# Contents

1	Sets and Classes		
	1.1	Classes	2
2	Axio	omatic Set Theory	4
	2.1	Extensionality	4
	2.2	Pairing	5
	2.3	Separtion	6
	2.4	Union	7
	2.5	Power Set	8
		2.5.1 Relations	8
		2.5.2 Functions	10
	2.6	Infinity	12
	2.7	Replacement	12
3	Ordinal Numbers		
	3.1	Ordering	14
	3.2	Ordinal Numbers	17
	3.3	Induction and Recursion	20
	3.4	Ordinal Arithmetic	22

1

# **Sets and Classes**

# 1.1 Classes

Russell's Paradox states the following:

#### Russell's Paradox

Let *X* be the set of all sets which do not contain themselves, i.e.,

$$X = \{S : S \notin S\}.$$

Now consider X. If  $X \in X$ , it means that X contains itself and should not be a member of X, i.e.,  $X \in X \implies X \notin X$ . If  $X \notin X$ , it means that X does not contain itself and therefore should be a member of X, i.e.  $X \notin X \implies X \in X$ . Hence, we have a paradox and such a set X does not exist.

However, in some cases it is still useful to consider the "set" of all sets for practical reasons. Therefore, we introduce the notion of a *class* to avoid Russell's Paradox.

#### **Definition 1.1.1 ▶ Class**

Let  $\phi$  be some formula and  $\boldsymbol{u}$  be a vector, the collection

$$\mathbb{C} = \{X : \phi(X, \boldsymbol{u})\}\$$

is called a **class** of all sets satisfying  $\phi(X, \mathbf{u})$ , where  $\mathbb{C}$  is said to be **definable** from  $\mathbf{u}$ . Equivalently, we say that

$$X \in \mathbb{C} \iff \phi(X, \boldsymbol{u}).$$

In particular, if  $\mathbb{C} = \{X : \phi(X)\}$ , i.e.,  $\phi$  only has one free variable, then we say that  $\mathbb{C}$  is **definable**.

*Remark.* It is easy to see that every set *X* is a class given by  $\{x: x \in X\}$ .

Intuitively, two classes are equal if they contain exactly the same members. We are able to give the following rigorous version of the notion of equality:

### **Definition 1.1.2** ▶ Equality between Classes

Let  $\mathbb{C} = \{X : \phi(X, \mathbf{u})\}\$  and  $\mathbb{D} = \{X : \psi(X, \mathbf{v})\}\$ , we say that  $\mathbb{C} = \mathbb{D}$  if for all X,

$$\phi(X, \mathbf{u}) \iff \psi(X, \mathbf{u}).$$

There are clearly two types of classes — the ones which are also sets and the ones which are not. Formally, this is put as follows:

#### **Definition 1.1.3** ▶ **Proper Class**

A class  $\mathbb{C}$  is said to be a **proper class** if  $\mathbb{C} \neq X$  for all sets X.

Like sets, we can define subclasses:

#### **Definition 1.1.4** ► **Subclass**

Let A and B be classes. We say that A is a **subclass** of B if every member of A is also a member of B, i.e.,

$$\mathbb{A}\subseteq\mathbb{B}\iff(X\in\mathbb{A}\implies X\in\mathbb{B}).$$

We shall also define the operations applicable to classes:

#### **Definition 1.1.5** ► **Intersection, Union and Difference**

Let A and B be classes. The **intersection**, **union** and **difference** between A and B are given by

$$A \cap B := \{X : X \in A \land X \in B\},\$$

$$\mathbb{A} \cup \mathbb{B} \coloneqq \{X : X \in \mathbb{A} \lor X \in \mathbb{B}\},\$$

$$\mathbb{A} - \mathbb{B} := \{X : X \in \mathbb{A} \land X \notin \mathbb{B}\}$$

respectively.

Finally, we shall introduce the universal class:

# **Definition 1.1.6 ▶ Universal Class**

The **universal class** is the class of all sets, denoted by

$$V \coloneqq \{X : X = X\}.$$

*Remark.* It is easy to prove that the universal class is **unique**.

2

# **Axiomatic Set Theory**

In Naïve Set Theory, we define a set as "a collection of mathematical objects which satisfy certain definable properties". However, such a definition is problematic (e.g. it leads to the Russell's Paradox). Thus, instead of viewing a set as a clearly defined mathematical object, we can think a set as an object entirely defined by a set of axioms to which it complies. In this sense, we avoid paradoxes by making the notion of a set undefined but only specify rigorously the axioms a set must satisfy. The following sections discuss each of the axioms in ZF set theory.

# 2.1 Extensionality

#### **Axiom 2.1.1** ► Extensionality

Let X and Y be sets, then X = Y if for all  $u, u \in X$  if and only if  $u \in Y$ .

An immediate result from Axiom 2.1.1 is that there exists a set X such that X = X, i.e. every set equals itself. Moreover, we can also prove the following:

#### Theorem 2.1.2 ▶ The Empty Set

The set which has no elements is unique.

*Proof.* Let X be a set with no elements. Note that this means that for all  $u, u \notin X$ .

Let Y be another set. Note that the statement  $u \in X \implies u \in Y$  is vacuously true. Suppose that Y has no elements, then similarly for all u, the statement  $u \in Y \implies u \in X$  is also vacuously true.

Therefore, for all u, we have proven that  $u \in X$  if and only if  $u \in Y$ . By Axiom 2.1.1, this means that X = Y, i.e. the set with no elements is unique.

This set with no elements is known as the **empty set**, denoted by  $\emptyset$ .

# 2.2 Pairing

#### Axiom 2.2.1 ▶ Pairing

For all u and v, there exists a set X such that for all  $z, z \in X$  if and only if z = u or z = v.

*Remark.* Note that Axiom 2.2.1 essentially says that given any sets u and v, there exists a set whose elements are exactly u and v.

