work, as it would assign false to both $p$ and $\neg p$.

**Theorem 1.2** (Model Existence Lemma)
Every consistent set $S \subseteq L$ has a model.

*Remark* 7. We use the fact that $P$ is a countable set in order to show that $L$ is countable. The result does in fact hold if $P$ is uncountable, but requires Zorn's Lemma and will be proved in Chapter 3. Some sources call this theorem the 'completeness theorem'.

*Proof.* First, we claim that for any consistent $S \subseteq L$ and proposition $p \in L$, either $S \cup \{p\}$ is consistent or $S \cup \{\neg p\}$ is consistent. If this were not the case, then $S \cup \{p\} \vdash \bot$, and also $S \cup \{\neg p\} \vdash \bot$. By the deduction theorem, $S \vdash p \Rightarrow \bot$ and $S \vdash (\neg p) \Rightarrow \bot$. But then $S \vdash \neg p$ and $S \vdash \neg\neg p$, so $S \vdash \bot$ contradicting consistency of $S$.

Now, $L$ is a countable set as each $L_n$ is countable, so we can enumerate $L$ as $t_1, t_2, \dots$.
Let $S_0 = S$, and define $S_1 = S_0 \cup \{t_1\}$ or $S_1 = S_0 \cup \{\neg t_1\}$, chosen s.t. $S_1$ is consistent. Continuing inductively, define $\overline{S} = \bigcup_i S_i$.
Then, $\forall\, t \in L$, either $t \in \overline{S}$ or $\neg t \in \overline{S}$.
Note that $\overline{S}$ is consistent since proofs are finite; indeed, if $\overline{S} \vdash \bot$, then this proof uses hypotheses only in $S_n$ for some $n$, but then $S_n \vdash \bot$ contradicting consistency of $S_n$.
Note also that $\overline{S}$ is deductively closed, so if $\overline{S} \vdash p$, we must have $p \in \overline{S}$; otherwise, $\neg p \in \overline{S}$ so $\overline{S} \vdash \neg p$, giving $\overline{S} \vdash \bot$ by MP, contradicting consistency of $\overline{S}$.

Now, define the function

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

We show that $v$ is a valuation, then the proof is complete as $v(s) = 1$ for all $s \in S$. Since $\overline{S}$ is consistent, $\bot \notin \overline{S}$, so $v(\bot) = 0$.

Suppose $v(p) = 1, v(q) = 0$. Then $p \in \overline{S}$ and $q \notin \overline{S}$, and we want to show $(p \Rightarrow q) \notin \overline{S}$. If this were not the case, we would have $(p \Rightarrow q) \in \overline{S}$ and $p \in \overline{S}$, so $q \in \overline{S}$ as $\overline{S}$ is deductively closed.

Now suppose $v(q) = 1$, so $q \in \overline{S}$, and we need to show $(p \Rightarrow q) \in \overline{S}$. Then $\overline{S} \vdash q$, and by axiom 1, $\overline{S} \vdash q \Rightarrow (p \Rightarrow q)$. Therefore, as $\overline{S}$ is deductively closed, $(p \Rightarrow q) \in \overline{S}$.

Finally, suppose $v(p) = 0$, so $p \notin \overline{S}$, and we want to show $(p \Rightarrow q) \in \overline{S}$. We know that $\neg p \in \overline{S}$, so it suffices to show that $(p \Rightarrow \perp) \vdash (p \Rightarrow q)$. By the deduction theorem, this is equivalent to proving $\{p, p \Rightarrow \perp\} \vdash q$, or equivalently, $\perp \vdash q$. But by axiom 1, $\perp \Rightarrow (\neg q \Rightarrow \perp)$ where $(\neg q \Rightarrow \perp) = \neg\neg q$, so the proof is complete by axiom 3. $\square$

**Corollary 1.1** (Adequacy)
Let $S \subseteq L$ and let $t \in L$, s.t. $S \models t$. Then $S \vdash t$.

*Proof.* $S \cup \{\neg t\} \models \perp$, so Model Existence Lemma, $S \cup \{\neg t\} \vdash \perp$. Then by Deduction Theorem $S \vdash \neg\neg t$. $\neg\neg t \Rightarrow t$ by Axiom 3 and so by MP $S \vdash t$. $\square$

## §1.7 Completeness

**Theorem 1.3** (Completeness Theorem for Propositional Logic)
Let $S \subseteq L$ and $t \in L$. Then $S \models t$ iff $S \vdash t$.

