

# Set Theory

Karan Elangovan

March 14, 2021

## Contents

<b>1</b>	<b>ZFC Axioms</b>	<b>1</b>
1.1	ZF Axioms . . . . .	1
1.2	Primitive Constructs . . . . .	5
1.3	The Axiom of Choice . . . . .	6
<b>2</b>	<b>Ordinals and Cardinals</b>	<b>8</b>
2.1	Cardinality . . . . .	8

## 1 ZFC Axioms

### 1.1 ZF Axioms

Throughout these notes we will implicitly assume all the rules and axioms of logic and avoid scrutinising or considering them on the basis that they are so obvious that any axiomatic treatment would be of a purely formal and philosophical interest, and would lead to no non-trivial results that could be applied to more interesting problems.

We define a variable as a string of characters possibly with subscripts, the symbol  $\emptyset$  as the empty set, and  $\cup, \mathcal{P}$  as function symbols. We now define a term inductively as either a variable, a non function symbol or a function

symbol applied to a term. In ZFC every term is a set (though there will be as a consequence of the proceeding axioms sets which cannot be expressed as terms due to their infinite nature)

We define the atomic formulas as  $x = y$  and  $x \in y$ , where  $x, y$  are terms. We then define a formula as either an atomic formula or a logical connective or quantifier applied to a formula.

Below we have the Zermelo-Frankel axioms, which alongside the Axiom of Choice comprise the entirety of ZFC. We list each one as an intuitive interpretation alongside the formal statement.

**Axiom** (Extensionality). If  $a$  and  $b$  have the same elements, then they are equal

$$\forall x(x \in a \iff x \in b) \implies a = b.$$

**Axiom** (Empty Set). The empty set,  $\emptyset$ , has no elements.

$$\forall x(x \notin \emptyset).$$

**Axiom** (Pairing). Given any sets  $x, y$ , the set  $\{x, y\}$  exists.

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

**Axiom** (Union). The union of a family of sets behaves as we would expect.

$$\forall t(t \in \cup X \iff \exists y(t \in y \wedge y \in X)).$$

**Axiom** (Power Set). The power set of a set behaves as we would expect.

$$\forall t(t \in \mathcal{P}(X) \iff \forall x(x \in t \implies x \in X)).$$

**Axiom** (Separation). Given a set  $X$  and a formula  $\psi(x)$  which does not contain

$S$  we have the set  $x \in X | \psi(x)$ .

$$\exists S \forall x (x \in S \iff (x \in X \wedge \psi(x))).$$

**Axiom (Infinity).** There is a set  $Z$  which contains  $\emptyset$  and if  $x \in Z$  then  $x \in Z$

$$\exists Z (\emptyset \in Z \wedge \forall x (x \in Z \implies \exists y (t \in y \iff t = x))).$$

**Axiom (Replacement).** We may restrict the range of a function to an arbitrary set. Let  $\psi(x, y)$  be a formula.

$$\forall x \forall y \forall z ((\psi(x, y) \wedge \psi(x, z) \implies y = z) \implies \forall D \exists R \forall y (y \in R \iff \exists x (x \in D \wedge \psi(x, y)))).$$

**Axiom (Regularity).** Every non-empty  $x$  contains  $y$  disjoint from  $x$ .

$$\forall x (x \neq \emptyset \implies \exists y (y \in x \wedge \forall z \neg (z \in x \wedge z \in y))).$$

The next two theorems actually don't use any of the ZF axioms and are instead just based on pure logical deduction.

**Theorem 1.** Equal sets contain the same elements, i.e. the converse of the axiom of extensionality holds.

$$\forall x \forall y (x = y \implies \forall z (z \in x \iff z \in y)).$$

*Proof.* Consider arbitrary  $z$ . Then  $z \in x \iff z \in y$  as we may substitute  $x$  for  $y$  as they are equal.  $\square$

**Theorem 2.** The subset relation is transitive.

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

*Proof.* If  $x \subset y$  and  $y \subset z$ , then  $\forall t((t \in x \implies t \in y) \wedge (t \in y \implies t \in z))$ . Hence  $\forall t(t \in x \implies t \in z)$ , and so  $x \subset z$ .  $\square$

From here, wanting to prove some less trivial results we have to start invoking the ZF axioms.

