



Bridgewater State University

## Virtual Commons - Bridgewater State University

---

Honors Program Theses and Projects

Undergraduate Honors Program

---

8-18-2022

# The Foundations of Mathematics: Axiomatic Systems and Incredible Infinities

Catherine Ferris  
*Bridgewater State University*

Follow this and additional works at: [https://vc.bridgew.edu/honors\\_proj](https://vc.bridgew.edu/honors_proj)



Part of the [Special Functions Commons](#)

---

### Recommended Citation

Ferris, Catherine. (2022). The Foundations of Mathematics: Axiomatic Systems and Incredible Infinities. In *BSU Honors Program Theses and Projects*. Item 505. Available at: [https://vc.bridgew.edu/honors\\_proj/505](https://vc.bridgew.edu/honors_proj/505)

Copyright © 2022 Catherine Ferris

This item is available as part of Virtual Commons, the open-access institutional repository of Bridgewater State University, Bridgewater, Massachusetts.

The Foundations of Mathematics:  
Axiomatic Systems and Incredible Infinities

Catherine Theresa Ferris

Submitted in Partial Completion of the  
Requirements for Departmental Honors in Mathematics

Bridgewater State University

December 16<sup>th</sup>, 2021

Dr. Ward Heilman, Thesis Advisor

Dr. John Pike, Committee Member

Dr. Rachel Stahl, Committee Member

# Contents

1	Abstract	2
2	Axiomatic Systems	2
3	The Language of Set Theory	5
4	Relations	7
5	Zermelo-Fraenkel Axioms	13
6	Ordinals and Cardinals	16
7	More Results on Infinity	24
8	Axiom of Choice	35
9	Continuum Hypothesis	61
10	Appendix 1: Other Axiomatic Schemes	63
11	Appendix 2: Definitions	67
	References	74

# 1 Abstract

Often, people who study mathematics learn theorems to prove results in and about the vast array of branches of mathematics (Algebra, Analysis, Topology, Geometry, Combinatorics, etc.). This helps them move forward in their understanding; but few ever question the basis for these theorems or whether those foundations are sufficient or even secure. Theorems come from our foundations of mathematics, Axioms, Logic and Set Theory. In the early 20th century, mathematicians set out to formalize the methods, operations and techniques people were assuming. In other words, they were formulating axioms. The most common axiomatic system is known as the Zermelo-Fraenkel axioms with the addition of the Axiom of Choice (AC), although AC is still a controversial axiom. These axioms helped build our notion of infinity, which turns out to be much more complicated than what had been suspected. Earlier mathematicians from the Greeks to Gauss even refused to view infinity as an actuality, and only referred to “potential infinities”. Now our axioms allow multiple sizes of infinity. With the help of ordinals and cardinals, we can start to see a framework for these previously obscure notions. This leads to the even more difficult question: what *exactly* are the real numbers? The Continuum Hypothesis provides one approach, yet the full story remains “unsolved” by mathematicians to this day.

# 2 Axiomatic Systems

**Definition 2.1.** *Axiom:* In mathematics, an *axiom* is an unproven rule that is accepted as true because it is self-evident.

**Definition 2.2.** *Axiomatic System:* An *axiomatic system* is any set of axioms which can be used together to derive theorems.

*“The axiomatic method consists in accepting without proof certain propositions as axioms or postulates and then deriving from the axioms all other propositions of the system of theorems. The axioms constitute the ‘foundations’ of the system; the theorems are the ‘superstructure’, and are obtained from the axioms with the exclusive help of principles of logic” (Nagel &*

*Newman, 2001, p.2-3).*

There are three goals we have when it comes to axiomatic systems: independence, consistency, and completeness.

## 2.1 Independence

An axiomatic system is independent if each of its axioms in the system is independent. That is, each axiom can not be proven from any of the other axioms in the system. Independence is a useful tool, but not essential.

## 2.2 Consistency

An axiomatic system is consistent if the axioms are not contradictory; if there is no proof of a contradiction in the theory. Gödel's Second Incompleteness Theorem states that no axiomatic theory as least as powerful as the axiomatic arithmetic can be proven consistent by methods that are formalized in that theory itself. So, it is not possible to prove consistency inside any axiomatic system that contains the basic theory of arithmetic. *"By Gödel's Second Incompleteness Theorem, it is impossible to show that consistency of ZF (or related theories) by means limited to ZF alone. Once we assume that ZF (or ZFC) is consistent, we may ask whether the theory remains consistent if we add an addition axiom A"* (Jech, 2003, p. 163). In this case, we can say  $ZF + C$  is consistent relative to  $ZF$  or  $ZFC + A$  is consistent relative to  $ZFC$ . Our systems should only be able to produce true statements. Thus, if our axiomatic systems were inconsistent, we would be able to prove false statements. 100+ years of mathematics has failed to find any inconsistencies, which intuitively suggests but does not prove that the system is consistent.

*"...Gödel's discovery of the Incompleteness Theorem ... implies that the consistency of a mathematical system cannot be proved except by methods more powerful than those of the system itself (Cohen, 1966, p. 3)".*

## 2.3 Completeness

An axiomatic system is complete if every true statement can be proven from the axioms. In 1930, mathematician Kurt Gödel proved the Incompleteness Theorem. The Incompleteness Theorem states that in any sufficiently complex axiomatic system, there must exist true statements that can not be proven inside that system. In other words, every sufficiently complex axiomatic system is incomplete, which is a call for the creation of new and stronger axioms. Completeness of powerful axiomatic systems is then an unreachable goal because the theorem states that there will always be stronger axioms out there. *“For every problem there is a solution and for every solution there is a new problem”* (Koellner, 2010, p. 2). Gödel’s incompleteness result reinforces the importance of a search for new axioms that will expand the theory of mathematics and allow for the proof of a wider class of theorems. For example, adding the Axiom of Choice (AC) to the Zermelo-Fraenkel (ZF) axioms allows us access to more results that we could not obtain solely with ZF.

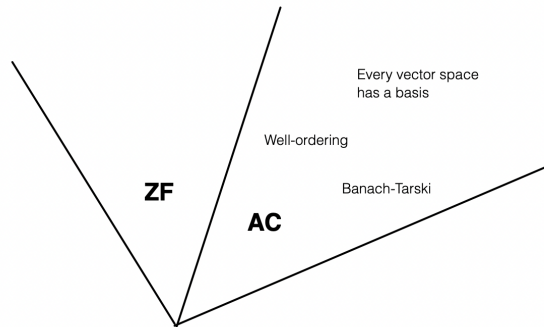


Figure 1: Diagram showing that the Zermelo-Fraenkel (ZF) axioms are not complete because we can add an important, independent axiom, the Axiom of Choice (AC), to it to obtain more results.

*“ ... A climate of opinion was thus generated in which it was tacitly assumed that each sector of mathematical thought can be supplied with a set of axioms sufficient for developing systematically the endless totality of true propositions about the given area of inquiry.*

*Gödel's paper showed that this assumption is untenable. He presented mathematicians with the astounding and melancholy conclusion that the axiomatic method has certain inherent limitations, which rule out the possibility that even the properties of the non-negative integers can ever be fully axiomatized ... In light of these conclusions, no final systematization of many important areas of mathematics is attainable, and no absolutely impeccable guarantee can be given that many significant branches of mathematical thought are entirely free from internal contradiction" (Nagel & Newman, 2001, p. 4-5).*

### 3 The Language of Set Theory

The Language of Set Theory (LAST) is a formal language that has a determined set of symbols and syntax, which ensures us that the concept of a 'collection desirable in the language' is rigorously defined. The symbols are comprised of the names of sets, variables, identities, quantifiers, and punctuation. The syntax is comprised of formulas, which look like phrases or sentences. We will use both capital letters and small letters to denote sets. Elements which are not sets will be denoted by small letters.

**Example 3.1.** Symbols:  $\vee, \wedge, \neg, \forall, \exists, \rightarrow, x, y, z, \dots$

**Example 3.2.** Formula:  $(\phi \vee \psi)$  Non-Formula:  $\phi \neg \forall$

In addition, we can define many standard operations on sets using LAST. These operations are:

- Subset ( $\subseteq$ ): If  $X$  and  $Y$  are sets, we say that  $X$  is a *subset* of  $Y$  if and only if every element of  $X$  is an element of  $Y$ , denoted by  $X \subseteq Y$ .
- Power Set ( $\mathcal{P}$ ): For any set  $S$ , the *power set* of  $S$ , denoted by  $\mathcal{P}(S)$  is the set of all subsets of  $S$ . Axiom 5.1 allows us to work with the power set as a set.
- Union ( $\cup$ ): If  $X$  and  $Y$  are sets, the *union* of  $X$  and  $Y$  is the set consisting of the members of  $X$  together with the members of  $Y$ , denoted by  $X \cup Y$ .

- Intersection ( $\cap$ ): If  $X$  and  $Y$  are sets, the *intersection* of  $X$  and  $Y$  is the set consisting of those objects that are members of both  $X$  and  $Y$ , denoted by  $X \cap Y$ .
- Set Difference ( $-$ ): If  $X$  and  $Y$  are sets, the *difference* of sets  $X$  and  $Y$  is the set consisting of those elements of  $X$  that are not elements of  $Y$ , denoted by  $X - Y$ .

### 3.1 Sheffer Stroke

While having all of these symbols is helpful in the comprehension of set theory, it is entirely possible to express these formulas in terms of a single symbol. This single symbol is called the *sheffer stroke* operation. The sheffer stroke operation is equivalent to the “not and” operation, denoted by the symbol  $|$ . So, we can write  $\neg(P \wedge Q)$  as  $P|Q$ . Given this, all mathematically defined formulas in logic can be defined using only the sheffer stroke. It can model operations such as “negation”, “or”, “and”, “implication”, etc. Therefore, it is true that connectors in LAST are not independent. The sheffer stroke is a commutative operation, but not associative. We see that  $P|Q = Q|P$ , but  $P|(Q|R) \neq (P|Q)|R$ .

<b>P</b>	<b>Q</b>	<b><math>P \wedge Q</math></b>	<b><math>\neg(P \wedge Q)</math></b>	<b><math>P Q</math></b>
T	T	T	F	F
T	F	F	T	T
F	T	F	T	T
F	F	F	T	T

Figure 2: Truth table showing  $\neg(P \wedge Q)$  is equivalent to  $P|Q$ .