*Proof.* Follows from soundness and adequacy. $\square$

**Theorem 1.4** (Compactness Theorem)
Let $S \subseteq L$ and $t \in L$ with $S \models t$. Then there exists a finite subset $S' \subseteq S$ s.t. $S' \models t$.

*Proof.* Trivial after applying the completeness theorem, since proofs depend on only finitely many hypotheses in $S$. $\square$

**Corollary 1.2** (Compactness Theorem, Equivalent Form)
Let $S \subseteq L$. Then if every finite subset $S' \subseteq S$ has a model, then $S$ has a model.

*Proof.* Let $t = \perp$ in the compactness theorem. Then, if $S \models \perp$, some finite $S' \subseteq S$ has $S' \models \perp$. But this is not true by assumption, so there is a model for $S$. $\square$

*Remark* 8. This corollary is equivalent to the more general compactness theorem, since the assertion that $S \models t$ is equivalent to the statement that $S \cup \{\neg t\}$ has no model, and $S' \models t$ is equivalent to the statement that $S' \cup \{\neg t\}$ has no model.

*Note.* The use of the word compactness is more than a fanciful analogy. See Sheet 1.

**Theorem 1.5** (Decidability Theorem)
Let $S \subseteq L$, $S$ finite and $t \in L$. Then, there is an algorithm to decide (in finite time) if $S \vdash t$.

*Proof.* Trivial after replacing $\vdash$ with $\models$, and checking all valuations by drawing the relevant truth tables. $\square$

# §2 Well-Orderings

## §2.1 Definition

**Definition 2.1** (Linear Order)
A **linear order** or **total order** is a pair $(X, <)$ where $X$ is a set, and $<$ is a relation on $X$ s.t.

- (irreflexivity) $\forall\, x \in X, \neg(x < x)$;
- (transitivity) $\forall\, x, y, z \in X, (x < y \wedge y < z) \Rightarrow (x < z)$;
- (trichotomy) $\forall\, x, y \in X$, either $x < y$, $y < x$, or $x = y$.

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

*Note.* In trichotomy, exactly one holds, e.g. if $x < y$ and $y < x$, then $x < x$ by transitivity contradicting irreflexivity.

If $X$ is linearly ordered by $<$, we use the obvious notation $x > y$ to denote $y < x$. In terms of the $\leq$ relation, we can equivalently write the axioms of a linear order as

- (reflexivity) $\forall\, x \in X, x \leq x$;
- (transitivity) $\forall\, x, y, z \in X, (x \leq y \wedge y \leq z) \Rightarrow (x \leq z)$;
- (antisymmetry) $\forall\, x, y \in X$, if $(x \leq y \wedge y \leq x) \Rightarrow (x = y)$.
- (trichotomy, or totality) $\forall\, x, y \in X$, either $x \leq y$ or $y \leq x$.

**Example 2.1**
1. $(\mathbb{N}, \leq)$ is a linear order.
2. $(\mathbb{Q}, \leq)$ is a linear order.
3. $(\mathbb{R}, \leq)$ is a linear order.
4. $(\mathbb{N}^+, |)$ is not a linear order, where $|$ is the divides relation, since $2$ and $3$ are not related.
5. $(\mathcal{P}(S), \subseteq)$ is not a linear order if $|S| > 1$, since it fails trichotomy.

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

**Definition 2.2** (Well-Ordering)
A linear order $(X, <)$ is a **well-ordering** if every nonempty subset $S \subseteq X$ has a least

element.

$$\forall\, S \subseteq X,\, S \neq \varnothing \Rightarrow \exists\, x \in S,\, \forall\, y \in S,\, x \leq y$$

We say $X$ is well-ordered by $<$, or simply say $X$ is a well-ordered set.

*Note.* This least element is unique by antisymmetry.

**Example 2.2**

1. $(\mathbb{N}, <)$ is a well-ordering.
2. $(\mathbb{Z}, <)$ is not a well-ordering, since $\mathbb{Z}$ has no least element.
3. $(\mathbb{Q}, <)$ is not a well-ordering.
4. $(\mathbb{R}, <)$ is not a well-ordering.
5. $[0, 1] \subset \mathbb{R}$ with the usual order is not a well-ordering, since $(0, 1]$ has no least element.
6. $\left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\} \subset \mathbb{R}$ with the usual order is a well-ordering.
7. $\left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\} \cup \{1\}$ with the usual order is also a well-ordering.
8. $\left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\} \cup \{2\}$ with the usual order is another example.
9. $\left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\} \cup \left\{1 + \frac{1}{2}, 1 + \frac{2}{3}, 1 + \frac{3}{4}, \dots\right\}$ is another example.