**Theorem 3.** If  $x \subset y$  and  $y \subset x$ , then  $x = y$

*Proof.* As  $x \subset y$ ,  $\forall t(t \in x \implies t \in y)$ , and as  $y \subset x$ ,  $\forall t(t \in y \implies t \in x)$ . So  $\forall t(t \in x \iff t \in y)$ , which implies  $x = y$  by the axiom of extensionality.  $\square$

**Theorem 4.** For any  $x$  the singleton set  $\{x\}$  exists.

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

*Proof.* By substituting  $y = x$  in the pairing axiom we have  $\forall x \exists z \forall t (t \in z \iff (t \in x \wedge t \in x))$ , which show the existence of the desired singleton set.  $\square$

**Theorem 5.** For any  $X$  the set with the expected properties of  $\cap X$  exists.

*Proof.* By the union and separation axioms we have the existence of the set  $Y = \{x \in \cup X \mid \forall t(t \in X \implies x \in t)\}$ , which is exactly the set whose properties we desired.  $\square$

**Theorem 6.** No set contains itself.

*Proof.* Assume the contrary. That for some  $x$  we have  $x \in x$ . By 4, we have the set  $\{x\}$ . By the axiom of regularity  $\{x\}$  contains an element disjoint from itself. The only element of  $\{x\}$  is  $x$ , hence  $x$  and  $\{x\}$  are disjoint. But  $x \in x$  and  $x \in \{x\}$ , so they are not disjoint, a contradiction.  $\square$

## 1.2 Primitive Constructs

We may now use the ZF axioms to construct various primitive objects in a rigorous manner.

**Definition 1** (Ordered Pair). We define the ordered pair  $(x, y)$  as  $\{x, \{x, y\}\}$  (this definition makes sense by the axiom of pairing)

**Definition 2** (Cartesian Product). We define the cartesian product of  $A$  and  $B$  as  $\{(a, b) | a \in A \wedge b \in B\}$ , this is possible by applying the axiom of separation on a suitable set that contains all the desired pairs, such as  $\mathcal{P}(A \cup \mathcal{P}(A \cup B))$

**Definition 3** (Partial Order). A partial order,  $B$ , of  $X$  is a subset of  $X \times X$  satisfying:

1.  $\forall x(x \in X \implies (x, x) \in B)$  (reflexivity)
2.  $\forall x \forall y((x, y) \in B \wedge (y, x) \in B \implies x = y)$  (antisymmetry)
3.  $\forall x \forall y \forall z((x, y) \in B \wedge (y, z) \in B \implies (x, z) \in B)$  (transitivity)

We call a set with a partial order a poset.

We call  $x \in X$  a maximal element of  $X$  under the partial order iff  $\forall y((y \in X \wedge (x, y) \in B) \implies x = y)$ .

**Definition 4** (Total Order). A total order is a special case of a partial order. A partial order  $B \subset X \times X$  is a total order if it also satisfies connexity:  $\forall x \forall y((x \in X \wedge y \in X) \implies ((x, y) \in B \vee (y, x) \in B))$

A subset  $S$  of a partial order in which the ordering induced by the original set makes  $S$  a total order is a chain.

We call a subset  $T$  of a chain  $S$  an initial segment iff when  $u \leq v$  are in  $S$  and  $v \in T$ , then  $u \in T$

**Definition 5** (Well-Ordering). A total ordering  $\leq$  over  $X$  is a well-ordering iff every non-empty subset of  $X$  has a least element.

**Definition 6** (Function). We define a function  $f$  mapping  $A$  to  $B$  as any subset of  $A \times B$  satisfying:

- $\forall x(x \in A \implies \exists y((x, y) \in f))$
- $\forall x \forall y \forall z((x, y) \in f \wedge (x, z) \in f) \implies y = z$

We further say that  $f$  is injective iff  $\forall x_1 \forall x_2 \forall y((x_1, y) \in f \wedge (x_2, y) \in f) \implies x_1 = x_2$  and surjective iff  $\forall y(y \in B \implies \exists x(x \in A \wedge (x, y) \in f))$ . And that  $f$  is bijective iff it is injective and surjective.

### 1.3 The Axiom of Choice

The axiom of choice allows us to choose an element from each set of a family of non empty sets. Formally, if  $T$  is a set of non empty sets, then the AC states there exists a "choice" function  $f : T \rightarrow \cup T$  satisfying  $\forall t(t \in T \implies f(t) \in t)$ .

An extremely useful consequence of the AC is Zorn's Lemma. It is actually equivalent to the AC under the other ZF axioms, however we will only show that it follows from ZFC.

**Theorem 7** (Zorn's Lemma). If  $P$  is a poset such that every chain in  $P$  is bounded above under the partial order relation, then  $P$  must have a maximal element.

*Proof.* We will proceed by contradiction. Assume that  $P$  satisfies the given hypotheses but has no maximal element. Consider an arbitrary chain  $S$ , by hypothesis  $S$  is bounded above. If there are no upper bounds outside of  $S$ , then  $S$  would have a maximal element (the element at the "top of the chain"), hence the set of upper bounds of  $S$  not in  $S$ ,  $B(S)$  is non-empty.