<b>P</b>	<b>Q</b>	<b><math>P Q</math></b>	<b><math>Q P</math></b>
T	T	F	F
T	F	T	T
F	T	T	T
F	F	T	T

Figure 3: Truth table showing the sheffer stroke is a commutative operation.



P	Q	R	$P Q$	$(P Q) R$	$Q R$	$P (Q R)$
T	T	T	F	T	F	T
T	T	F	F	T	T	F
T	F	T	T	F	T	F
T	F	F	T	T	T	F
F	T	T	T	F	F	T
F	T	F	T	T	T	T
F	F	T	T	F	T	T
F	F	F	T	T	T	T

Figure 4: Truth table showing the sheffer stroke is not an associative operation.

P	Q	$P \rightarrow Q$	$P Q$	$P (P Q)$	$Q Q$	$P (Q Q)$
T	T	T	F	T	F	T
T	F	F	T	F	T	F
F	T	T	T	T	F	T
F	F	T	T	T	T	T

Figure 5: Truth table showing the sheffer stroke models the implication connector ( $\rightarrow$ ).

P	Q	$P \wedge Q$	$P Q$	$P Q$	$(P Q)   (P Q)$
T	T	T	F	F	T
T	F	F	T	T	F
F	T	F	T	T	F
F	F	F	T	T	F

Figure 6: Truth table showing the sheffer stroke models the “and” connector ( $\wedge$ ).

## 4 Relations

**Definition 4.1.** *Positive Natural Numbers:* The *positive natural numbers*, denoted by  $\mathbb{N}$ , is the set  $\{1, 2, 3, 4, \dots, n, \dots\}$ . When discussing the positive natural numbers, we will start

the set at 1. When discussing the natural numbers,  $\mathbb{N}^0$ , we will adopt the usual convention that logicians and computer scientists start the set at 0.

**Definition 4.2.** *Ordered Pair:* We define an ordered pair  $(a, b)$  to be the set  $\{\{a\}, \{a, b\}\}$ .

**Definition 4.3.** *Cartesian Product:* A Cartesian product of sets  $X$  and  $Y$  is the set of all ordered pairs which is defined to be the set  $X \times Y = \{(a, b) \mid a \in X \wedge b \in Y\}$ .

**Example 4.1.** Given two sets  $X$  and  $Y$  where  $X = \{a, b, c\}$  and  $Y = \{d, e, f\}$ , we have the Cartesian product  $X \times Y = \{(a, d), (a, e), (a, f), (b, d), (b, e), (b, f), (c, d), (c, e), (c, f)\}$ .

**Definition 4.4.** *Binary Relation:* A binary relation over sets  $X$  and  $Y$  is a subset  $R$  of  $X \times Y$ . If  $X = Y$ , we say  $R$  is a binary relation on  $Y$ . In general, an  $n$ -ary relation over  $A_1, \dots, A_n$  is a subset of  $A_1 \times \dots \times A_n = \{(a_1, \dots, a_n) \mid a_k \in A_k \text{ for each } k = 1, \dots, n\}$ . We say  $aRb$  if  $(a, b) \in R$ .

**Example 4.2.** We let  $X = \{1, 2\}$  and let  $\leq$  be the standard less than or equal to relation on the set  $X$  where  $R = \{(a, b) \in X \times X \mid a \leq b\}$ . Then,  $R = \{(1, 1), (1, 2), (2, 2)\}$ .

**Definition 4.5.** *Reflexive:* Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is reflexive if and only if  $(\forall a \in X)(aRa)$ .

**Example 4.3.** We let  $X = \{1, 3, 5, 7, 9\}$  and let  $\leq$  be the standard less than or equal to relation on the set  $X$ . Then, for every element  $a$  in  $X$ , it is true that  $a \leq a$ . For instance,  $7 \leq 7$ . Thus,  $\leq$  is reflexive on  $X$ .

**Example 4.4.** We let  $Z = \{8, 9, 10, 11, 12\}$  and let  $<$  be the standard less than relation on the set  $z$ . We see that this is a non-reflexive relation because there exists an element  $a$  in  $X$  such that  $a$  is not related to  $a$ . For instance,  $9 \not< 9$ .

**Definition 4.6.** *Anti-Symmetric:* Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is anti-symmetric if and only if  $(\forall a, b \in X)[(aRb \wedge bRa) \rightarrow a = b]$ .

**Example 4.5.** We let  $X = \{2, 4, 6, 8\}$  and let  $\leq$  be the standard less than or equal to relation on the set. Then, for any two elements in the set  $a$  and  $b$ , if it is true that  $a \leq b$  and  $b \leq a$ , then  $a = b$ . For instance,  $2 \leq 2 \wedge 2 \leq 2 \rightarrow 2 = 2$ .

**Definition 4.7.** *Transitive:* Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is transitive if and only if  $(\forall a, b, c \in X)[(aRb \wedge bRc) \rightarrow (aRc)]$ .

**Example 4.6.** We let  $X = \{1, 2, 3, 5, 7, 8\}$  and let  $\leq$  be the standard less than or equal to relation on the set. Then, it is true that for any  $a, b, c$  in the set  $X$ , if  $a \leq b$  and  $b \leq c$  then  $a \leq c$ . For instance,  $2 \leq 5 \wedge 5 \leq 8 \rightarrow 2 \leq 8$ .

**Example 4.7.** Let  $X = \{n, \{n, p\}, \{\{n, p\}, q\}\}$  and let  $\in$  be the standard membership relation on the set. Now  $n \in \{n, p\}$  and  $\{n, p\} \in \{\{n, p\}, q\}$  but  $n \notin \{\{n, p\}, q\}$ . Thus, we see that  $\in$  is not a transitive relation.

**Definition 4.8.** *Connected:* Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is connected if and only if  $(\forall a, b \in X)(a \neq b) \rightarrow [(aRb) \vee (bRa)]$ .

**Remark 1.** We can show by LAST and a known logical equivalence  $P \rightarrow Q \equiv \neg P \vee Q$ , that  $(\forall a, b \in X)(a \neq b) \rightarrow [(aRb) \vee (bRa)]$  is logically equivalent to the trichotomy  $(\forall a, b \in X)(a = b) \vee (aRb) \vee (bRa)$ . We shall call  $P$  the statement  $(a \neq b)$  and  $Q$  the statement  $[(aRb) \vee (bRa)]$ . Thus, we have  $P \rightarrow Q$  which is only false when  $P$  is true and  $Q$  is false. Hence,  $\neg P$  is the statement  $(a = b)$  and  $Q$  is the statement  $(aRb \vee bRa)$  which gives us the statement  $\neg P \vee Q$ . This statement is also only false when  $P$  is true and  $Q$  is false. Thus, these definitions of connected are logically equivalent. So, we can use the equivalent definition of connected and say that the elements in the relation are *comparable*. Hence, in a connected relation, either  $a = b$  or  $a < b$  or  $b < a$ . If we look at Figure 7, we see that  $\{a, c\} \not\subseteq \{b, c\}$  and  $\{b, c\} \not\subseteq \{a, c\}$  and  $\{a, c\} \neq \{b, c\}$ . These are called *incomparable* elements.

**Example 4.8.** Let  $X = \{1, 2, 3, 4, \dots, n\}$ , and take the natural numbers 2 and 3. If  $2 \neq 3$ , then  $2 < 3 \vee 3 < 2$ . Since this is true for all pairs of positive natural numbers, the elements of  $X$  are comparable under the standard less than ( $<$ ) relation.

**Example 4.9.** We let  $X = \{n, \{m, p\}\}$  and let  $\in$  be the standard membership relation on the set. We have that  $n \neq \{m, p\}$ , but  $n \notin \{m, p\}$  and  $\{m, p\} \notin n$ . Thus, we see that  $\in$  is not a connected relation or the elements under  $\in$  are not necessarily comparable.

**Definition 4.9.** *Partial Ordering:* A partial ordering of a set  $X$ , denoted by  $\preceq$ , is a binary relation on that set which is reflexive, anti-symmetric, and transitive.

**Definition 4.10.** *Partially Ordered Set (poset):* A partially ordered set is an ordered pair  $(X, \preceq)$ , consisting of a set  $X$  together with a partial ordering  $\preceq$  of  $X$ .

**Example 4.10.** Let  $A = \{a, b, c\}$  where  $a, b, c$  are all distinct elements and let  $X = \mathcal{P}(A)$ . The standard set containment operation ( $\subseteq$ ) is reflexive, anti-symmetric, and transitive. We have that  $\subseteq$  is reflexive because  $\forall y \in \mathcal{P}(A), y \subseteq y$ . For example,  $\{b\} \subseteq \{b\}$ . We also have that  $\subseteq$  is anti-symmetric because  $\forall y, z \in \mathcal{P}(A)$ , if  $y \subseteq z$  and  $z \subseteq y$  then  $y = z$ . For example,  $\{a, b\} \subseteq \{a, b\}$  and  $\{a, b\} \subseteq \{a, b\}$ , so  $\{a, b\} = \{a, b\}$ . Finally, we have that  $\subseteq$  is transitive because  $\forall w, y, z \in \mathcal{P}(A)$ , when  $w \subseteq y$  and  $y \subseteq z$ , then  $w \subseteq z$ . For example, since  $\{b\} \subseteq \{a, b\}$  is true and  $\{a, b\} \subseteq \{a, b, c\}$  is true, then  $\{b\} \subseteq \{a, b, c\}$  is true as well. Thus, we have that  $(\mathcal{P}(A), \subseteq)$  is a poset. See the Hasse diagram in Figure 7 for reference. A Hasse diagram is a graphical representation of the relation of elements of a poset with an implied upward orientation.

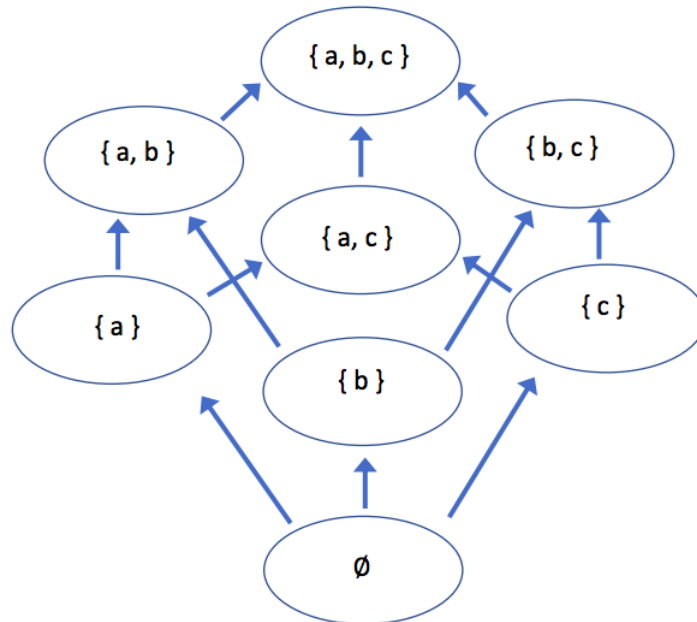


Figure 7: Example of a partial ordering represented by set containment.