*Note.* Every subset of a well-ordered set is well-ordered.

*Remark* 9. Let $(X, <)$ be a linear order. $(X, <)$ is a well-ordering iff there is no infinite decreasing sequence $x_1 > x_2 > \dots$. Indeed, if $(X, <)$ is a well-ordering, then the set $\{x_1, x_2, \dots\}$ has no minimal element, contradicting the assumption. Conversely, if $S \subseteq X$ has no minimal element, then we can construct an infinite decreasing sequence by arbitrarily choosing points $x_1 > x_2 > \dots$ in $S$, which exists as $S$ has no minimal element.

**Definition 2.3** (Order-Isomorphism)

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

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

Examples (1) and (6) are isomorphic, and (7) and (8) are isomorphic. Examples (1) and (7) are not isomorphic, since example (7) has a greatest element and (1) does not. Example (9) is not isomorphic to (6) or (7).

**Example 2.3** 1. $\mathbb{N}, \mathbb{Q}$ are not order-isomorphic.

2. $\mathbb{Q}, \mathbb{Q} \setminus \{0\}$ are.

**Definition 2.4** (Initial Segment)
A subset $I$ of a totally ordered set $X$ is an **initial segment** (i.s.) if $x \in I$ implies $y \in I$ for all $y < x$.

**Example 2.4**
$\{1, 2, 3, 4\}$ is an i.s. of $\mathbb{N}$. $\{1, 2, 3, 5\}$ is not.

*Remark* 10. In any linear ordering $X$ and element $x \in X$, the set $\{y : y < x\}$ is an initial segment by transitivity.

Not every initial segment is of this form, for instance $\{x : x \leq 3\}$ in $\mathbb{R}$, or $\{x : x > 0, x^2 < 2\}$ in $\mathbb{Q}$.

*Remark* 11. In a well-ordering, every proper initial segment $I \neq X$ is of this form. Indeed, letting $I_x = \{y : y < x\}$ where $x$ is the least element of $X \setminus I$ we see $I_x = I$.
If $y \in I_x$ then $y < x$ so $y \in I$ by choice of $x$, i.e. $I_x \subseteq I$. If $y \in I$ and $y \geq x$, then $x \in I$ as $I$ is an i.s. $\lightning$ so $y < x$, i.e. $y \in I_x$ and $I \subseteq I_x$.

**Lemma 2.1**
Let $X, Y$ be well-ordered sets, $I$ an i.s. of $Y$ and $f : X \to Y$ be an order-isomorphism between $X$ and $I$.
Then $\forall\, x \in X$, $f(x)$ is the least element of $Y \setminus \{f(t) : t < x\}$.

*Proof.* The set $A = Y \setminus \{f(t) : t < x\}$ is non-empty, e.g. $f(x) \in A$. Let $a$ be the least element of $A$. Then $a \leq f(x)$ and $f(x) \in I$ and so $a \in I$. Thus $a = f(z)$ for some $z \in X$. Note that $z > x$ implies that $a = f(z) > f(x)$ $\lightning$, so $z \leq x$. If $z < x$ then $a = f(x) \in f(t) : t < x$ $\lightning$ as $a \in A$. So $z = x$ and $a = f(z) = f(x)$. $\square$

**Proposition 2.1** (Proof by Induction)
Let $X$ be a well-ordered set, and let $S \subseteq X$ be s.t. for every $x \in X$

$$(\forall\, y < x,\, y \in S) \Rightarrow x \in S$$

Then $S = X$.

*Remark* 12. Equivalently, if $p(x)$ is a property s.t. if $p(y)$ is true for all $y < x$ then $p(x)$, then $p(x)$ holds for all $x$.

Formally, if $S$ is given by a property $p$, $S = \{x \in X : p(x)\}$.
$(\forall\, x \in X)((\forall\, y < x, p(y)) \Rightarrow p(x)) \Rightarrow (\forall\, x \in X, p(x))$ (base case is included).

> *Proof.* Suppose $S \neq X$. Then $X \setminus S$ is nonempty, and therefore has a least element $x$. But all elements $y < x$ lie in $S$, and so by the property of $S$, we must have $x \in S$, contradicting the assumption. □

**Proposition 2.2**

Let $X, Y$ be order-isomorphic well-orderings. Then there is exactly one order-isomorphism between $X$ and $Y$.

Note that this does not hold for general linear orderings, such as $\mathbb{Q}$ to itself or $[0, 1]$ to itself by $x \mapsto x$ or $x \mapsto x^2$.