This allows us to formally define the notion of a *pair* as follows:

#### **Definition 2.2.2** ▶ Pair

For all a, b, the pair  $\{a, b\}$  is defined to be the set C such that for all x,  $x \in C$  if and only if x = a or x = b.

*Remark.* In particular, we can define the singleton  $\{a\}$  to be the pair  $\{a, a\}$ .

Furthermore, given any a and b, we can prove by Extensionality that the pair  $\{a,b\}$  is unique:

#### Theorem 2.2.3 ▶ Uniqueness of Pairs

For all a, b, the pair  $\{a, b\}$  is unique.

*Proof.* Let  $C := \{a, b\}$  and  $D := \{a, b\}$ . Suppose  $x \in C$ , then x = a or x = b, which means  $x \in D$ . Similarly, suppose  $y \in D$ , we can prove that  $y \in C$ . Therefore, for all x, we have  $x \in C$  if and only if  $x \in D$ . By Axiom 2.1.1, this means that C = D, i.e., the pair  $\{a, b\}$  is unique.

We can further define the notion of an *ordered pair*:

#### **Definition 2.2.4** ▶ Ordered Pair

For all a and b, the **ordered pair** (a, b) is defined to be the set  $\{a, \{a, b\}\}$ .

Again, one can use Extensionality to prove that such an ordered pair is always unique and that (a, b) = (c, d) if and only if a = c and b = d. The notions of pair and ordered pair can be extended to ordered and un-ordered n-tuples, which will have similar properties as we have proven as above. Recursively, we can write the following definition:

# **Definition 2.2.5** ▶ Ordered n-tuple

The *n*-tuple is defined as

$$(a_1, a_2, \dots, a_n) = ((a_1, a_2, \dots, a_{n-1}), a_n).$$

By Extensionality, we can similarly prove that two ordered *n*-tuples  $(a_1, a_2, \dots, a_n)$  and  $(b_1, b_2, \dots, b_n)$  if and only if  $a_i = b_i$  for  $i = 1, 2, \dots, n$ .

# 2.3 Separtion

#### **Axiom 2.3.1** ► **Axiom Schema of Separation**

If P is a property with parameter p, then for all X and p there exists a set

$$Y := \{u \in X : P(u, p)\}.$$

Informally, this means that for every set X, we can form another set Y by taking the elements from X which satisfy a given property P. The above axiom justifies our set-builder notation

$$\{x: \varphi(x, \boldsymbol{p})\},\$$

where  $\varphi$  is some formula and  $\boldsymbol{p}$  is an ordered n-tuple of parameters.

Alternatively, we can write Axiom Schema 2.3.1 in the following form:

Let  $\mathbb{C} = \{u : \varphi(u, \mathbf{p})\}\$  be a class, then for all sets X there exists a set Y such that  $\mathbb{C} \cap X = Y$ .

Consequently, the intersection and the difference between two sets is a set, which can be defined as

$$X \cap Y := \{u \in X : u \in Y\}$$
 and  $X - Y := \{u \in X : u \notin Y\}$ .

Suppose that there exists some set X such that X = X, we can use Separtion to define the empty set as

$$\emptyset := \{u : u \neq u\}.$$

We shall define other notions related to Separation Axioms:

#### **Definition 2.3.2** ► **Disjoint**

Two sets *X* and *Y* are called **disjoint** if  $X \cap Y = \emptyset$ .

# **Definition 2.3.3** ► **Unary Intersection**

Let  $\mathbb C$  be a non-empty class of sets, we define the **unary intersection** of  $\mathbb C$  to be

$$\bigcap \mathbb{C} := \{u : u \in X \text{ for all } X \in C\}.$$

Note that the unary intersection helps us define the intersection of two sets as

$$X \cap Y = \bigcap \{X, Y\}.$$

# 2.4 Union

# Axiom 2.4.1 ► Axiom of Union

For all X, there exists a set  $Y = \bigcup X$  whose elements are all the elements of all elements of X, i.e.

$$Y := \{ u \in U : U \in X \}.$$

*Remark.* We often call  $\bigcup X$  the unary union of X.

The unary union defines the union of two sets as

$$X \cup Y = \bigcup \{X, Y\}.$$

One can prove that union between sets is associative. In general, we can also see that

$${a_1, a_2, \cdots, a_n} = \bigcup_{i=1}^n {a_i}.$$

In addition, we can also define the notion of symmetric difference:

# **Definition 2.4.2** ▶ Symmetric Difference

The **symmetric difference** between two sets *X* and *Y* is defined as

$$X \triangle Y := \{u : u \in X \cup Y, u \notin X \cap Y\} = (X - Y) \cup (Y - X).$$

# 2.5 Power Set

#### Axiom 2.5.1 ▶ Axiom of Power Set

For all X, there exists a set  $Y = \mathcal{P}(X)$ , known as the **power set** of X, such that

$$Y := \{U: \ U \subseteq X\}.$$

This allows us to define the notion of the *Cartesian product* (or simply the *product*) of two sets:

#### **Definition 2.5.2** ► Cartesian Product

Let *X* and *Y* be sets. The Cartesian product of *X* and *Y* is defined as the set

$$X \times Y := \{(x, y) : x \in X, y \in Y\}.$$

*Remark.* Note that  $X \times Y$  is a set because  $X \times Y \subseteq \mathcal{P}(X \cup Y)$ .

The above offers a new way to define n-tuples, as we can define Cartesian products of countably many sets recursively.

#### **Definition 2.5.3** ► Cartesian Product of Countably Many Sets

et  $n \in \mathbb{N}^+$  and let X be a set, we define

$$X^n := \prod_{i=1}^n X = \left(\prod_{i=1}^{n-1} X\right) \times X.$$

#### 2.5.1 Relations

Colloquially, we may want to express the idea that a collection of n objects are related by some rules. Observe that such a *relation* between n objects can be precisely abstracted as an ordered n-tuple, which motivates the following definition:

#### **Definition 2.5.4** ▶ **Relation**

An *n*-ary relation R is a set of *n*-tuples. We say that R is an *n*-ary relation on X if  $R \subseteq X^n$ . Conventionally, to say that  $x_1, x_2, \dots, x_n$  are related by the rules defined by R, we use the notation  $R(x_1, x_2, \dots, x_n)$ . Note that this notation is equivalent to

$$(x_1, x_2, \cdots, x_n) \in R$$
.