**Example 4.11.** Given any three positive natural numbers,  $a, b, c$ , where  $\leq$  is the standard ordering of the positive natural numbers, the following is true:  $\leq$  is a reflexive, anti-symmetric, and transitive relation. Thus, we have that  $(\mathbb{N}, \leq)$  is a poset.

**Example 4.12.** Consider the set  $E = \{2, 3, 5, 6, 8, 10, 12, 15, 24\}$ . For all  $n, m \in E$ , we define  $n|m$  as there exists  $k \in \mathbb{N}$  such that  $m = n \cdot k$ . In Figure 8 we display the poset using a Hasse diagram.

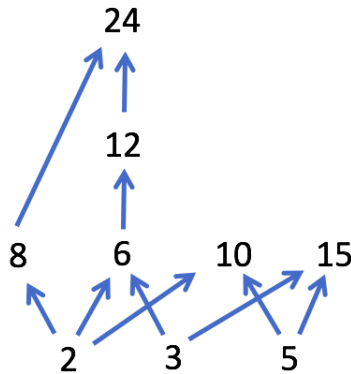


Figure 8: Hasse diagram of a poset.

**Definition 4.11.** *Total Ordering (linear ordering):* A total ordering ( $\preceq$ ) of a set  $X$  is a partial ordering of  $X$  whose elements are comparable.

**Definition 4.12.** *Totally Ordered Set (toset):* A pair  $(X, \preceq)$  is a totally ordered set if  $\preceq$  is a total ordering of the set  $X$ .

**Example 4.13.** The real numbers  $\mathbb{R}$  together with the standard less than or equal to ( $\leq$ ) relation is a totally ordered set.  $(\mathbb{R}, \leq)$  is reflexive, anti-symmetric, transitive, and its elements are comparable. Observe that Example 4.11 is also an example of a toset.

**Example 4.14.** Consider the set  $R = \{e, \pi, \sqrt{2}, 199\}$  ordered by the less than or equal to relation ( $\leq$ ). We will display the toset by again using a Hasse diagram. Any Hasse diagram of a totally ordered set will look like Figure 9, which is why we call a toset a *linear ordering* or a *chain*.



Figure 9: Hasse diagram of a toset.

**Definition 4.13.** *Minimal Element:* Let  $(X, \preceq)$  be a poset and let  $Y \subseteq X$ . An element  $a$  of  $Y$  is a minimal element of  $Y$  if and only if there is no  $b$  in  $Y$  such that  $b \preceq a$ . That is,  $a$  is minimal in  $Y$  with respect to the given ordering.

**Example 4.15.** Looking back at Example 4.12 and Figure 8, we see that some of the minimal elements of  $(E, |)$  are 2 from subset  $\{2, 8, 24\}$ , 3 from subset  $\{3, 6, 12, 24\}$ , and 6 from subset  $\{6, 12, 24\}$ .

**Example 4.16.** Looking back at Example 4.10, we can take the subset  $B = \{\{a\}, \{a, b\}, \{b\}, \{a, b, c\}\}$ . In this scenario, there are multiple minimal elements, namely,  $\{a\}$  and  $\{b\}$ . Observe that neither minimal element is the least element of  $B$ , since  $B$  does not contain a least element.

**Definition 4.14.** *Well-Founded:* A partially ordered set is well-founded if every nonempty subset has a minimal element.

**Example 4.17.** The relation  $R \subset \mathbb{N} \times \mathbb{N}$  defined by  $nRm$  where  $m = n + 1$  means a relation in which the natural number  $m$  is the successor of the natural number  $n$ . This set is well-founded since every non-empty subset  $S$  of natural numbers contains at least one element  $r$

that is not the successor of any element of  $S$ . For instance, the number 3 is not a successor of any of the elements of the subset  $\{3, 4, 5, 6\}$ .

**Definition 4.15.** *Well-Ordering:* A well-ordering of a set  $X$  is a well-founded, total ordering  $(\preceq)$  of  $X$ .

**Definition 4.16.** *Well-Ordered Set (woset):* A pair  $(X, \preceq)$  is a *well-ordered set* when  $X$  is a well-founded set together with a total ordering  $(\preceq)$ .

**Example 4.18.** We start with the set of positive natural numbers

$\mathbb{N} = \{1, 2, 3, 4, \dots, n, \dots\}$  and the standard less than or equal to relation  $\leq$ . We see that  $\mathbb{N}$  is a totally ordered set because it is reflexive, anti-symmetric, transitive, and the elements of  $\mathbb{N}$  under  $\leq$  are comparable. We also observe that  $(\mathbb{N}, \leq)$  is well-founded because every non-empty subset of the natural numbers contains a minimal element. Therefore,  $(\mathbb{N}, \leq)$  is a well-ordered set.

**Example 4.19.** We start with the set of real numbers  $\mathbb{R}$  on the interval  $(0, 1)$  and the standard less than or equal to relation  $\leq$ . We see that the set  $\{\frac{1}{2^n} : n \in \mathbb{N}\}$  is not a well-ordered set because it is not well-founded. That is,  $\{\frac{1}{2^n} : n \in \mathbb{N}\}$  has no least element. It is a set that gets increasingly smaller.

## 5 Zermelo-Fraenkel Axioms

The particular set of axioms described here is not an independent set, but chosen for convenience and utility. The eight axioms listed here are considered ZF. Adding a ninth axiom, the Axiom of Choice, turns this into ZFC. The Axiom of Choice is discussed later in this thesis. We denote sets in these axiom definitions with lower case letters.

**Axiom 5.1.** *Power Set Axiom:* If  $x$  is a set, there is a set that consists of all and only the subsets of  $x$ . The formal definition in LAST is:  $\forall x \exists y \forall v (v \in y \leftrightarrow v \subseteq x)$ . We denote the power set with  $\mathcal{P}$ .

**Example 5.1.** Given  $x = \{a, b, c\}$ , then  $\mathcal{P}(x) = y = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$ .

**Example 5.2.** Given the set  $x = \{v_2, v_1, v_a, v_c, v_3, v_b, \dots\}$ . Let  $w_2 = \{\}$ ,  $w_a = \{v_c\}$ ,  $w_5 = \{v_c, v_3\}$ ,  $w_e = \{v_2, v_a, v_b\}, \dots$ . Then  $y = \{w_2, w_a, w_5, w_e, \dots\}$ , which is all the subsets of  $x$ . No particular pattern or order is implied by the subscripts here.

**Axiom 5.2.** *Axiom of Union:* For any set  $x$  there exists a set  $y$  such that  $y = \bigcup x$ . If  $x$  is a set, then there is a set whose members are precisely the members of the members of  $x$ . The formal definition in LAST is:  $\forall x \exists y \forall v (v \in y \leftrightarrow \exists z (z \in x \wedge v \in z))$ . This says that if  $v$  is an element of a set  $z$ , which is an element of a set  $x$ , then  $v$  is in the union of  $x$ , so  $v$  is an element of  $y$ .

**Remark 2.** Given the set  $x$ ,  $\bigcup x = \{v | \exists y (y \in x \wedge v \in y)\}$ , the union of  $x$  is a set.

**Example 5.3.** If  $x = \{v_1, v_2, v_3, v_a, v_b, v_c\}$  and  $v_1 = \{a, b, c\}$ ,  $v_2 = \{b, c, d\}$ ,  $v_3 = \{m, n\}$ ,  $v_a = \{n, p, q\}$ ,  $v_b = \{a, r\}$ ,  $v_c = \{t\}$ , then  $\bigcup x = \{a, b, c, d, m, n, p, q, r, t\}$ .

**Axiom 5.3.** *Axiom Schema of Replacement:* Let  $\phi(v_n, v_m)$  be any formula of LAST, such that for each set  $a$  there is a unique set  $b$  such that  $\phi(a, b)$ . Let  $x$  be a set. Then there is a set  $y$  consisting of just those  $b$  such that  $\phi(a, b)$  for some  $a$  in  $x$ . The formal definition in LAST is:  $\forall x \exists y \forall a \in x (\exists b \phi(a, b) \rightarrow \exists b \in y \phi(a, b))$ . The replacement axiom basically says, if the domain of a function is a set, then the range of that function is also a set. Since this is an infinite collection of possible axioms, it is described as an axiom schema.

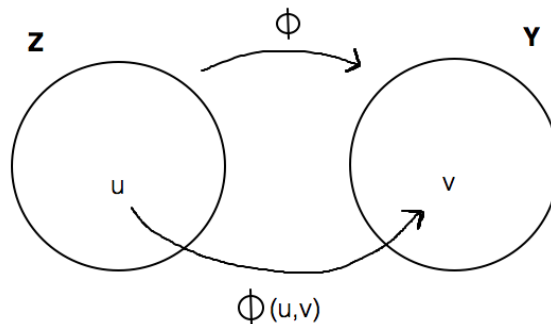


Figure 10: The range of a function is a set



**Axiom 5.4.** *Null Set Axiom:* There is a set which has no members. This set is denoted by the symbol  $\emptyset$  or by the notation  $\{\}$ .

**Axiom 5.5.** *Axiom of Infinity:* There is a set  $x$  such that  $\emptyset \in x$  and for every  $a \in x$ ,  $a \cup \{a\} \in x$ .

**Axiom 5.6.** *Axiom of Extensionality:* If two sets have identical elements, then they are equal.  $x = y \leftrightarrow \forall a [(a \in x) \leftrightarrow (a \in y)]$

**Axiom 5.7.** *Axiom Schema of Subset Selection:* For any set  $x$  and for any formula  $\varphi(v)$ , there is a set  $y$  in  $\mathcal{P}(x)$ , whose members are exactly those elements,  $a$ , such that  $\varphi(a)$ . In other words,  $y$  is a subset of  $x$  and  $y$  is just those elements of  $x$  that satisfy the formula  $\varphi$ . For a particular formula of LAST, say,  $\psi$ , the particular instance of this schema can be expressed as  $\forall x \exists y \forall u [(u \in y) \leftrightarrow (u \in x \wedge \psi(u))]$ . In general, this axiom schema allows us to form set  $y$ , which is a subset of some given set  $x$ , and whose members satisfy some property  $\psi$ .