> *Proof.* Let $f, g\colon X \to Y$ be order-isomorphisms. We show that $f(x) = g(x)$ for all $x$ by induction on $x$. Suppose $f(y) = g(y)$ for all $y < x$. We must have that $f(x) = a$, where $a$ is the least element of $Y \setminus \{f(y) : y < x\}$. Indeed, if not, we have $f(x') = a$ for some $x' > x$ by bijectivity, contradicting the order-preserving property. Note that the set $Y \setminus \{f(x) : y < x\}$ is nonempty as it contains $f(x)$. So $f(x) = a = g(x)$, as required. □

*Remark* 13. Induction proves things. We need a tool to construct things.

## §2.2 Initial segments

*Note.* A function from a set $X$ to a set $Y$ is a subset of $f$ of $X \times Y$ s.t.

1. $\forall\, x \in X\ \exists\, y \in Y\ (x, y) \in f$;
2. $\forall\, x \in X\ \forall\, y, z \in Y\ ((x, y) \in f \wedge (x, z) \in f) \Rightarrow (y = z)$.

Of course we write $y = f(x)$ instead of $(x, y) \in f$. Note that $f \in \mathcal{P}(X \times Y)$.

For $Z \subseteq X$, the restriction of $f$ to $Z$ is $f|_Z = \{(x, y) \in f; x \in Z\}$. $f|_Z$ is a fcn $Z \to Y$, so $f|_Z \subseteq Z \times Y \subseteq X \times Y$ so $f_Z \in \mathcal{P}(Z \times Y)$.

**Theorem 2.1** (Definition by Recursion)

Let $X$ be a w.o. set and $Y$ be any set. Then for any fcn $G\colon \mathcal{P}(X \times Y) \to Y$ there's a unique fcn $f : X \to Y$ s.t. $f(x) = G(f|_{I_x})$ for every $x \in X$.

*Remark* 14. What this means in defining $f(x)$, we may use the value of $f(y)$ for all $y < x$.

*Proof.* For uniqueness, we apply induction on $x$. If $f, f'$ agree below $x$, then they must agree at $x$ since $f(x) = G\Big(f|_{I_x}\Big) = G\Big(f'|_{I_x}\Big) = f'(x)$.

We say that $h$ is an **attempt** to mean that $h\colon I \to Y$ where $I$ is some i.s. of $X$, s.t. $\forall\, x \in I,\ h(x) = G\Big(h|_{I_x}\Big)$ (note $I_x \subseteq I$).

Let $h, h'$ be attempts. We show that $\forall\, x \in X$ if $x \in \operatorname{dom}(h) \cap \operatorname{dom}(h')$ then $h(x) = h'(x)$ ($\operatorname{dom}(h)$ is the domain of h, i.e. $I$ above). Fix $x \in \operatorname{dom}(h) \cap \operatorname{dom}(h')$ and assume $h(y) = h'(y)$ for every $y < x$ (note $y < x$ implies $y \in \operatorname{dom}(h) \cap \operatorname{dom}(h')$). Then $h|_{I_x} = h'|_{I_x}$ so $h(x) = G(h|_{I_x}) = G(h'|_{I_x}) = h'(x)$. Done by induction.

Now we need to show that $\forall\, x \in X\ \exists$ attempt $h$ s.t. $x \in \operatorname{dom}(h)$. We prove this by induction. Fix $x \in X$ and assume that for $y < x$ there's an attempt defined at $y$, and let $h_y$ be the unique attempt with domain $\{z \in X : z \leq y\} = I_y \cup \{y\}$. Then $h = \bigcup_{y<x} h_y$ is a well defined fcn on $I_x$ and it is an attempt since for $y < x$, $h(y) = h_y(y) = G(h_y|_{I_y}) = G(h|_{I_y})$.

The attempt $h' = h \cup \{(x, G(h))\}$ is an attempt with domain $I_x \cup \{x\}$. Therefore, there is an attempt defined at each $x$, so we can define $f\colon X \to Y$ by $f(x) = h(x)$ where $h$ is some attempt defined at $x$. This is well defined by above and $f(x) = h(x) = G(h|_{I_x}) = G(f|_{I_x})$. □

**Proposition 2.3** (Subset Collapse)

Let $Y$ be a w.o. set where $X \subseteq Y$. Then $X$ is order-isomorphic to a unique initial segment of $Y$.

This is not true for general linear orderings, such as $\{1, 2, 3\} \subset \mathbb{Z}$, or $\mathbb{Q}$ in $\mathbb{R}$.

*Proof.* WLOG $X \neq \varnothing$.