Let  $\mathcal{C}$  be the set of chains of  $P$ . Then by the AC we have  $\varphi : \mathcal{C} \rightarrow B(\mathcal{C})$  such that  $\varphi(S) \in B(S)$ , i.e.  $\varphi$  chooses an upper bound of  $S$  not in  $S$  (as we have shown such an upper bound always exists)

Now consider a fixed  $p \in P$ . Let  $C$  be the set of well-ordered chains,  $S$ , satisfying

- $p$  is the least element of  $S$
- For any proper non-empty initial segment  $T$  of  $S$ , the least element of  $S - T$  is  $\varphi(T)$

Intuitively  $C$  is the set of well-ordered chains "starting" at  $p$  and such that for any initial segment  $\varphi$  chooses the "next" element as the upper bound, in the context of the chain  $S$

Say  $S, S' \in C$ . We claim that either  $S$  or  $S'$  is an initial segment of the other. Let  $R$  be the union of all sets that are simultaneously initial segments of  $S$  and  $S'$ , so  $R$  is itself a common initial segment which is not contained in any other initial segments. Say  $R$  is not equal to either of  $S$  or  $S'$ .  $p \in R$ , hence we have that  $\varphi(R)$  is the least element of both  $S - R$  and  $S' - R$ . Hence  $R \cup \{\varphi(R)\}$  is a common initial segment containing  $R$ , which cannot be. Hence  $R = S$  or  $R = S'$ . Hence one of  $S$  and  $S'$  is an initial segment of the other.

Now let  $U = \cup C$ . In light of the previous paragraph, we have that  $U$  is a well-ordered chain with least element  $p$ . If  $T$  is a proper non-empty initial segment of  $U$ , then we may choose  $u \in U - T$ , which means  $u \in S - T$  for some  $S \in C$ . Again by the previous paragraph we have that  $T$  is a proper non-empty initial segment of  $s$ , so  $\varphi(T)$  is the least element of  $S - T$  and hence also of  $U - T$ , hence  $U \in C$ . Hence  $U \cup \{\varphi(U)\} \in C$ , but  $(U \cup \{\varphi(U)\}) \not\subseteq U$ , a contradiction.

Hence  $P$  has a maximal element. □

## 2 Ordinals and Cardinals

### 2.1 Cardinality

We may rephrase our intuitive notions of the relative sizes in terms of injective and surjective functions so as to generalise to infinite sets.

We say  $X \lesssim Y$  if there is an injective function from  $X$  to  $Y$ , that  $X \sim Y$  if there is a bijection between  $X$  and  $Y$ , and  $X \prec Y$  if  $X \lesssim Y$  and  $X \not\sim Y$ . Intuitively the relation  $\lesssim$  should be thought of as saying one set is "smaller" than the other.

**Theorem 8** (Cantor's Theorem). For any  $X$ ,  $X \prec \mathcal{P}(X)$

*Proof.* The function  $f : X \rightarrow \mathcal{P}(X)$  defined by  $f(x) = \{x\}$  is an injection and hence  $X \lesssim \mathcal{P}(X)$ . Now say that  $X \sim \mathcal{P}(X)$ , then we have a bijection  $f : X \rightarrow \mathcal{P}(X)$ . By the axiom of separation, we have the existence of  $Y = \{x \in X \mid x \notin f(x)\}$ . By surjectivity of  $f$  we have the existence of some  $x \in X$  such that  $f(x) = Y$ , as  $Y \subset X$ . If  $x \in Y$  then  $x \notin Y$  and the converse holds by the construction of  $Y$ . This means that  $\neg(x \in Y \vee x \notin Y)$ , which is a contradiction, hence  $X \not\sim \mathcal{P}(X)$ . Hence we have  $X \prec \mathcal{P}(X)$ .  $\square$

**Theorem 9** (Cantor-Bernstein Theorem). If  $A \lesssim B$  and  $B \lesssim A$ , then  $A \sim B$ .

*Proof.* By definition, we have injection  $f : A \rightarrow B$  and  $g : B \rightarrow A$ . Consider the sequence of sets

$$A_0 = A$$

$$A_1 = g(B)$$

$$A_2 = g \circ f(A)$$

$$A_3 = g \circ f \circ g(B)$$



A trivial induction argument shows that  $A_{n+1} \subset A_n$ . So we may define the sets  $A'_n = A_n - A_{n+1}$  and  $A'_\omega = \bigcap \{A_n\}$ . It is trivially shown that these sets form a partition of  $A$ . We now define  $h : A \rightarrow B$  by

$$h(x) = \begin{cases} f(x) & \text{if } x \in \cup \{A'_\omega, A'_0, A'_2, \dots\} \\ g^{-1}(x) & \text{if } x \in \cup \{A'_1, A'_3, A'_5, \dots\} \end{cases}.$$

As the  $A'_n$  forms a partition of  $A$ ,  $h$  is well-defined on  $A$ . By considering a similar sequence starting with  $B$ , it is trivially shown that  $h$  is bijective. Hence  $A \sim B$  □