**Example 5.4.** Let  $x$  be the set of positive natural numbers and let  $\psi$  be the formula  $\psi(u) = u < 7$ . Then, there is a set  $y$ , consisting of all elements  $u$ , if and only if  $u$  is in the set of positive natural numbers and  $u$  is less than 7. Thus,  $y = \{1, 2, 3, 4, 5, 6\}$ .

**Remark 3.** *The Potential Axiom Schema of Comprehension:* The axiom schema of subset selection replaces an earlier attempt to axiomatize the building of sets. This earlier attempt was called the axiom schema of comprehension, which says that if  $\varphi$  is a property, then there exists a set  $y$  such that  $y = \{x : \varphi(x)\}$ . This gives us Russell's Paradox: Consider the set  $s$  whose elements are all those (and only those) sets that are not members of themselves:  $s = \{x : x \notin x\}$ . So, the question is, does  $s$  belong to  $s$ ? If  $s$  belongs to  $s$ , then  $s$  is not a member of itself, so  $s \notin s$ . On the other hand, if  $s \notin s$ , then  $s$  belongs to  $s$ . In either case, we have a contradiction. Thus, we must conclude that  $\{x : x \notin x\}$  is not a set, and we must revise the intuitive notion of a set (Jech, 2003).

**Axiom 5.8.** *Axiom of Foundation:* The axiom of foundation says that  $\in$  is a well-founded relation. A more direct way of expressing this is: Given any set  $s$ , every nonempty subset  $x$  of  $s$  has a minimal element with respect to the  $\in$  relation.

## 6 Ordinals and Cardinals

### 6.1 Ordinals

What if someone asked you to describe the number 7? How would you do it? Growing up, we learned our numbers by using visuals, grouping objects together and using a number to describe how many there are. This is why the natural numbers are also called the “counting numbers”, because we can count objects to describe how many of them there are. But, without those visuals and tangible objects, it seems harder to describe what a natural number is. This is where ordinal numbers come in. Ordinal numbers are a formal way to describe our natural numbers with an ordering. We will adopt the convention that logicians and computer scientists include zero as part of the natural numbers, denoting it by  $\mathbb{N}^0$ .

**Definition 6.1.** *Ordinal Number:* Let  $(X, \preceq)$  be a woset. For  $a \in X$ , let the segment determined by  $a$  be the set  $X_a = \{x \in X : x < a\}$ . An *ordinal number* is defined to be a woset  $(X, \preceq)$  such that  $X_a = a$  for  $a \in X$ .

**Example 6.1.** Let us go back to our question in the beginning of this section. We can describe the number 7 through our ordinal definition. We have that  $X_a = a$  for all  $a \in X$  when  $X_a = \{x \in X : x < a\}$ . So, when  $a = 7$ , we have  $X_7 = \{x \in X : x < 7\} = \{0, 1, 2, 3, 4, 5, 6\}$ .

The null set axiom gives us  $\{\}$  or  $\emptyset$ . The axiom of infinity lets us define our successors via  $f(\emptyset) = \emptyset \cup \{\emptyset\}$ . So, we obtain our initial successor ordinals by defining

$$0 = \emptyset$$

$$1 = f(0) = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\}$$

$$2 = f(1) = 1 \cup \{1\} = \{0, 1\}$$

$$3 = f(2) = 2 \cup \{2\} = \{0, 1, 2\}$$

...

$$(n + 1) = \{0, 1, 2, 3, \dots, n\}.$$

Continuing in this manner, we define all of the finite ordinals or natural numbers, getting up to any element in  $\mathbb{N}^0$ . We denote the first infinite ordinal by  $\omega = \mathbb{N}^0$  (so,  $\omega$  equals all of  $\mathbb{N}^0$ ). We will continue the hierarchy of transfinite ordinals after defining some ordinal arithmetic. We see that the natural numbers are merely symbols for our finite ordinals. Thus, ordinals are a way to formally describe what we typically informally see as the “counting numbers”.

## 6.2 Ordinal Arithmetic

### 6.2.1 Addition

Given ordinals  $\alpha$  and  $\beta$ , we can define ordinal addition  $\alpha + \beta$ . First, we set

$$A = (\alpha \times \{0\}) \cup (\beta \times \{1\})$$

which represents the cartesian product of all the ordinals  $\alpha$  with  $\{0\}$ ,  $(\alpha, \{0\})$ , together with the cartesian product of all the ordinals  $\beta$  with  $\{1\}$ ,  $(\beta, \{1\})$ . Some of the elements of  $A$  could look like  $(7, 0)$ ,  $(4, 0)$ ,  $(10, 0)$ ,  $(800, 1)$ ,  $(35, 1)$ ,  $(2, 1)$ ,  $\dots$ . We define a well-ordering of  $A$  by

$$(\alpha, a) <_A (\beta, b) \leftrightarrow (a < b) \vee (a = b \wedge \alpha < \beta)$$

which states that ordinal addition is ordered by first comparing the second position of the ordered pairs. If  $a < b$ , then  $(\alpha, a)$  is placed before  $(\beta, b)$  in the ordering. But, if  $a = b$ , then we compare the first position of the ordered pairs. If  $\alpha < \beta$ , then again  $(\alpha, a)$  is placed before  $(\beta, a)$  in the ordering.

For example, say we are comparing  $(5, 0)$  and  $(70, 1)$ . Looking at the  $a$  and  $b$  positions, we see that  $0 < 1$ . Thus,  $(5, 0)$  is placed before  $(70, 1)$  in the ordering. But, say we have  $(5, 0)$  and  $(3, 0)$ . Since the  $a$  and  $b$  positions are equal, we then look at the  $\alpha$  and  $\beta$  positions. Thus, we see that  $3 < 5$ , which means  $(3, 0)$  will be placed before  $(5, 0)$  in the ordering. We note that ordinal addition is an associative operation, but not commutative.

In set theory, we define ordinals in two types: successor ordinals and limit ordinals. Defining these two types helps us when constructing our ordinals into the transfinite.

**Definition 6.2.** *Successor Ordinal:* A *successor ordinal* is an ordinal that comes directly after some ordinal in the ordering. For example,  $\beta$  is a successor ordinal if  $\beta = \alpha + 1$  for an ordinal  $\alpha$ .

**Definition 6.3.** *Limit Ordinal:* A *limit ordinal* is an ordinal that will have no greatest member; it is not the successor of any ordinal. Thus,  $\omega$  is a limit ordinal. Also, zero is a limit ordinal because it is not the successor of any ordinal.

### 6.2.2 Multiplication

Given the ordinal  $\lambda$ , let  $\langle \alpha_\eta \mid \eta < \lambda \rangle$  be a  $\lambda$ -sequence of ordinals. The ordinal sum  $\sum_{\eta < \lambda} \alpha_\eta$  is defined as follows: Set

$$A = \bigcup_{\xi < \lambda} (\alpha_\xi \times \{\xi\})$$

which represents the Cartesian product of each  $\alpha$  (sub-scripted with some  $\xi$ ) with  $\xi$  up to some number  $\lambda$ . Similarly, define a well-ordering of  $A$  by

$$(m, n) <_A (o, p) \leftrightarrow (n < p) \vee (n = p \wedge m < o).$$

Denote the ordinal sum  $\sum_{\eta < \lambda} \alpha_\eta$  by  $Ord(A, <_A)$ . Hence,  $\sum_{\eta < \lambda} \alpha_\eta = Ord(A, <_A)$ . So, the ordinal sum is defined to be the well-ordering of  $A$ . We see that  $\sum_{\eta < \lambda} \alpha_\eta$  is the ordinal sum that begins with  $\alpha_0$ , then has  $\alpha_1$  steps, then  $\alpha_2$  steps, and continues up through  $\xi < \lambda$ . For example,  $\sum_{\xi < 3} \alpha_\xi = \alpha_0 + \alpha_1 + \alpha_2$ . So, what  $\sum_{\xi < 3} \alpha_\xi$  represents is all the things in  $\alpha_0$ , followed by all the things in  $\alpha_1$ , followed by all the things in  $\alpha_2$ . When it comes to the well-ordering of elements in  $A$ , it is very similar to the well-ordering of  $A$  in ordinal addition. The second position,  $n$ , in the coordinate  $(m, n)$  tells us the ordinal pile we picked from, while the first position,  $m$ , tells us which element we picked from that ordinal pile. Suppose we have two elements  $(27, 1)$  and  $(5, 2)$ . This means we picked 27 from the first ordinal pile and 5 from

the second ordinal pile. Since  $1 < 2$ , we can place  $(27, 1)$  before  $(5, 2)$  in our ordering.

Now, ordinal multiplication is the special case of this general situation. We define ordinal multiplication as iterated addition, that is

$$\alpha \cdot \beta = \sum_{\xi < \beta} \alpha$$

where  $\alpha \cdot \beta$  denotes  $\beta$  copies of  $\alpha$ . In simpler terms, what is happening is that  $\alpha$  is being duplicated  $\beta$  times and is well-ordered by our definition of the well-ordering of  $A$ . This is true for any finite ordinal  $n$ , where  $\alpha \cdot n = \alpha + \alpha + \dots + \alpha$ ,  $n$  times. We note that ordinal multiplication is associative, is left distributive but not necessarily right distributive, and is not a commutative operation. For instance,  $3 \cdot \omega = \omega$  but  $\omega \cdot 3 = \omega + \omega + \omega > \omega$ .

$$A = \bigcup_{\xi < 3} (\alpha_\xi \times \{\xi\}) \rightarrow \text{cartesian product of } \alpha \text{ with } \{0,1,2\} \rightarrow (\alpha, 0), (\alpha, 1), (\alpha, 2)$$

$$\begin{aligned} (\alpha, 0) &= \alpha_0 = 0,1,2,3,4, \dots, \alpha \rightarrow \text{first copy of ordinal} \\ (\alpha, 1) &= \alpha_1 = 0,1,2,3,4, \dots, \alpha \rightarrow \text{second copy of ordinal} \\ (\alpha, 2) &= \alpha_2 = 0,1,2,3,4, \dots, \alpha \rightarrow \text{third copy of ordinal} \end{aligned}$$

$$\sum_{\xi < 3} \alpha_\xi = \alpha_0 + \alpha_1 + \alpha_2 = 0,1,2,3,4, \dots, \alpha + 0,1,2,3,4, \dots, \alpha + 0,1,2,3,4, \dots, \alpha \rightarrow \text{combining the three copies gives us } \alpha \cdot 3 \text{ or three copies of } \alpha.$$

Figure 11: Broken down example of ordinal multiplication.