Uniqueness: Assume $f : X \to I$ is an o.i. where $I$ is an i.s. of $Y$. By lemma 2.1, $f(x) = \min(Y \setminus \{f(y) : y < x, y \in X\})$. So by induction, $f$ and hence $I$ are uniquely determined.

Existence: If $f$ is some such isomorphism, we must have that $f(x)$ is the least element of $X$ not of the form $f(y)$ for $y < x$. We define $f$ in this way by recursion, and this is an isomorphism as required. Note that this is always well-defined as $f(y) \leq y$, so there is always some element of $X$ (namely, $x$) not of the form $f(y)$ for $y < x$. □

*Remark* 15. A w.o. set $X$ cannot be isomorphic to a proper i.s. by uniqueness as it is isomorphic to itself.

## §2.3 Relating well-orderings

**Definition 2.5**

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

$X \leq Y$ iff $X$ is isomorphic to some subset of $Y$.

**Example 2.5**

$\mathbb{N} \leq \left\{\frac{1}{2}, \frac{2}{3}, \dots\right\}$.

**Proposition 2.4**

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

*Proof.* By recursion we define the function $f\colon X \to Y$ by letting $f(x)$ be the least element of $Y$ not of the form $f(y)$ for all $y < x$. If a least element of this form always exists, this is a well-defined isomorphism from $X$ to an initial segment of $Y$ as required. Suppose that $Y \setminus \{f(y) : y < x\}$ is empty, so $\{f(y) : y < x\} = Y$. Then $Y$ is isomorphic to $I_x \subseteq X$, so $Y \leq X$. □

**Proposition 2.5**

Let $X, Y$ be well-orderings, and suppose $X \leq Y$ and $Y \leq X$. Then $X$ is isomorphic to $Y$.

*Proof.* Let $f\colon X \to Y$ and $g\colon Y \to X$ be isomorphisms to initial segments. Then $g \circ f$ is an isomorphism from $X$ to some initial segment of $X$, as an initial segment of an initial segment is an initial segment. So by uniqueness, $g \circ f$ is the identity map on $X$. Similarly, $f \circ g$ is the identity on $Y$, so $f$ and $g$ are inverses. □

## §2.4 Constructing larger well-orderings

**Definition 2.6**

For well-orderings $X, Y$, we write $X < Y$ if $X \leq Y$ and $X$ is not isomorphic to $Y$.

Equivalently, $X < Y$ if $X$ is isomorphic to a proper initial segment of $Y$.

Let $X$ be a well-ordering, and let $x \notin X$. Construct the well-ordering on $X \cup \{x\}$ by setting $y < x$ for all $y \in X$. This well-ordering is strictly greater than $X$, since $X$ is isomorphic to a proper initial segment. This is called the **successor** of $X$, written $X^+$.

For well-orderings $(X, <_X), (Y, <_Y)$, we say that $(Y, <_Y)$ **extends** $(X, <_X)$ if $X \subseteq Y$, $<_Y|_X = <_X$, and $X$ is an initial segment of $Y$. We say that well-orderings $X_i$ for $i \in I$ are **nested** if for all $i, j \in I$, either $X_i$ extends $X_j$ or $X_j$ extends $X_i$.

**Proposition 2.6**

Let $X_i$ for $i \in I$ be a nested set of well-orderings. Then, there exists a well-ordering $X$ s.t. $X_i \leq X$ for all $i \in I$.

*Proof.* Let $X = \bigcup_{i \in I} X_i$ with ordering $<_X = \bigcup_{i \in I} <_i$. Then, as the $X_i$ are nested, each $X_i$ is an initial segment of $X$. We show that this is a well-ordering. Let $S \subseteq X$ be a nonempty set. Then $S \cap X_i \neq \varnothing$ for some $i \in I$. Let $x$ be the least element of $S \cap X_i$. Thus, $x$ is the least element of $S$, as $X_i$ is an initial segment of $X$. $\square$

*Remark* 16. The proposition holds without the nestedness assumption.

## §2.5 Ordinals

**Definition 2.7**

An **ordinal** is a well-ordered set, where we regard two ordinals as equal if they are isomorphic.

*Remark* 17. We cannot construct ordinals as equivalence classes of well-orderings, due to Russell's paradox. Later, we will see a different construction that deals with this problem.

**Definition 2.8**

Let $X$ be a well-ordering corresponding to an ordinal $\alpha$. Then, we say that $X$ has **order type** $\alpha$.

The order type of the unique well-ordering on a collection of $k \in \mathbb{N}$ points is named $k$. The order type of $(\mathbb{N}, <)$ is named $\omega$.