*Remark.* In the case where R is a binary relation, we can also use the notation xRy to express that  $(x, y) \in R$ .

If *R* is a binary relation, then we define the *domain* of *R* to be

$$dom(R) = \{u : \exists v \text{ s.t. } (u, v) \in R\},\$$

and the *range* of *R* to be

$$ran(R) = \{v : \exists u \text{ s.t. } (u, v) \in R\}.$$

Note that

$$dom(R) \subseteq \bigcup (\bigcup R)$$
 and  $ran(R) \subseteq \bigcup (\bigcup R)$ ,

so the domain and range of a relation are sets. Additionally, we define the *field* of *R* to be the set

$$field(R) = dom(R) \cup ran(R)$$
.

There are some special relations which possess certain useful properties. For now, we shall introduce one such special relation known as the *equivalence relation*.

#### **Definition 2.5.5** ▶ Equivalence Relation

An equivalence relation on a set *X* is a binary relation  $\equiv$  such that for all  $x, y, z \in X$ ,

- 1.  $x \equiv x$  (reflexive);
- 2.  $x \equiv y \implies y \equiv x$  (symmetric);
- 3.  $x \equiv y$  and  $y \equiv z \implies x \equiv z$  (transitive).

The equivalence class of  $x \in X$  under  $\equiv$  is defined as

$$[x] := \{ y \in X : y \equiv x \}.$$

Furthermore, the **quotient** of *X* by  $\equiv$  is defined as

$$X/\equiv:=\{[x]:x\in X\}.$$

With a bit of observation, one can check that the equivalence classes of any set X are pairwise disjoint. Therefore, the quotient of a set is a *partition* of the set.

#### **Definition 2.5.6** ▶ Partition

Let X be a set. A partition of X is the family P of disjoint non-empty sets such that

$$X = \bigcup \{Y : Y \in P\}.$$

We see that every equivalence relation determines a partition for a set X. Conversely, a partition of X can also determine an equivalence relation where  $x \equiv y$  if and only if there is some  $Y \in P$  such that  $x, y \in Y$ .

#### 2.5.2 Functions

Given a binary relation R, we can see R as a **mapping** which corresponds each  $u \in \text{dom}(R)$  with some  $v \in \text{ran}(R)$ . From this, we are able to derive the following definition for a *function*:

#### **Definition 2.5.7** ▶ Function

Let X be a set. A binary relation f on X is a **function** if  $(x, y) \in f$  and  $(x, z) \in f$  implies that y = z, i.e. for all  $x \in X$  there exists a unique y such that  $(x, y) \in f$ . This unique y is called the **value** of f at x. We may use the notations

$$y = f(x)$$
 or  $f: x \mapsto y$ 

to express that  $(x, y) \in f$ .

*Remark.* If  $dom(f) = X^n$ , we also say that f is an n-nary function on X.

We denote a function f from X to Y by

$$f: X \to Y$$

where dom(f) = X and  $ran(f) \subseteq Y$ . The set of all functions from X to Y is denoted as  $Y^X$ , which is a set because

$$Y^X \subset \mathcal{P}(X \times Y)$$
.

If ran(f) = Y, we say that f is *onto* Y or that f is *surjective*. A function f is *one-to-one* or *injective* if

$$f(x) = f(y) \implies x = y.$$

Additionally, we may call the function  $f: X^n \to X$  an *n-nary operation* on X.

We may also define new functions from some existing function(s).

#### **Definition 2.5.8** ▶ **Restriction**

Let f be a function. The **restriction** of f to a set X is defined to be the function

$$f|_X := \{(x, y) \in f : x \in X\}.$$

#### **Definition 2.5.9** ► Extension

Let f, g be functions. g is called an **extension** of f if  $f \subseteq g$ , i.e.,

$$dom(f) \subseteq dom(g)$$
 and  $g(x) = f(x)$  for all  $x \in dom(f)$ .

# **Definition 2.5.10** ► Composition

Let f and g be functions such that  $ran(g) \subseteq dom(f)$ . The **composition** of f and g is the function denoted by  $f \circ g$  with  $dom(f \circ g) = dom(g)$  such that

$$(f \circ g)(x) = f(g(x))$$
 for all  $x \in \text{dom}(g)$ .

Note that a function provides a mapping from one set to another set, and so we can define the notion of an *image*.

# **Definition 2.5.11** ► Image and Inverse Image

Let f be a function and X be a set. The **image** of X by f is the set

$${y: \exists x \in X \text{ s.t. } y = f(x)},$$

denoted by f[X]. The **inverse image** of X by f is the set

$$\{x: f(x) \in X\},\$$

denoted by  $f^{-1}[X]$ .

*Remark.* Trivially, if  $X \cap \text{dom}(f) = \emptyset$ , then  $f[X] = \emptyset$ .

For injections, we can also define their *inverses*.

#### **Definition 2.5.12** ► **Inverse**

Let f be an injective function, then we denote the **inverse** of f by  $f^{-1}$ , which is defined by

$$f^{-1}(x) = y$$
 if and only if  $x = f(y)$ .

The above definitions for functions can be applied similarly with respect to classes.

# 2.6 Infinity

#### **Axiom 2.6.1** ► **Axiom of Infinity**

There exists an infinite set.

In the usual sense, we would define finiteness with the aid of natural numbers. However, notice that by now we have not formally defined the natural numbers yet. Therefore we take a different approach here.

In the most intuitive sense, a set is infinite if given any sub-collection of the set we can produce an element of the set outside of the sub-collection.