### 6.2.3 Examples of Ordinal Multiplication

**Example 6.2.** The expression  $\omega \cdot 2$  means we have 2 copies of  $\omega$ . In other words,  $\omega \cdot 2 = \sum_{\xi < 2} \omega = \bigcup_{\xi < 2} (\omega \times \{\xi\}) = (0, 0) < (1, 0) < \dots < (0, 1) < (1, 1) < \dots = \omega + \omega$ .

**Example 6.3.** The expression  $2 \cdot \omega$  means we have  $\omega$  copies of 2. In other words,  $2 \cdot \omega = \sum_{\xi < \omega} (2 \times \{\xi\}) = (0, 0) < (1, 0) < (0, 1) < (1, 1) < (0, 2) < (1, 2) < \dots = \omega$ . So, we have  $2 \cdot \omega = \omega$ . We see that  $\omega \cdot 2 > \omega$  since  $\omega \subset \omega \cdot 2$ . Therefore, we see that ordinal multiplication is not commutative. See Figure 12 for a visual representation of this.

$\omega$  copies of 2  
 01 01 01 01...  $\omega$   
 01 23 45 67...  $\omega$

$$2 \cdot \omega = \omega$$



2 copies of  $\omega$   
 01234567...  $\omega$   $\omega + 1$   $\omega + 2$   $\omega + 3$ ...

$$\omega \cdot 2 = \omega + \omega$$



Figure 12: A nice way to illustrate Example 6.2 and Example 6.3. We can see visually that  $2 \cdot \omega$  is not the same order type as  $\omega \cdot 2$ .

**Example 6.4.** In ordinal arithmetic, we have the left distributive law, but not the right distributive law. For example,  $2(\omega + 1) = 2\omega + 2$ . On the left hand side,  $2(\omega + 1)$  means we have  $(\omega + 1)$  copies of 2. Since we know  $2 = \{0, 1\}$ , we have  $(\omega + 1)$  copies of this. So,  $0, 1, 0, 1, 0, 1, 0, 1, \dots, \omega, 0, 1$ . More properly, we may write this as  $0, 1, 2, 3, 4, 5, 6, 7, \dots, \omega, \omega + 1, \omega + 2$ . On the right hand side,  $2\omega + 2$  means we have  $\omega$  copies of 2 followed by one more copy of 2. So,  $0, 1, 0, 1, 0, 1, 0, 1, \dots, \omega, 0, 1$ , which is exactly what we had for the left hand side. However,  $(\omega + 1)2$  is not the same order type as  $\omega 2 + 2$ .  $(\omega + 1)2$  is 2 copies of  $(\omega + 1)$ , which is  $0, 1, 2, 3, 4, \dots, \omega, \omega + 1, 0, 1, 2, 3, 4, \dots, \omega, \omega + 1$ . More properly, we may write this as  $0, 1, 2, 3, 4, \dots, \omega, \omega + 1, \omega + 2, \dots, \omega 2, \omega 2 + 1$ . Whereas  $\omega 2 + 2$  is 2 copies of  $\omega$  followed by another copy of 2. So,  $\omega 2 + 2$  is  $0, 1, 2, 3, 4, \dots, \omega, 0, 1, 2, 3, 4, \dots, \omega, 0, 1$ . More properly, we may write this as  $0, 1, 2, 3, 4, \dots, \omega, \omega + 1, \omega + 2, \dots, \omega 2, \omega 2 + 1, \omega 2 + 2$ . This is not the same as  $(\omega + 1)2$ . Thus, we have left distribute law but not right distribute law for ordinal arithmetic. Moreover, as we have already noted, we can not commute these ordinals.

### 6.2.4 Exponentiation

**Definition 6.4.** Ordinal exponentiation, denoted by  $\alpha^\beta$ , is defined by the recursion:

- $\alpha^0 = 1$

- $\alpha^\beta = \alpha^{\beta-1} \cdot \alpha$  if  $\beta$  is a successor ordinal.

Example:  $15^3 = 15^2 \cdot 15$

- $\alpha^\beta = \lim_{\gamma < \beta} \alpha^\gamma = \sup \{ \alpha^\gamma : \gamma < \beta \}$  if  $\beta$  is a limit ordinal. See Definition 8.18 for reference on supremum.

Example:  $(\omega^2)^\omega = \lim \left( (\omega^2)^1, (\omega^2)^2, (\omega^2)^3, \dots \right)$

Therefore,  $\alpha^\beta$  corresponds to the product of  $\alpha$  with itself taken  $\beta$  times. For example,  $\alpha^4 = \alpha^3 \cdot \alpha = \alpha \cdot \alpha \cdot \alpha \cdot \alpha$ . We note that ordinal exponentiation has similar properties to exponents:

- $\alpha^\beta \cdot \alpha^\gamma = \alpha^{\beta+\gamma}$
- $(\alpha^\beta)^\gamma = \alpha^{(\beta \cdot \gamma)}$

### 6.2.5 Examples of Ordinal Exponentiation

**Example 6.5.** Since 2 is a successor ordinal, we can write  $\omega^2$  as  $\omega^1 \cdot \omega$ . This means we have  $\omega$  copies of  $\omega$ . This means we have  $\{1, 2, 3, 4, \dots, \omega\}$ ,  $\{1, 2, 3, 4, \dots, \omega\}$ ,  $\{1, 2, 3, 4, \dots, \omega\}$ ,  $\{1, 2, 3, 4, \dots, \omega\}, \dots, \omega$ .

**Example 6.6.** Since  $\omega$  is a limit ordinal, we can write  $2^\omega$  as  $\lim_{n < \omega} 2^n = \lim \{2, 4, 8, 16, 32, 64, \dots\} = \omega$ . We see that this is not the same as our previous example  $\omega^2 = \omega^1 \cdot \omega$  because obviously  $\omega < \omega^1 \cdot \omega$ . So, just as with finite numbers, ordinal exponentiation is not commutative.

$\omega$  copies of  $\omega$

01234567...  $\omega$  01234567...  $\omega$  01234567...  $\omega$  ...

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



$\omega$  copies of 2

01 01 01 01...  $\omega$

01 23 45 67...  $\omega$

$$2 \cdot \omega = \omega$$



Figure 13: A nice way to illustrate Example 6.5 and Example 6.6. We can see visually that  $\omega^2$  is not the same order type as  $2^\omega$ .

Now that we have ordinal addition, multiplication, and exponentiation, we can elaborate on our ordinal number system to picture many ordinals. We can say  $\omega + 1 = \mathbb{N}_0 \cup \{\omega\}$ ,  $\omega + 2 = \mathbb{N}_0 \cup \{\omega, \omega + 1\}, \dots$ . Then we get to  $\omega \cdot 2 = \mathbb{N}_0 \cup \{\omega + n : n \in \mathbb{N}_0\}$ , and keep going through  $\omega \cdot n$  through  $\omega^2$  through  $\omega^n$  through  $\omega^\omega$  through  $\dots$ . However, this does not cover even a fraction of the sequence of all ordinal numbers.

$$\begin{aligned} &0, 1, 2, \dots, n, \dots, \omega, \omega + 1, \dots, \omega + n, \dots, \omega + \omega, \dots, \omega + \omega + 1, \\ &\dots, \omega + \omega + n, \dots, \omega \cdot 3, \omega \cdot 3 + 1, \dots, \omega \cdot 3 + n, \dots, \omega \cdot n, \dots, \omega \cdot \omega, \\ &\dots, \omega^3, \omega^3 + 1, \dots, \omega^3 + \omega, \dots, \omega^3 + \omega^2, \dots, \omega^4, \\ &\omega^4 + 1, \dots, \omega^\omega, \omega^\omega + 1, \dots, \omega^{\omega \cdot 2}, \dots, \omega^{\omega \cdot 3}, \dots, \omega^{\omega \cdot \omega}, \\ &\dots, \omega^{(\omega^2)}, \dots, \omega^{(\omega^3)}, \dots, \omega^{(\omega^\omega)}, \dots, \omega^{\omega^{\omega^{\omega^{\dots}}}} \end{aligned}$$

Figure 14: Extension of ordinal number system with addition, multiplication, and exponentiation of ordinals.



### 6.3 Cardinals

Before we discuss cardinals, some definitions are needed:

**Definition 6.5.** *Function:* A function  $f : A \rightarrow B$  is a subset of  $A \times B$  such that if  $(a, b), (a, b') \in f$ , then  $b = b'$ .

**Definition 6.6.** *Injection:* Given two sets  $A$  and  $B$ , a function  $f : A \rightarrow B$  is defined as an *injection* (one-to-one) if every input of  $A$  has its own unique output in  $B$ . In other words,

$$(\forall a, b \in A) f(a) = f(b) \rightarrow a = b.$$

**Definition 6.7.** *Surjection:* Given two sets  $A$  and  $B$ , a function  $f : A \rightarrow B$  is defined as a *surjection* (onto) if for every element in  $B$ , there is an element in  $A$  that maps to it. In other words,

$$(\forall b \in B)(\exists a \in A) f(a) = b.$$

**Definition 6.8.** *Bijection:* Given two sets  $A$  and  $B$ , a function  $f : A \rightarrow B$  is defined to be a *bijection* if it is both an injection and a surjection. We denote bijection with the “if and only if” (iff) symbol  $\leftrightarrow$ .

**Definition 6.9.** *Cardinal number:* A *cardinal number* is an ordinal  $\alpha$  such that for no  $\beta < \alpha$  does there exist a bijection  $f : \beta \leftrightarrow \alpha$ .