**Example 2.6**

In the reals, the set $\{-2, 3, -\pi, 5\}$ has order type 4. The set $\left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right\}$ has order type $\omega$.

We will write $\alpha \leq \beta$ if $X \leq Y$ where $X$ has order type $\alpha$ and $Y$ has order type $\beta$. This does not depend on the choice of representative $X$ or $Y$. We define $\alpha < \beta$ and $\alpha^+$ in a similar way. Note that $\alpha \leq \beta, \beta \leq \alpha$ implies $\alpha = \beta$. Therefore, ordinals are totally ordered.

**Proposition 2.7**

Let $\alpha$ be an ordinal. Then the set of ordinals less than $\alpha$ form a well-ordered set of order type $\alpha$.

*Proof.* Let $X$ be a well-ordering with order type $\alpha$. Then, the well-orderings less than $X$ are precisely the proper initial segments of $X$, up to isomorphism. The initial segments of $X$ are precisely the sets $I_x = \{y \in X : y < x\}$ for $x \in X$. But these are order isomorphic to $X$ itself by mapping $I_x \mapsto x$. $\square$

We define $I_\alpha = \{\beta < \alpha\}$, which is a well-ordered set of order type $\alpha$. This is often a convenient representative to choose for an ordinal.

**Proposition 2.8**

Every nonempty set $S$ of ordinals has a least element.

*Proof.* Let $\alpha \in S$. Suppose $\alpha$ is not the least element of $S$. Then $S \cap I_\alpha$ is nonempty. But $I_\alpha$ is well-ordered, so $S \cap I_\alpha$ has a minimal element as required. $\square$

**Theorem 2.2** (Burali-Forti paradox)

The ordinals do not form a set.

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

*Remark* 18. Given a set $S = \{\alpha_i \colon i \in I\}$ of ordinals, there exists an upper bound $\alpha$ for $S$, so $\alpha_i \leq \alpha$ for all $i \in I$, by considering the nested family of well-orderings $I_{\alpha_i}$. Hence, by the previous proposition, there exists a least upper bound, as $I_\alpha$ is a set. We write $\alpha = \sup S$.

**Example 2.7**

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

*Remark* 19. If we represent ordinals by sets of smaller ordinals, $\sup S = \bigcup_{\alpha \in S} \alpha$.

## §2.6 Some ordinals

$$0, 1, 2, 3, \dots, \omega$$

Write $\alpha + 1$ for the successor $\alpha^+$ of $\alpha$.

$$\omega + 1, \omega + 2, \omega + 3, \dots, \omega + \omega = \omega \cdot 2$$

where $\omega + \omega = \omega \cdot 2$ is defined by $\sup \{\omega, \omega + 1, \omega + 2, \dots\}$.

$$\omega \cdot 2 + 1, \omega \cdot 2 + 2, \dots, \omega \cdot 3, \omega \cdot 4, \omega \cdot 5, \dots, \omega \cdot \omega = \omega^2$$

where we define $\omega \cdot \omega = \sup \{\omega \cdot 2, \omega \cdot 3, \dots\}$.

$$\omega^2 + 1, \omega^2 + 2, \dots, \omega^2 + \omega, \dots, \omega^2 + \omega \cdot 2, \dots, \omega^2 + \omega^2 = \omega^2 \cdot 2$$

Continue in the same way.

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

where $\omega^3 = \sup \{\omega^2 \cdot 2, \omega^2 \cdot 3, \dots\}$.

$$\omega^3 + \omega^2 \cdot 7 + \omega \cdot 4 + 13, \dots, \omega^4, \omega^5, \dots, \omega^\omega$$

where $\omega^\omega = \sup \{\omega, \omega^2, \omega^3, \dots\}$.

$$\omega^\omega \cdot 2, \omega^\omega \cdot 3, \dots, \omega^\omega \cdot \omega = \omega^{\omega+1}$$

$$\omega^{\omega+2}, \dots, \omega^{\omega \cdot 2}, \omega^{\omega \cdot 3}, \dots, \omega^{\omega^2}, \dots, \omega^{\omega^3}, \dots, \omega^{\omega^\omega}, \dots, \omega^{\omega^{\omega^\omega}}, \dots, \omega^{\omega^{\omega^{\cdots}}} = \varepsilon_0$$

where $\varepsilon_0 = \sup \{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\}$.

$$\varepsilon_0 + 1, \varepsilon_0 + \omega, \varepsilon_0 + \varepsilon_0 = \varepsilon_0 \cdot 2, \dots, \varepsilon_0^2, \varepsilon_0^3, \dots, \varepsilon_0^{\varepsilon_0}$$

where $\varepsilon_0^{\varepsilon_0} = \sup \{\varepsilon_0^\omega, \varepsilon_0^{\omega^\omega}, \dots\}$.

$$\varepsilon_0^{\varepsilon_0^{\varepsilon_0^{\cdots}}} = \varepsilon_1$$

All of these ordinals are countable, as each operation only takes a countable union of countable sets.