#### **Definition 2.6.2** ► **Inductive Set**

A set *S* is inductive if  $\emptyset \in S$  and for all  $x \in S$ ,  $x \cup \{x\} \in S$ .

In later chapters, we will discuss inductiveness and infinity in more detail.

# 2.7 Replacement

#### Axiom 2.7.1 ▶ Axiom Schema of Replacement

If a class F is a function, then for all X there exists a set  $Y = F(X) = \{F(x) : x \in X\}$ .

Similar to Axiom 2.3.1, the Replacement Schema states that given any set, we can produce another set by mapping each element to an image via a function. In particular, this is equivelent to saying that for any function F, if dom(F) is a set, then so is ran(F).

#### Axiom 2.7.2 ▶ Axiom of Regularity

For every non-empty set X, there exists some  $Y \in X$  such that  $Y \cap X = \emptyset$ .

*Remark.* Axiom 2.7.2 is sometimes known as the **Axiom of Foundation**. A direct result from it is that for all sets X, there exists some  $x \in X$  such that  $x \nsubseteq X$ .

Furthermore, we can use Axiom 2.7.2 to prove the following seemingly trivial result:

#### Theorem 2.7.3

There is no set A such that  $A \in A$ .

*Proof.* If  $A = \emptyset$ , it is immediate that  $A \notin A$  by definition.

Suppose that there exists a non-empty set *A* such that  $A \in A$ . Note that  $A \in \{A\}$ , so

$$A \cap \{A\} = A$$
.

However, by Axiom 2.7.2, since A is the only member of  $\{A\}$ , we have

$$A \cap \{A\} = \emptyset$$
,

which is a contradiction. Therefore, there exists no set A such that  $A \in A$ .

Additionally, we also introduce the Axiom of Choice:

#### **Axiom 2.7.4** ► **Axiom of Choice**

For every X with  $\emptyset \notin X$ , there exists a choice function

$$f: X \to \bigcup X$$

such that for all  $S \in X$ , we have  $f(S) \in S$ .

*Remark.* Essentially, the choice function maps every set which is a member of some family of sets to one and only one element in that set.

3

# **Ordinal Numbers**

# 3.1 Ordering

By the term "ordinal", we suggest that these numbers can be placed in such a way that any number is comparable to its neighbours in terms of precedence. Therefore, it is necessary to first define the notion of ordering over a set.

#### **Definition 3.1.1** ► **Linear and Partial Ordering**

A binary relation < on a set *P* is a partial ordering if

- 1.  $p \not< p$  for all  $p \in P$  (irreflexive);
- 2. p < q and q < r implies p < r (transitive).

We say that (P, <) is a **partially ordered set**. In particular, if a partial ordering < on P satisfies trichotomy, then it is known as a **linear ordering**.

*Remark.* It is important to recognise the difference between  $p \not< q$  and  $q \le p$ . The former means that either  $q \le p$  or p and q are not comparable.

A partial ordering, therefore, does not guarantee that every pair of elements in a set are mutually comparable. Moreover, the trichotomy condition for linear orderings essentially confines that any two elements of a set can be compared by the relation.

In the case of a linear ordering <, the relation  $\leq$  will also be a partial (and linear) ordering. In this case, we say < is *strict*.

The direct consequence of an ordered set is that we can now define the extreme values in the set.

#### Definition 3.1.2 ▶ Maxima, Suprema, Upper Bounds and Greatest-ness

Let (P, <) be a partially ordered set and  $X \subseteq P$  be non-empty, then

- 1. a is a maximal element of X if  $a \in X$  and  $a \not< x$  for all  $x \in X$ ;
- 2. a is the greatest element of X if  $a \in X$  and  $x \le a$  for all  $x \in X$ ;
- 3. a is an upper bound of X if  $x \le a$  for all  $x \in X$ .
- 4. *a* is the **supremum** of *X* if *a* is the least upper bound of *X*.

We shall discuss some subtle differences between the definitions. First, we notice that both

the maxima and the greatest element of a set must belong to the set, while upper bounds and the supremum can be outside of the set.

Second, one may check that the greatest element of a set is unique if it exists because the notation  $x \le a$  for all  $x \in X$  suggests that a is comparable to all elements in X. On the other hand, a maximum of a set may not be so. Suppose we have a set X with  $a \in X$  which is not comparable to any other element in X via <, then a still satisfies the definition of a maximum. From here, we can see that if < is linear, then the greatest element equals the maximum.

Similarly, we can define the minima, least element, lower bounds and infinum of a set.

Intuitively, let us consider two partially ordered sets P and Q. While their elements can be very different, we can conveniently regard them as the same if their elements can be ordered in the same way, i.e., we can map the elements of P to those of Q without changing their ordering. If such a mapping f exists, then we say that f is *order-preserving*.

### **Definition 3.1.3** ► **Isomorphism**

Let (P, <) and (Q, <) be partially ordered sets and  $f: P \to Q$  be a function. f is said to be **order-preserving** if f(x) < f(y) for all x < y. If f is a bijection, then it is an **isomorphism** of P and Q if both f and  $f^{-1}$  are order-preserving. An isomorphism of P onto itself is known as an **automorphism**.

Suppose we have a linearly ordered set *P*. The fact that the elements of *P* are pairwise comparable is a very nice property. However, mutual comparability alone is not sufficient for us to define the notion of ordinal numbers. Notice that for any two ordinal numbers, we would like to know not only which one is bigger, but also by how much do the two numbers differ. For this purpose, we will further introduce the notion of *well-ordering*.

#### **Definition 3.1.4** ► Well-Ordering

A linear ordering < of a set *P* is a **well-ordering** if every non-empty subset of *P* has a least element.