We can assign to every set a quantity that represents the size or the number of elements of that set. While ordinals give us an ordering of a set, cardinals give us the number of elements within a set. For finite sets, it is what we expect. But, for infinite sets, we will discuss a generalization of this finite concept. If  $A$  is a finite set, we say  $|A| = n$  if there is a bijection  $f : A \rightarrow [n]$  where  $[n] := \{1, 2, 3, \dots, n\}$ . The set  $A$  is finite if  $A = \emptyset$  or  $|A| = n$  for some  $n \in \mathbb{N}$ . Otherwise,  $A$  is an infinite set. The abstraction  $|A|$  is a finite cardinal. For example,  $|7| = |\{0, 1, 2, 3, 4, 5, 6\}| = 7$  because there is a bijection  $f : [7] \rightarrow [7]$ . We see that  $|n|$  is a cardinal, while  $n$  is an ordinal. Therefore, we see that, in the finite case, every ordinal is a cardinal in the finite case. In the infinite case, a cardinal is an ordinal of

a special type; there may be many ordinals that have the same cardinal number, losing its uniqueness. The ordinal  $\omega$  is the least infinite cardinal because it is the least ordinal that can not be put in one-to-one correspondence with a finite number  $n$ . We call the infinite ordinal numbers that are cardinals, alephs ( $\aleph$ ), with  $|\omega| = \aleph_0$ . All infinite cardinals are limit ordinals.

Definition 6.9 implies that since no number less than  $\omega$  can be put in one-to-one correspondence with  $\omega$ , then  $\omega$  is a cardinal. For example,  $\omega + 1$  is not a cardinal number because there is a number less than it,  $\omega$ , that can not be put in one-to-one correspondence with it. One way to map  $\omega$  to  $\omega + 1$ , if  $\omega = \{0, 1, 2, 3, 4, \dots\}$ , is mapping  $0 \rightarrow \omega + 1$ ,  $1 \rightarrow 0$ ,  $2 \rightarrow 1$ ,  $3 \rightarrow 2$ ,  $\dots$ ,  $\omega \rightarrow \omega$ . Under this correspondence,  $\omega$  is mapped to  $\omega + 1$  in a one-to-one and onto way. Thus,  $|\omega| = |\omega + 1| = \aleph_0$ . We can think of this concept informally through the “infinity motel” concept. Say there are  $\omega$  rooms with  $\omega$  guests and then another guest comes along. We can place the  $\omega + 1$  guest in the first room (room 0) and push the remainder of the guests down a room. So, instead of having room 0 with guest 0 in it, we have room 0 with guest  $\omega + 1$  in it. Then room 1 will have guest 0, room 2 will have guest 1, and so on. We see that every guest has been assigned to a room. Therefore, the number of rooms,  $\omega$ , and the number of guests  $\omega + 1$  are the same. Take any ordinal, for example,  $\omega + 7$ . Using the method illustrated above we can show that its predecessor  $\omega + 6$  has the same size as  $\omega + 7$ , and therefore  $\omega + 7$  is not a cardinal. Hence, the only possible infinite cardinal numbers are limit ordinals.

## 7 More Results on Infinity

We just saw that the size of the natural numbers is defined to be  $\aleph_0$ . Somewhat surprisingly, many seemingly different sets have this same cardinality. It can be shown that  $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\text{all algebraic numbers}| = \aleph_0$ . However, because the real numbers can be viewed as all the possible subsets of the natural numbers,  $\mathbb{N}$ , then  $|\mathbb{R}| = 2^{\aleph_0}$ . By a theorem we will show and prove in Section 7.0.2, the power set,  $\mathcal{P}(S)$  of a set  $S$  is strictly greater than the set  $S$ . That is, there is no mapping from a set  $S$  onto its power set. Thus,  $|2^{\aleph_0}| > |\aleph_0|$ .

Another surprising result about infinity is that the cardinality of the open unit interval  $(0, 1)$  in  $\mathbb{R}$  is the same as the cardinality of  $(0, 2)$  in  $\mathbb{R}$ , namely  $2^{\aleph_0}$ . We provide a geometric illustration of the proof in the figure below.

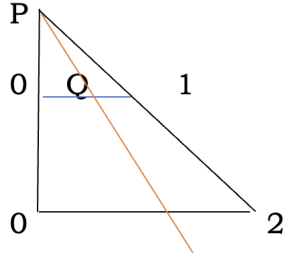


Figure 15: Visual proof showing  $|(0, 1)| = |(0, 2)|$ .

To show the one-to-one function from  $(0, 1)$  onto  $(0, 2)$  take any point, say  $Q$ , in  $(0, 1)$  and draw a line from  $P$  through  $Q$  to intersect  $(0, 2)$  at a point on  $(0, 2)$ . This demonstrates that  $|(0, 1)| = |(0, 2)|$ . In fact, this can be used to demonstrate that  $|(0, 1)| = |(m, n)|$  for all  $m, n \in \mathbb{R}$ . That is, it can be used to show that all open intervals of the real number line have the same cardinality.

Furthermore, we can say that the cardinality of any open interval is the same size as the real numbers,  $|(m, n)| = |\mathbb{R}|$ . Take the interval  $(0, 1)$ . Since the real numbers between 0 and 1 can be written as sequences of digits, for example, 0.74154829750123, we can count the number of such sequences by for each of the  $\aleph_0$  decimal places, there are 10 possible digits 0 – 9. Therefore, there are  $10^{\aleph_0}$  numbers in  $(0, 1)$ . But, this is the same number as  $2^{\aleph_0}$ , which means  $10^{\aleph_0} = 2^{\aleph_0}$ . Here are some more examples of functions from  $(0, 1)$  to  $\mathbb{R}$  that are not only bijections, but also continuous mappings of  $(0, 1)$  to  $\mathbb{R}$ :

$$f : (0, 1) \rightarrow \mathbb{R} \quad f(x) = \tan\left(\pi x - \frac{\pi}{12}\right),$$

$$h : (0, 1) \rightarrow \mathbb{R} \quad h(x) = \frac{2x - 1}{x - x^2},$$

$$g : \mathbb{R} \rightarrow (0, 1) \quad g(x) = \frac{1}{1 + e^{-x}}.$$

We can even show that the number of points in the open unit square:  $(0, 1) \times (0, 1)$  is the same as the number of points in the open unit interval:  $(0, 1)$ . Let  $(x, y)$  be any point in the open unit square  $(0, 1) \times (0, 1)$ . Say  $x = 0.x_1x_2x_3x_4x_5\cdots$  and  $y = 0.y_1y_2y_3y_4y_5\cdots$ . We map this point  $(x, y)$  to a point  $z$  in a one-to-one way onto  $(0, 1)$  by  $(x, y)$  goes to  $z$  where  $z = 0.x_1y_1x_2y_2x_3y_3x_4y_4x_5y_5\cdots$ . There is a small technical adjustment that needs to be made in the situation where we have two different decimal representations for the same number, for example,  $0.53999\bar{9}$  and  $0.54000\bar{0}$ . However, this does not alter the conclusion that a bijection from  $\mathbb{R} \times \mathbb{R}$  to  $\mathbb{R}$  can be described. Generalizing this argument as before we get  $|(0, 1)| = |(0, 1) \times (0, 1)| = |\mathbb{R} \times \mathbb{R}|$ . We could extend this to points in 3-space with  $w = 0.x_1y_1z_1x_2y_2z_2x_3y_3z_3\cdots$  or even n-space. Hence,  $|(0, 1)| = |\mathbb{R}^2| = |\mathbb{R}^3| = |\mathbb{R}^n|$ , for any positive integer n.

**Remark 4.** It can be shown that the size of the set of all functions, both continuous and discontinuous, from  $\mathbb{R} \rightarrow \mathbb{R}$  is  $2^{2^{\aleph_0}}$ . The set of all *continuous* functions from  $\mathbb{R} \rightarrow \mathbb{R}$  is only of size  $2^{\aleph_0}$  because continuous functions are uniquely defined by their values at rational points. If the General Continuum Hypothesis holds (see Section 9), then the set of *all* possible functions from  $\mathbb{R}$  to  $\mathbb{R}$  is of size  $\aleph_2$ .

### 7.0.1 Cantor Set

Another odd set with an interesting cardinality is the Cantor set. We start with the closed interval  $[0, 1]$  and remove the middle third open interval  $(\frac{1}{3}, \frac{2}{3})$  so that we have two closed intervals remaining:

$$\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

We call this set  $A_1$ . Now, remove the middle third open interval of each of these two parts. That is, remove  $(\frac{1}{9}, \frac{2}{9})$  and  $(\frac{7}{9}, \frac{8}{9})$  so we are left with

$$\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right] \cup \left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

We call this set  $A_2$ . Again, remove the middle third open intervals of these intervals. So, remove  $\left(\frac{1}{27}, \frac{2}{27}\right), \left(\frac{7}{27}, \frac{8}{27}\right), \left(\frac{19}{27}, \frac{20}{27}\right)$ , and  $\left(\frac{25}{27}, \frac{26}{27}\right)$  to get

$$\left[0, \frac{1}{27}\right] \cup \left[\frac{2}{27}, \frac{3}{27}\right] \cup \left[\frac{6}{27}, \frac{7}{27}\right] \cup \left[\frac{8}{27}, \frac{9}{27}\right] \cup \left[\frac{18}{27}, \frac{19}{27}\right] \cup \left[\frac{20}{27}, \frac{21}{27}\right] \cup \left[\frac{24}{27}, \frac{25}{27}\right] \cup \left[\frac{26}{27}, 1\right].$$

We call this set  $A_3$ . Continue this process and then look at  $\bigcap_{k=1}^{\infty} A_k$  and call the elements in the intersection, the Cantor Set ( $C$ ). We observe that  $C$  is bounded since we are looking at  $[0, 1]$ . We also see that  $C$  is obviously non-empty. Finally, we have that  $C$  is closed because the intersection of any collection of closed sets is closed (Bartle & Sherbert, 2011, p. 328). We can also observe that  $C$  contains no intervals since between any two points  $x$  and  $y$  of  $C$  there is a point  $z$  such that  $z \notin C$ . Without loss of generality, let  $x < y$ . The width of the closed intervals of  $A_i$  is  $\frac{1}{3^i}$ . Let  $n = y - x$  and note that for some  $i$ ,  $n > \frac{1}{3^i}$ . This means that at some stage we delete a middle third inside the interval  $(x, y)$ . Thus,  $C$  is a closed set which consist of no intervals. In other words,  $C$  consists of all isolated points. So, how much of the interval  $[0, 1]$  did we remove? At the first stage we remove  $\frac{1}{3}$ . At the second stage we remove two sections each of size  $\frac{1}{9}$ , so  $\frac{2}{9}$ . At the third stage we remove four sections each of size  $\frac{1}{27}$ , so  $\frac{4}{27}$ . At the fourth stage we remove eight sections each of size  $\frac{1}{3^4}$ , so  $\frac{8}{3^4}$ . At the  $n^{th}$  stage we remove  $2^{n-1}$  sections each of size  $\frac{1}{3^n}$ , so  $\frac{2^{n-1}}{3^n}$ . Thus, we have

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{8}{3^4} + \dots + \frac{2^{n-1}}{3^n} + \dots = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{3} \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^{n-1} = \frac{1}{3} \left(\frac{1}{1 - \frac{2}{3}}\right) = \frac{1}{3} \left(\frac{1}{\frac{1}{3}}\right) = \frac{1}{3}(3) = 1.$$

So, the sum of the lengths that we removed is equal to 1, which is also the length of the interval  $[0, 1]$ . Therefore, the length of  $C$  must be  $1 - 1 = 0$ . This is disturbing since we have that  $C$  is an infinite set of isolated points which seems to have 0 length. More amazing results have come from the Cantor Set, such as the fact that it is an uncountable set. We will not be covering this in this thesis, but encourage the reader to explore these results.