## §2.7 Uncountable ordinals

**Theorem 2.3**

There exists an uncountable ordinal.

*Remark* 20. The reals cannot be explicitly well-ordered.

*Proof.* Let $A \subseteq \mathcal{P}(\omega \times \omega)$ be the set of well-orderings of subsets of $\mathbb{N}$. Let $B$ be the set of order types of $A$. Then $B$ is the set of all countable ordinals. Let $\omega_1 = \sup B$. $\omega_1$ is uncountable, and in particular, the least uncountable ordinal. Indeed, if it were countable, it would be the greatest element of $B$, but $\omega_1 + 1$ would also lie in $B$. $\square$

*Remark* 21. Without introducing $A$, it would be difficult to show that $B$ was in fact a set.

*Remark* 22. Another ending to the proof above is as follows. $B$ cannot be the set of all ordinals, since the ordinals do not form a set by the Burali-Forti paradox, so there exists an uncountable ordinal. In particular, there exists a least uncountable ordinal.

The ordinal $\omega_1$ has a number of remarkable properties.

1. $\omega_1$ is uncountable, but $\{\beta : \beta < \alpha\}$ is countable for all $\alpha < \omega_1$.
2. There exists no sequence $\alpha_1, \alpha_2, \dots$ in $I_{\omega_1}$ with supremum $\omega_1$, as it is bounded by $\sup\{\alpha_1, \alpha_2, \dots\}$, which is a countable ordinal.

**Theorem 2.4** (Hartogs' lemma)

For every set $X$, there exists an ordinal $\gamma$ that does not inject into $X$.

*Proof.* Use the argument above from the existence of an uncountable ordinal. $\square$

We write $\gamma(X)$ for the least ordinal that does not inject into $X$. For example $\gamma(\omega) = \omega_1$.

## §2.8 Successors and limits

**Definition 2.9**

We say that an ordinal $\alpha$ is a **successor** if there exists $\beta$ s.t. $\alpha = \beta^+$. Otherwise, $\alpha$ is a **limit**.

Equivalently, an ordinal is a successor iff it has a greatest element. An ordinal $\alpha$ is a limit iff it has no greatest element, or equivalently, for all $\beta < \alpha$, there exists $\gamma < \alpha$ with $\gamma > \beta$, giving $\alpha = \sup\{\beta : \beta < \alpha\}$.

**Example 2.8**

5 is a successor. $\omega + 2 = (\omega^+)^+$ is a successor. $\omega$ is a limit as it has no greatest element. 0 is a limit.

## §2.9 Ordinal arithmetic

Let $\alpha, \beta$ be ordinals. We define $\alpha + \beta$ by induction on $\beta$, by

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

**Example 2.9**

$\omega + 1 = \omega + 0^+ = (\omega + 0)^+ = \omega^+$. $\omega + 2 = \omega + 1^+ = (\omega + 1)^+ = (\omega^+)^+$. $1 + \omega = \sup \{1 + \gamma : \gamma < \omega\} = \omega$. Therefore, addition is noncommutative.

*Remark* 23. As the ordinals do not form a set, we must technically define addition $\alpha + \gamma$ by induction on the set $\{\gamma : \gamma \leq \beta\}$. The choice of $\beta$ does not change the definition of $\alpha + \gamma$ as defined for $\gamma \leq \beta$.

**Proposition 2.9**

Ordinal addition is associative.

*Proof.* Let $\alpha, \beta, \gamma$ be ordinals. We use induction on $\gamma$. Suppose $\alpha + (\beta + \delta) = (\alpha + \beta) + \delta$ for all $\delta < \gamma$.

First, suppose $\gamma = 0$. $\alpha + (\beta + 0) = \alpha + \beta = (\alpha + \beta) = 0$, as required. Now consider $\gamma^+$.

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

Finally, consider $\lambda$ a nonzero limit.

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

We claim that $\beta + \lambda$ is a limit. Indeed, $\beta + \lambda = \sup \{\beta + \gamma : \gamma < \lambda\}$, but for every $\gamma < \lambda$ there exists $\gamma' < \lambda$ with $\gamma < \gamma'$ as $\lambda$ is a limit, so $\beta + \gamma < \beta + \gamma'$. Thus, there is no greatest element in the set $\{\beta + \gamma : \gamma < \lambda\}$, so $\beta + \lambda$ is a limit.