What Definition 3.1.4 suggests is that we can now compare the "lengths" of well-ordered sets by considering their respective least elements. More formally, we have the following:

#### Theorem 3.1.5 ► Trichotomy of Well-Ordered Sets

If  $W_1$  and  $W_2$  are well-ordered sets, then exactly one of the following holds:

- 1.  $W_1$  is isomorphic to  $W_2$ ;
- 2.  $W_1$  is isomorphic to an initial segment of  $W_2$ ;

3.  $W_2$  is isomorphic to an initial segment of  $W_1$ .

*Proof.* We first consider the following lemma:

#### Lemma 3.1.6

If (W, <) be a well-ordered set and  $f: W \to W$  be an increasing function, then  $f(x) \ge x$  for all  $x \in W$ .

*Proof.* Suppose on contrary there is some  $x \in W$  such that f(x) < x, then the set

$$X \coloneqq \{x \in W : f(x) < x\}$$

is non-empty. Since  $X \subseteq W$  is well-ordered, it has a least element z. Let f(z) = w. Since  $z \in X$ , we have w = f(z) < z, so z is not the least element of X, which is a contradiction.

Lemma 3.1.6 leads to the following lemma:

#### Lemma 3.1.7

Let W be a well-ordered set, then W is not isomorphic to an initial segment of itself.

*Proof.* Let  $V := \{x \in W : x < u\}$  be some initial segment of W. Suppose on contrary that  $f : W \to V$  is an isomorphism, then f(u) < u. However, f is increasing, so this is a contradiction by Lemma 3.1.6.

Let  $W_i(u) := \{w : w < u\}$  for i = 1, 2. Define

$$f := \{(x, y) \in W_1 \times W_2 : W_1(x) \text{ is isomorphic to } W_2(y)\}.$$

Let  $(x, y_1), (x, y_2) \in f$ , then  $W_2(y_1)$  is isomorphic to  $W_2(y_2)$ . Similarly, if  $(x_1, y), (x_2, y) \in f$ , then  $W_1(x_1)$  is isomorphic to  $W_1(x_2)$ . By Lemma 3.1.7, it is easy to see that  $y_1 = y_2$  and  $x_1 = x_2$ , so f is well-defined and injective. The surjectivity of f is immediate from its definition, so f is a bijection.

Let h be an isomorphism of  $W_1(x)$  and  $W_2(y)$  and let x' < x. Note that  $x' \in W_1(x)$ , so  $h(x') \in W_2(y)$ . This means that  $h|_{W_1(x')}$  is an isomorphism of  $W_1(x')$  and  $W_2(h(x'))$ . Therefore, f(x') = h(x') < y = f(x), which means f is order-preserving, so dom(f) is isomorphic to ran(f).

If  $dom(f) = W_1$  and  $ran(f) = W_2$ , then  $W_1$  is isomorphic to  $W_2$ .

If  $ran(f) \neq W_2$ , let  $y_0$  be the least element of  $W_2 - ran(f)$ , then for all  $y < y_0$ ,  $y \in ran(f)$ . Take any  $y > y_0$ , if  $y \in ran(f)$ , then there is some  $x \in W_1$  such that  $W_1(x)$  is isomorphic to  $W_2(y)$ , but  $y_0 \in W_2(y)$ , so there is some  $x_0 \in W_1(x)$  such that  $W_1(x_0)$  is isomorphic to  $W_2(y_0)$ , which is impossible. Therefore,  $ran(f) = W_2(y_0)$ .

Consider  $W_1 - \text{dom}(f)$ . If it is non-empty, let  $x_0$  be its least element, and so  $W_1(x_0)$  is isomorphic to  $W_2(y_0)$ . Which is impossible since  $(x_0, y_0) \notin f$ . Therefore,  $\text{dom}(f) = W_1$ . This means  $W_1$  is isomorphic to  $W_2(y_0)$ . Similarly, one may check that if  $\text{dom}(f) \neq W_1$ , then  $W_2$  is isomorphic to  $W_1(x_0)$ . By Lemma 3.1.7, we can prove that the three cases are mutually exclusive.

Intuitively, for isomorphic sets  $W_1$  and  $W_2$ , we would consider them to follow the same ordering, or the same *order-type* as a formal notion. We informally view the ordinal numbers as a representation of the order-type.

# 3.2 Ordinal Numbers

The idea behind ordinal numbers is that every ordinal number is uniquely determined by all ordinals which are strictly less than itself. For instance, the natural number 10 is uniquely determined as "the natural number which is bigger than exactly natural numbers  $0, 1, 2, \dots, 9$ ".

In this sense, an ordinal  $\alpha$  is nothing more but a set of ordinals which are strictly less than  $\alpha$ . Meanwhile, we can see that the statement  $\beta < \alpha$  between two ordinals is equivalent to  $\beta \in \alpha$ . To better define this concept in rigorous languages, we shall introduce a preliminary notion of *transitivity*.

#### **Definition 3.2.1** ► Transitive Set

A set *T* is **transitive** if  $t \subseteq T$  for all  $t \in T$ .

While being a very abstract notion, it is worth noting that if T is transitive, then every element of T is its subset, i.e.,  $T \subseteq \mathcal{P}(T)$ .

# **Definition 3.2.2** ▶ Ordinal Number

A set is an **ordinal number** if it is transitive and well-ordered by  $\in$ . The class of all ordinals is denoted by Ord.

It might be useful to consider what it means by being "well-ordered by  $\in$ ". Essentially, this means for any two different ordinals  $\alpha$  and  $\beta$ , either  $\alpha \in \beta$  or  $\beta \in \alpha$ . Being a well-ordering,

it implies that there exists an ordinal which is in every other ordinals, i.e., it is smaller than all other ordinals (which is the zero).

Vacuously,  $\emptyset$  is a transitive well-ordered set, so it is an ordinal. We denote it by 0. Since there is no set which is an element of  $\emptyset$ , 0 being the least ordinal is justified.