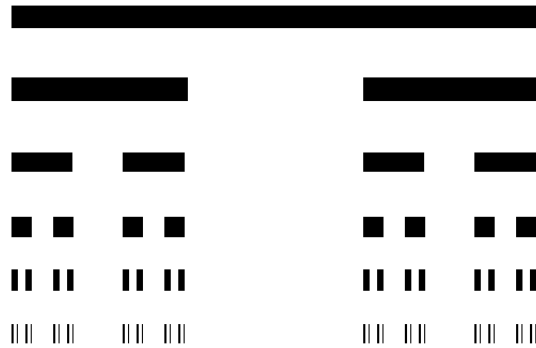


Figure 16: Visual representation of the Cantor Set.

### 7.0.2 Cantor's Theorem

**Theorem 7.1.** *Cantor's Theorem: for any set  $S$ , the set of all subsets of  $S$  has cardinality strictly greater than  $S$  itself,  $|S| < |\mathcal{P}(S)|$ .*

Here are some examples to illustrate that no matter what size a set  $S$  is, we can always prove that there is no surjection from  $S$  to  $\mathcal{P}(S)$ , thus,  $|S| < |\mathcal{P}(S)|$ . Let set  $S = \{a, b, c, d\}$ . We show some examples of functions  $f$  from  $S$  to  $\mathcal{P}(S)$ .

	example 1	example 2	example 3	example 4
$f(a) =$	$\{a, b\}$	$\{b, c\}$	$\{b\}$	$\{a, b\}$
$f(b) =$	$\{c, d\}$	$\{c\}$	$\{a, c\}$	$\{b, c, d\}$
$f(c) =$	$\{a, b, d\}$	$\{c, d\}$	$\{b\}$	$\{a, c\}$
$f(d) =$	$\emptyset$	$\{a, b, c, d\}$	$\{a, b\}$	$\{b, d\}$
$\{x : x \notin f(x)\} =$	$\{b, c, d\}$	$\{a, b\}$	$\{a, b, c, d\}$	$\emptyset$

Figure 17: Examples of functions  $f$  from  $S$  to  $\mathcal{P}(S)$ .

Note that in each case, the set  $\{x : x \notin f(x)\}$  represents the set of all elements,  $x$ , of  $S$  which are not in the image,  $f(x)$ , of themselves under the function  $f$ . That set is a subset of  $S$ , yet nothing in  $S$  gets mapped to that set. Now, we want to try and do this in general for any set  $S$ , even if  $S$  is an infinite set.

*Proof.* Suppose that  $f : S \rightarrow \mathcal{P}(S)$  and consider the set  $T = \{x \in S : x \notin f(x)\}$ . We ask if any element of  $S$  could be mapped to the set  $T$  under the function  $f$ . In other words, is there in element  $y \in S$  such that  $f(y) = T$ ? (NO). We may now build a contradiction showing the two possible cases:  $y \in T$  or  $y \notin T$ . In case one, if  $y \in T$  for  $y \in S$ , then  $y \in \{x \in S : x \notin f(x)\}$ , i.e.  $y \notin f(y)$ , so  $f(y) \neq T$ . In case two, if  $y \notin T$  for  $y \in S$ , then  $y \notin \{x \in S : x \notin f(x)\}$ , i.e.  $y \in f(y)$ , so  $f(y) \neq T$ . Thus, no element of  $S$  maps to  $T$  and there will never be any mapping from  $S$  onto the  $\mathcal{P}(S)$  for any set  $S$ . This shows that the cardinality of  $\mathcal{P}(S)$  will always be greater than the cardinality of  $S$ .  $\square$

### 7.0.3 Schröder-Bernstein Theorem

**Definition 7.1.** For two sets  $A$  and  $B$  we have that  $|A| \leq |B|$  if and only if there is an injection  $f : A \rightarrow B$ . For example, if  $A = \{a, b, c\}$  and  $B = \{a, b, c, d\}$  we have that  $|A| \leq |B|$  because we can define an injection  $f : \{a, b, c\} \rightarrow \{a, b, c, d\}$ . We will use this fact in our next theorem.

**Theorem 7.2.** *Schröder-Bernstein Theorem:* for any sets  $A$  and  $B$ , if  $|A| \leq |B|$  and  $|B| \leq |A|$ , then  $|A| = |B|$ .

*Proof.* Assume  $A$  and  $B$  are disjoint sets so  $A \cap B = \emptyset$ . If not, rename the elements of  $A$  and  $B$  so that they are disjoint. Since  $|A| \leq |B|$ , by the definition of  $\leq$ , there is a one-to-one function mapping  $A$  into  $B$  (not necessarily onto  $B$ ). Let  $f$  be the function that maps from  $A$  to  $B$  in a one-to-one way. Since  $|B| \leq |A|$ , by the definition of  $\leq$ , there exists a one-to-one function mapping  $B$  into  $A$  (not necessarily onto  $A$ ). Let  $g$  be the function that maps from  $B$  to  $A$  in a one-to-one way. Consider the elements of  $A$  as blue dots (that is,  $A$  is the collection of blue dots) and the elements of  $B$  as red dots (that is,  $B$  is the collection of red dots). View the one-to-one function  $f$  as the set of blue arcs (edges) from the blue dots (vertices) to the red dots (vertices). View the one-to-one function  $g$  as the set of red arcs (edges) from the set of red dots (vertices) to the blue dots (vertices).

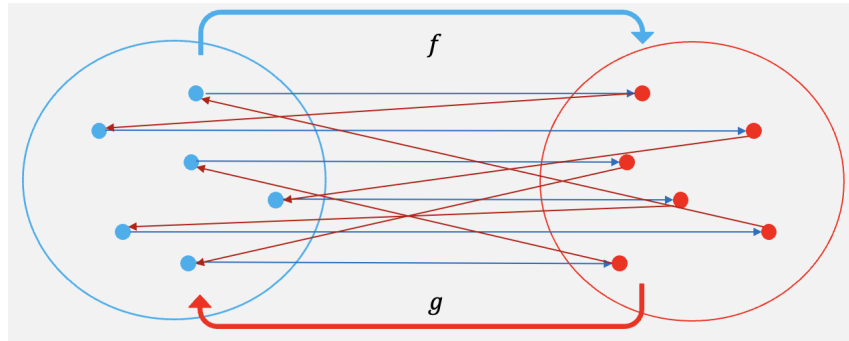
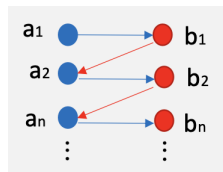


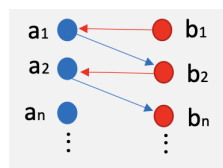
Figure 18: Diagram of the one-to-one functions  $f$  and  $g$ .

The graph of these blue and red edges consists of a collection of disjoint connected components. Each vertex of  $A \cup B$  has exactly one edge going out and at most one edge coming in. Each connected component is one of the following four types:

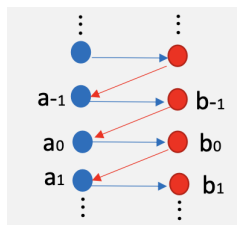
1. A singly infinite path starting at a vertex of  $A$ :  $a_1, b_1, a_2, b_2, a_3, b_3, \dots$



2. A singly infinite path starting at a vertex of  $B$ :  $b_1, a_1, b_2, a_2, b_3, a_3, \dots$

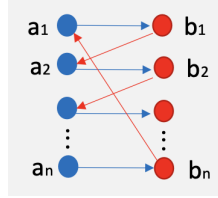


3. A doubly infinite path:  $\dots, a_{-1}, b_{-1}, a_0, b_0, a_1, b_1, \dots$





4. A finite directed cycle:  $a_1, b_1, a_2, b_2, a_4, b_4, \dots, a_n, b_n, a_1$



To show that  $|A| = |B|$ , we need to display a one-to-one function from  $A$  onto  $B$ . Let  $\varphi$  be the function  $\varphi : A \rightarrow B$  defined by

1. If  $a_i$  is a vertex (element of  $A$ ) which belongs to a connected component of type 1, 3, or 4, then  $\varphi(a_i) = f(a_i) = b_i$ . (that is, its successor in the sequence)
2. If  $a_i$  is a vertex (element of  $A$ ) which belongs to a connected component of type 2, then  $\varphi(a_i) = g^{-1}(a_i)$ . (that is, the immediate predecessor in the sequence)

Since  $f$  is a function on  $A$ , every element of  $A$  gets mapped to some element of  $B$ , by  $\varphi$ . Suppose that  $\varphi(a_i) = \varphi(a_k)$ , then  $a_i$  and  $a_k$  must be in the same component. If both  $a_i$  and  $a_k$  are in a component of type 1, 3, or 4, then  $\varphi$  is one-to-one because  $f$  is one-to-one. If both  $a_i$  and  $a_k$  are in a component of type 2, then  $\varphi$  is one-to-one because  $g$  is one-to-one.

Now, we will show that  $A$  is onto  $B$ . Consider any element,  $b_j$ , of  $B$ . Since  $g$  is a function on  $B$ , there must be a red edge out of  $b_j$  to an element  $a$  in  $A$  and a blue edge of out of  $a$ . Thus,  $b_j$  is a part of some component. If  $b_j$  is in a component of type 1, 3, or 4 then  $\varphi(a_j) = b_j$ . This means  $a_j$  maps onto  $b_j$ . If  $b_j$  is in a component of type 2, then  $\varphi(a_j) = g^{-1}(a_j) = b_j$ . This means  $a_j$  maps onto  $b_j$ . In any case, for every  $b$  in  $B$  some element of  $A$  maps onto  $b$ , thus  $\varphi$  is onto  $B$ . Thus, given a one-to-one function from  $A$  to  $B$  and a one-to-one function from  $B$  to  $A$ , we can find a one-to-one function  $\varphi$  from  $A$  onto  $B$ . Therefore,  $|A| = |B|$ .