Now, $\alpha + (\beta + \lambda) = \sup \{\alpha + \delta : \delta < \beta + \lambda\}$. So it suffices to show that

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

Certainly

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

as $\gamma < \lambda$ implies $\beta + \gamma < \beta + \lambda$. Further, for any $\delta < \beta + \lambda$, $\delta \leq \beta + \gamma$ for some $\gamma < \lambda$ by definition of $\beta + \lambda$. Therefore, $\alpha + \delta \leq \alpha + (\beta + \gamma)$, so each element of $\{\alpha + \delta : \delta < \beta + \lambda\}$ is at most some element of $\{\alpha + (\beta + \gamma) : \gamma < \lambda\}$. So the two suprema agree. □

*Remark* 24. We used the facts

1. $\beta \leq \gamma \Rightarrow \alpha + \beta \leq \alpha + \gamma$, which is trivial by induction on $\gamma$;
2. $\beta < \gamma \Rightarrow \alpha + \beta < \alpha + \gamma$, as $\beta^+ \leq \gamma$ so $\alpha + \beta^+ \leq \alpha + \gamma$ by (i).

However, $1 < 2$ but $1 + \omega \not< 2 + \omega$.

The above is the **inductive** definition of addition; there is also a **synthetic** definition of addition. We can define $\alpha + \beta$ to be the order type of $\alpha \sqcup \beta$, where every element of $\alpha$ is taken to be less than every element of $\beta$.

For instance, $\omega + 1$ is the order type of $\omega$ with a point afterwards, and $1 + \omega$ is the order type of a point followed by $\omega$, which is clearly isomorphic to $\omega$. Associativity is clear, as $(\alpha + \beta) + \gamma$ and $\alpha + (\beta + \gamma)$ are the order type of $\alpha \sqcup \beta \sqcup \gamma$.

**Proposition 2.10**

The inductive and synthetic definitions of addition coincide.

*Proof.* We write $+'$ for synthetic addition, and aim to show $\alpha + \beta = \alpha +' \beta$. We perform induction on $\beta$.

For $\beta = 0$, $\alpha + 0 = \alpha$ and $\alpha +' 0 = \alpha$. For successors, $\alpha + \beta^+ = (\alpha + \beta)^+ = (\alpha +' \beta)^+$, which is the order type of $\alpha \sqcup \beta \sqcup \{\star\}$, which is equal to $\alpha +' \beta^+$.

Let $\lambda$ be a nonzero limit. We have $\alpha + \lambda = \sup \{\alpha + \gamma : \gamma < \lambda\}$. But $\alpha + \gamma = \alpha +' \gamma$ for $\gamma < \lambda$, so $\alpha + \lambda = \sup \{\alpha +' \gamma : \gamma < \lambda\}$. As the set $\{\alpha +' \gamma : \gamma < \lambda\}$ is nested, it is equal to its union, which is $\alpha +' \lambda$. □

Synthetic definitions can be easier to work with if such definitions exist. However, there are many definitions that can only easily be represented inductively, and not synthetically.

We define multiplication inductively by

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

**Example 2.10**

$\omega 2 = \omega 1 + \omega = \omega 0 + \omega + \omega = \omega + \omega$. Similarly, $\omega 3 = \omega + \omega + \omega$. $\omega\omega = \sup\{0, \omega 1, \omega 2, \dots\} = \{0, \omega, \omega + \omega, \dots\}$. Note that $2\omega = \sup\{0, 2, 4, \dots\} = \omega$. Multiplication is noncommutative. One can show in a similar way that multiplication is associative.

We can produce a synthetic definition of multiplication, which can be shown to coincide with the inductive definition. We define $\alpha\beta$ to be the order type of the Cartesian product $\alpha \times \beta$ where we say $(\gamma, \delta) < (\gamma', \delta')$ if $\delta < \delta'$ or $\delta = \delta'$ and $\gamma < \gamma'$. For instance, $\omega 2$ is the order type of two infinite sequences, and $2\omega$ is the order type of a sequence of pairs.

Similar definitions can be created for exponentiation, towers, and so on. For instance, $\alpha^\beta$ can be defined by

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

For example, $\omega^2 = \omega^1\omega = \omega^0\omega\omega = \omega\omega$. Further, $2^\omega = \sup\{2^0, 2^1, \dots\} = \omega$, which is countable.