Now, let  $\alpha$  be an ordinal and  $\beta \in \alpha$ , then  $\beta \subseteq \alpha$ . Take any  $\gamma \in \beta$ , then  $\gamma \in \alpha$  and so  $\gamma \subseteq \alpha$ . Note that  $\in$  is transitive on  $\alpha$ , so for all  $\delta \in \gamma$  we have  $\delta \in \beta$ . Therefore,  $\gamma \subseteq \beta$ , which means  $\beta$  is transitive. Moreover, let  $\gamma \subseteq \beta$ , then for all  $\delta \in \gamma$ , we have  $\delta \in \beta \subseteq \alpha$ , so  $\gamma \subseteq \alpha$  and so it has a  $\in$ -least element, which means  $\beta$  is well-ordered by  $\in$ . Therefore, any element of an ordinal is an ordinal.

The above argument proves that every element of an ordinal is an ordinal. This means we can define  $\alpha < \beta$  if  $\alpha \in \beta$ . Next, let there be ordinals  $\alpha \neq \beta$  with  $\alpha \subset \beta$ , then  $\beta - \alpha$  is non-empty and so it has a least element  $\gamma$ . By Theorem 3.1.5, it can be proven that  $\alpha = \{\xi \in \beta : \xi < \gamma\} = \gamma \in \beta$ .

Lastly, we prove that trichotomy holds for ordinals. Note that it suffices to prove that for any ordinals  $\alpha \neq \beta$ , either  $\alpha \subset \beta$  or  $\beta \subset \alpha$ . First, observe that  $\alpha \cap \beta$  is clearly an ordinal, so we can denote  $\gamma = \alpha \cap \beta$ . Suppose  $\gamma \neq \alpha$ , but since  $\gamma \subset \alpha$  we have  $\gamma \in \alpha$ . Similarly, if  $\gamma \neq \beta$  we would have  $\gamma \in \beta$ . This implies that  $\gamma \in \gamma$  which is not possible, so either  $\gamma = \alpha$  or  $\gamma = \beta$ . Therefore, we have  $\alpha \subset \beta$  or  $\beta \subset \alpha$  as desired. With trichotomy established, we see that  $\gamma \in \beta$  in a linear ordering for the class  $\gamma \in \beta$ .

With the above reasonings, one gets the following important results:

- 1. If  $\mathbb{C}$  is a non-empty class of ordinals, then  $\bigcap \mathbb{C}$  is an ordinal. In particular,  $\bigcap \mathbb{C} \in \mathbb{C}$  and  $\bigcap \mathbb{C} \subseteq \gamma$  for all  $\gamma \in \mathbb{C}$ . We denote  $\bigcap \mathbb{C} = \inf \mathbb{C}$ .
- 2. If X is a non-empty set of ordinals, then  $\bigcup X$  is an ordinal. We can denote  $\bigcup X = \sup X$ .
- 3. If  $\alpha$  is an ordinal,  $\alpha \cup \{\alpha\}$  is an ordinal. In particular,  $\alpha \cup \{\alpha\} = \inf\{\beta : \beta > \alpha\}$ .

The last result shows that  $\alpha \cup \{\alpha\}$  is the **smallest ordinal which is greater than**  $\alpha$ . We call it the *successor* of  $\alpha$  and denote by  $\alpha + 1$ .

#### **Definition 3.2.3** ► Successor Ordinal

An ordinal  $\alpha$  is a successor ordinal if there is some ordinal  $\beta$  such that  $\alpha = \beta + 1$ .

Additionally, we can also argue that Ord must be a proper class. Otherwise, sup Ord is an ordinal and so sup  $Ord + 1 \in Ord$ , which is a contradiction.

Recall that we previously mentioned that ordinals are a order-type for well-ordered sets. More formally, this can be formulated as below:

### Theorem 3.2.4 ▶ Isomorphism of Well-Ordered Sets and Ordinals

Every well-ordered set is isomorphic to a unique ordinal number.

*Proof.* Let W be a well-ordered set. For each  $x \in W$ , we can define

$$W(x) \coloneqq \{ w \in W : w < x \}$$

to be an initial segment of W given by x. Define  $F: W \to Ord$  by  $F(x) = \alpha$  if  $\alpha$  is isomorphic to W(x). By Theorem 3.1.5, such an  $\alpha$  always exists. By Axiom Schema 2.7.1, F[W] is a set of ordinals, so  $\sup F[W]$  exists. Therefore, we have established an isomorphism between W and  $\sup F[W]$  which is an ordinal. Furthermore, suppose there are  $\alpha, \beta \in Ord$  which are both isomorphic to W, then W must be isomorphic to W which implies W high implies W is unique for each W.

One commonly used class of ordinals is the *natural numbers*. A convenient analogy for natural numbers is to consider

$$\emptyset = 0,$$
  $1 = 0 + 1, \cdots$ 

and so on, but this definition is not very rigorous. Recall that we previously justified that every ordinal has a successor, but is the converse true? In other words, **is every ordinal** a successor ordinal? Well, obviously  $\emptyset$  is not a successor ordinal. Suppose  $\alpha = \beta + 1$  is a successor ordinal, then we have  $\bigcup \alpha = \beta$ , i.e.,  $\bigcup \alpha < \alpha$ . Notice that it is not possible that  $\alpha < \bigcup \alpha$ , so negating the definition of successor ordinal we have  $\alpha = \bigcup \alpha$  if  $\alpha$  is not a successor ordinal.

#### **Definition 3.2.5** ► Limit Ordinal

An ordinal  $\alpha$  is called a **limit ordinal** if  $\alpha = \sup \alpha$ .

We will justify the existence of limit ordinals.

#### Theorem 3.2.6 ▶ Existence of Least Non-zero Limit Ordinal

 $\mathbb{N} = \bigcap \{X : X \text{ is inductive}\}\$ is the least non-zero limit ordinal.

*Proof.* Since there exists an inductive set,  $\mathbb{N}$  is a set by Axiom Schema 2.3.1 and  $\mathbb{N}$  is inductive.  $\mathbb{N}$  is non-empty because  $\emptyset \in \mathbb{N}$ . Let S be the set of all successor ordinals. Notice that  $S \cup \{\emptyset\}$  is inductive, so for all  $\eta \in \mathbb{N}$ ,  $\eta$  is either 0 or a successor ordinal. Therefore,  $\mathbb{N} \cap Ord = \mathbb{N}$ , which means  $\mathbb{N} \subseteq Ord$  and is an ordinal.

Take any  $n \in \mathbb{N}$ . Since  $n \cup \{n\} \in \mathbb{N}$ , we have  $n \in \bigcup \mathbb{N}$ , and so  $\mathbb{N} \subseteq \bigcup \mathbb{N}$ . Conversely, we wish to prove that  $\bigcup \mathbb{N} \subseteq \mathbb{N}$ . Note that for all  $m \in \bigcup \mathbb{N}$ ,  $m \in \eta$  for some  $\eta \in \mathbb{N}$ . Therefore, it suffices to prove that  $\mathbb{N}$  is transitive.

Let  $\mu \in \mathbb{N}$  be an ordinal. If  $\mu = \emptyset$  then we are done. Otherwise, suppose on contrary that  $\mathbb{N}$  is not transitive. Let  $\nu \in \mu$  be the least ordinal such that  $\nu \notin \mathbb{N}$ . Clearly,  $\nu \neq \emptyset$ , so there is some  $\nu' \in \mu$  such that  $\nu = \nu' + 1$ . However,  $\nu' < \nu$ , so  $\nu' \in \mathbb{N}$ . By the inductive property of  $\mathbb{N}$  we have  $\nu \in \mathbb{N}$  which is a contradiction. Therefore,  $\mathbb{N}$  is transitive.

Therefore,  $\mathbb{N} = \bigcup \mathbb{N}$ . Since for all  $n \in \mathbb{N}$ , we have  $n = \emptyset$  or n is a successor ordinal, this implies that  $\mathbb{N}$  is the least limit ordinal.

Sometimes, this least non-zero limit ordinal is denoted by  $\omega$ . Clearly,  $0 \in \mathbb{N}$  and for all  $n \in \mathbb{N}$ ,  $n+1 \in \mathbb{N}$  — indeed, this least non-zero limit ordinal defines exactly the notion of *natural numbers*.

#### **Definition 3.2.7** ► Natural Numbers

Let  $\mathbb{N}$  be the least non-zero limit ordinal. An ordinal  $\alpha$  is a **finite ordinal** or **natural** number if  $\alpha < \mathbb{N}$ .

# 3.3 Induction and Recursion

Induction is a common tool used in mathematical proofs. We can generalise the notion of induction using the language of ordinals.

#### Theorem 3.3.1 ▶ Transfinite Induction

Let  $\mathbb{C}$  be a class of ordinals such that

- 1.  $0 \in \mathbb{C}$ ;
- 2.  $\alpha \in \mathbb{C}$  implies that  $\alpha + 1 \in \mathbb{C}$ ;
- 3. If  $\alpha$  is a limit ordinal with  $\beta \in \mathbb{C}$  for all  $\beta < \alpha$ , then  $\alpha \in \mathbb{C}$ ,

*then*  $\mathbb{C}$  = *Ord*.

*Proof.* Suppose on contrary that  $\mathbb{C} \neq Ord$ , then there exists a least ordinal  $\alpha \in Ord - \mathbb{C}$ . Note that  $\alpha \neq 0$ . If  $\alpha$  is a successor ordinal, then there is some  $\beta \in Ord$  such that  $\beta + 1 = \alpha$ , but  $\beta < \alpha$  so  $\beta \in \mathbb{C}$ , which implies that  $\alpha \in \mathbb{C}$ . Therefore,  $\alpha$  must be a non-zero limit ordinal. However, this means for all  $\beta < \alpha$ , we have  $\beta \in \mathbb{C}$ , which again implies that  $\alpha \in \mathbb{C}$ , a contradiction.

Another term closely associated to induction is *recursion*. To define a recurrence behaviour, we first consider the notion of a *sequence*.

### **Definition 3.3.2** ► Sequence

An (infinite) sequence in X is a function  $f : \mathbb{N} \to X$ . A finite sequence of length n is a function whose domain is  $\{i : i < n\}$  for some  $n \in \mathbb{N}$ . A transfinite sequence is a function whose domain is an ordinal.

Let  $\alpha \in Ord$ , we can denote an  $\alpha$ -sequence or a sequence of length  $\alpha$  by

$$\langle a_{\xi}: \xi < \alpha \rangle$$
.

Such a sequence is also called an *enumeration* of its range. If s is a sequence of length  $\alpha$ , then we can extend it by appending some x as its  $(\alpha + 1)$ -th term. This extended sequence can be denoted by  $s \cap x$ .

Now we proceed to defining recursion.

#### **Theorem 3.3.3** ➤ Transfinite Recursion

Let X be a set. For every function G on the set of all transfinite sequences in X of length less than  $\theta \in Ord$  such that  $ran(G) \subseteq X$ , then there exists a unique  $\theta$ -sequence  $\langle a_{\alpha} : \alpha < \theta \rangle$  in X such that

$$a_{\alpha} = G(\langle a_{\xi} : \xi < \alpha \rangle)$$

for all  $\alpha < \theta$ .

*Proof.* Define a mapping  $F: Ord \to X$  by  $F(\alpha) = x$  if and only if there is a sequence  $\langle a_{\xi} : \xi < \alpha \rangle$  such that

$$a_{\xi} = G(\langle a_{\eta} : \eta < \xi \rangle)$$
 for all  $\xi < \alpha$ , and  $x = G(\langle a_{\xi} : \xi < \alpha \rangle)$ .

Note that such a sequence  $\langle a_{\xi} : \xi < \alpha \rangle$  exists by definition of G. By induction on  $\xi$ , one can check that this sequence is unique. Therefore, F is a well-defined function since G is well-defined. By taking  $a_{\alpha} = F(\alpha)$  we therefore obtain an  $(\alpha + 1)$ -sequence.

Let F':  $Ord \to X$  be any function with the above property, then by induction on  $\alpha$  we can prove that  $F'(\alpha) = F(\alpha)$  for all  $\alpha \in Ord$ , i.e., F is unique.

Lastly, we define the limiting behaviour of a sequence.

### **Definition 3.3.4** ► Limit and Continuity of Sequences

Let  $\alpha > 0$  be a limit ordinal and let  $\langle \gamma_{\xi} : \xi < \alpha \rangle$  be a non-decreasing sequence of ordinals. The **limit** of the sequence is defined as

$$\lim_{\xi \to \alpha} = \sup \{ \gamma_{\xi} : \xi < \alpha \}.$$

A sequence in *Ord* is **continuous** if for every limit ordinal  $\alpha$ ,  $\gamma_{\alpha} = \lim_{\xi \to \alpha} \gamma_{x}i$ . A sequence in *Ord* is **normal** if it is increasing and continuous.

# 3.4 Ordinal Arithmetic

Arithmetic operations, namely addition, multiplication and exponentiation, can be defined rigorously using recursion.

#### **Definition 3.4.1** ▶ Ordinal Addtion

For all  $\alpha, \beta \in Ord$ ,

- 1.  $\alpha + 0 = \alpha$ ;
- 2.  $\alpha + (\beta + 1) = (\alpha + \beta) + 1$ ;
- 3.  $\alpha + \beta = \lim_{\xi \to \beta} (\alpha + \xi)$  if  $\beta$  is a limit ordinal.

#### **Definition 3.4.2** ▶ **Ordinal Multiplication**

For all  $\alpha, \beta \in Ord$ ,

- 1.  $\alpha \cdot 0 = 0$ ;
- 2.  $\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$ ;
- 3.  $\alpha \cdot \beta = \lim_{\xi \to \beta} (\alpha \cdot \xi)$  if  $\beta$  is a limit ordinal.

# **Definition 3.4.3** ► Ordinal Exponentiation

For all  $\alpha, \beta \in Ord$ ,

- 1.  $\alpha^0 = 1$ ;
- 2.  $\alpha^{\beta+1} = \alpha^{\beta} \cdot \alpha$ ;
- 3.  $\alpha^{\beta} = \lim_{\xi \to \beta} \alpha^{\xi}$  if  $\beta$  is a limit ordinal.

Note that in the above definitions, we make use of limit of sequences only for the case where  $\beta$  is a limit ordinal, as for successor ordinals we can apply the definition of successors iteratively.

We can prove that both addition and multiplication are associative on Ord by induction.

However, they are **not commutative** on *Ord*. First, notice that

$$1 + \omega = \sup\{1 + \xi : \xi < \omega\} = \sup \mathbb{N} = \omega$$

but  $\omega + 1 = \omega \cup \{\omega\} \neq \omega$ . Therefore, we can show that for any finite ordinal n,

$$n + \omega = \omega \neq \omega + n$$
.

Similarly, we have  $2 \cdot \omega = \omega \neq \omega \cdot 2 = \omega + \omega$ , so  $n \cdot \omega = \omega \neq \omega \cdot n$  for any finite ordinal n.

Note that ordinals are essentially linearly ordered sets. Therefore, we can relate ordinal arithmetic to the sums and products of arbitrary linearly ordered sets.

#### **Definition 3.4.4** ▶ Sum of Linear Orders

Let  $(A, <_A)$  and  $(B, <_B)$  be disjoint linearly ordered sets. The sum of A and B is defined as  $A \cup B$  such that for any  $x, y \in A \cup B$ , x < y if and only if

- $x, y \in A$  and  $x <_A y$ , or
- $x, y \in B$  and  $x <_B y$ , or
- $x \in A$  and  $y \in B$ .

#### **Definition 3.4.5** ▶ **Product of Linear Orders**

Let  $(A, <_A)$  and  $(B, <_B)$  be linearly ordered sets. The **product** of A and B is the set  $A \times B$  such that for any  $(a_1, b_1), (a_2, b_2) \in A \times B, (a_1, b_1) < (a_2, b_2)$  if and only if

- $b_1 < b_2$ , or
- $b_1 = b_2$  and  $a_1 < a_2$ .

Given any ordinals  $\alpha$  and  $\beta$ , by induction on  $\beta$  we can prove that  $\alpha + \beta$  is isomorphic to  $\alpha \cup \alpha \triangle \beta$  and  $\alpha \cdot \beta$  is isomorphic to  $\alpha \times \beta$ .

There are some properties for ordinal arithmetic which comply with arithmetic over natural numbers. For example:

- 1. If  $\beta < \gamma$ , then  $\alpha + \beta < \alpha + \gamma$ .
- 2. If  $\alpha < \beta$ , then there exists a unique  $\delta$  such that  $\alpha + \delta = \beta$ .
- 3. If  $\beta < \gamma$  and  $\alpha > 0$ , then  $\alpha \cdot \beta < \alpha \cdot \gamma$ .
- 4. If  $\alpha > 0$ , then for any  $\gamma$ , there exists a unique  $\beta$  and a unique  $\rho < \alpha$  such that  $\gamma = \alpha \cdot \beta + \rho$ .
- 5. If  $\beta < \gamma$  and  $\alpha > 1$ , then  $\alpha^{\beta} < \alpha^{\gamma}$ .

This leads to what is known as *Cantor's Normal Form Theorm*.

# **Theorem 3.4.6** ▶ Cantor's Normal Form Theorem

Every ordinal  $\alpha > 0$  can be uniquely represented in the form

$$\alpha = \omega^{\beta_1} \cdot k_1 + \dots + \omega^{\beta_n} \cdot k_n,$$

where  $n \geq 1$ ,  $\alpha \geq \beta_1 > \dots > \beta_n$ , and  $k_1, \dots, k_n$  are non-zero natural numbers.