□

## 7.1 Cardinal Arithmetic

### 7.1.1 Addition

We start with a definition of cardinal addition for finite sets:

**Definition 7.2.** *Cardinal Addition:* Let  $\mu$  and  $\nu$  be two arbitrary cardinal numbers. The sum of  $\mu$  and  $\nu$ , denoted  $\mu + \nu$ , is the cardinal number of  $B \cup C$  where  $B$  and  $C$  are any disjoint sets having cardinals  $\mu$  and  $\nu$  respectively. So, let  $\mu = |B|$  and  $\nu = |C|$  with  $B \cap C = \emptyset$ . Then,  $\mu + \nu = |(B \cup C)|$ .

It is true that finite cardinal addition is commutative.

**Theorem 7.3.** *Let  $\mu$  and  $\nu$  be arbitrary cardinal numbers. Then,  $\mu + \nu = \nu + \mu$*

*Proof.* Let  $B$  and  $C$  be disjoint sets with  $\mu = |B|$  and  $\nu = |C|$ . Then,  $\mu + \nu = |B \cup C|$ . But,  $B \cup C = C \cup B$  so that  $\nu + \mu = |C \cup B| = |B \cup C| = \mu + \nu$ .  $\square$

In an effort to generalize the concept of a number, certain properties of numbers are lost. For example, if  $n$  is a finite cardinal, it is true that  $n + 1 > n$ . Although, in the transfinite case, it is not true that  $\aleph_0 + 1 > \aleph_0$ , but  $\aleph_0 + 1 = \aleph_0$ .

We can also define cardinal addition in terms of indexing sets.

**Definition 7.3.** *Cardinal Addition (index sets):* Let  $T$  be an index set and  $C_t$  a set for each  $t \in T$  such that  $\gamma_t = |C_t|$  and  $C_t \cap C_r = \emptyset$  if  $t \neq r$ . Then, the sum of the cardinal numbers  $\gamma_t$ , denoted

$$\sum_{t \in T} \gamma_t,$$

is defined by

$$\sum_{t \in T} \gamma_t = \left| \bigcup_{t \in T} C_t \right|.$$

**Example 7.1.** Suppose  $A = \{a, b, c\}$  and  $B = \{d, e, f, g, h, i, j, k\}$ . Then,  $A \cup B = \{a, b, c, d, e, f, g, h, i, j, k\}$ . So,  $|(A \cup B)| = |A| + |B| = 3 + 8 = 11$ .

### 7.1.2 Multiplication

**Definition 7.4.** *Cardinal Multiplication:* Let  $\mu$  and  $\nu$  be arbitrary cardinal numbers with sets  $A$  and  $B$  such that  $\mu = |A|$  and  $\nu = |B|$ . Then, the product of  $\mu$  and  $\nu$ , denoted  $\mu \cdot \nu$  is defined by  $\mu \cdot \nu = |A \times B|$ .

Again, in an effort to generalize the concept of number, we lose certain properties of those numbers. For example, for finite cardinals the equation  $n \cdot x = n$  where  $n \neq 0$  holds only for  $x = 1$ . But, for transfinite cardinals, this is not true. We have that  $\aleph_0 \cdot 1 = \aleph_0$ , and  $\aleph_0 \cdot \aleph_0 = \aleph_0$ ,  $\dots$ ,  $\aleph_0 \cdot n = \aleph_0$ . Thus, every time  $n \neq 0$ , we have that  $\aleph_0 \cdot n = \aleph_0$ . So, our uniqueness of multiplication by 1 is lost.

**Theorem 7.4.**  $\mu \cdot \nu = 0$  if and only if  $\mu = 0$  or  $\nu = 0$  where  $\mu$  and  $\nu$  are cardinal numbers.

*Proof.* If either  $\mu = 0$  or  $\nu = 0$ , then  $\mu \cdot \nu = 0$ . Suppose  $\mu \neq 0$  and  $\nu \neq 0$  and let  $\mu = |A|$  and  $\nu = |B|$ . Then  $A \neq \emptyset$  and  $B \neq \emptyset$  so that  $A \times B \neq \emptyset$ , and  $\mu \cdot \nu = |A \times B| \neq 0$ .  $\square$

**Example 7.2.** Suppose  $D = \{2, 4, 6, 8\}$  and  $E = \{1, 3, 5\}$ . Then,  $D \times E = \{(2, 1), (2, 3), (2, 5), (4, 1), (4, 3), (4, 5), (6, 1), (6, 3), (6, 5), (8, 1), (8, 3), (8, 5)\}$ . So,  $|D \times E| = |D| \cdot |E| = 4 \cdot 3 = 12$ .

### 7.1.3 Exponentiation

**Definition 7.5.** *Cardinal Exponentiation:* Let  $\mu$  and  $\nu$  be cardinal numbers. If  $\mu = |A|$  and  $\nu = |B|$ , then  $\mu^\nu = |B \rightarrow A|$ , where  $B \rightarrow A$  is the set of all functions from  $B$  to  $A$ .

**Example 7.3.** Suppose we have  $2^3$  with  $|A| = 2$  and  $|B| = 3$ . Then, we see if  $B = \{a, b, c\}$  and  $A = \{0, 1\}$ , there are exactly eight functions that can be defined from  $B$  to  $A$ . The eight functions are:

$$f = \{(a, 0), (b, 0), (c, 0)\},$$

$$g = \{(a, 1), (b, 1), (c, 1)\},$$

$$h = \{(a, 1), (b, 1), (c, 0)\},$$

$$i = \{(a, 1), (b, 0), (c, 0)\},$$

$$j = \{(a, 0), (b, 1), (c, 0)\},$$

$$k = \{(a, 0), (b, 1), (c, 1)\},$$

$$l = \{(a, 0), (b, 0), (c, 1)\},$$

$$m = \{(a, 1), (b, 0), (c, 1)\}.$$

If we let  $\mathcal{F}$  denote the set of functions from  $B$  to  $A$ , then we see that  $|\mathcal{F}| = 8$ , which is precisely what  $2^3$  equals.

**Example 7.4.** Suppose we have  $3^2$  with  $|A| = 3$  and  $|B| = 2$ . Then, we see if  $B = \{0, 1\}$  and  $A = \{d, e, f\}$ , there are exactly nine functions that can be defined from  $B$  to  $A$ . The nine functions are:

$$f = \{(0, f), (1, f)\},$$

$$g = \{(0, f), (1, e)\},$$

$$h = \{(0, f), (1, d)\},$$

$$i = \{(0, e), (1, e)\},$$

$$j = \{(0, e), (1, f)\},$$

$$k = \{(0, e), (1, d)\},$$

$$l = \{(0, d), (1, d)\},$$

$$m = \{(0, d), (1, e)\},$$

$$n = \{(0, d), (1, f)\}.$$

If we let  $\mathcal{H}$  denote the set of functions from  $B$  to  $A$ , then we see that  $|\mathcal{H}| = 9$ , which is precisely what  $3^2$  equals. Notice that  $2^3 \neq 3^2$  because they do not have the same number of functions from  $B$  to  $A$ . So, cardinal exponentiation is not commutative.

**Remark 5.** Properties of cardinal exponentiation:

1.  $\mu^1 = \mu$  for every cardinal  $\mu$ .
2.  $\mu^0 = 1$  for every cardinal  $\mu$ .
3. If  $\mu \neq 0$ ,  $0^\mu = 0$ .
4.  $\mu^{\nu+\rho} = \mu^\nu \cdot \mu^\rho$  for arbitrary cardinals  $\mu, \nu, \rho$ .
5.  $(\mu \cdot \nu)^\rho = \mu^\rho \cdot \nu^\rho$  for arbitrary cardinals  $\mu, \nu, \rho$ .
6.  $(\mu^\nu)^\rho = \mu^{\nu \cdot \rho}$  for arbitrary cardinals  $\mu, \nu, \rho$ .

## 8 Axiom of Choice

In the early 1900's, one axiom was carefully identified and later shown not to be provable from the other well-known axioms in Zermelo-Fraenkel Set Theory (ZF). Kurt Gödel in the 1930's and Paul Cohen in the 1960's proved that this new axiom, the Axiom of Choice (AC), was independent of the Zermelo-Fraenkel axiom scheme.

The Axiom of Choice asserts that, given a family of sets,  $T$ , a new set exists which consists of one element from each set in the family  $T$ . There are steps in certain proofs that cannot be justified using only the axioms of Zermelo-Fraenkel Set Theory. The Axiom of Choice allows us to create new sets that could not be created without its use. There are a variety of ways to describe the Axiom of Choice. Here are a few of the ways in which we can describe it that are all equivalent:

**Axiom 8.1.** *Axiom of Choice:* Given any family  $T$  of non-empty sets, there is a function  $f$  which assigns to each member  $A$  of  $T$  an element  $f(A)$  of  $A$ . Such an  $f$  is now called a choice function for  $T$ .

**Axiom 8.2.** Every family of nonempty sets has a choice function.

**Axiom 8.3.** The Cartesian product of a non-empty family of non-empty sets is non-empty.

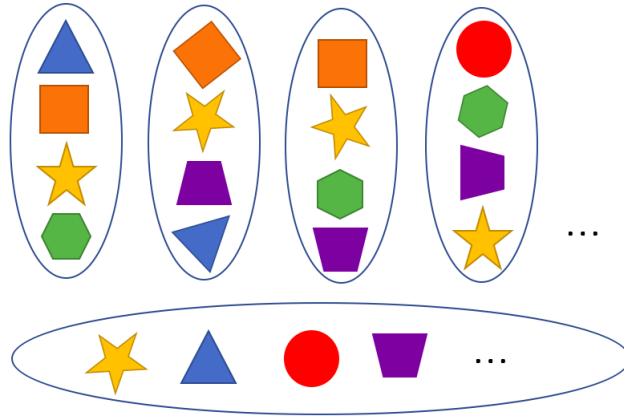


Figure 19: Informal diagram of the Axiom of Choice.

We can also describe the Axiom of Choice more formally in terms of the Language of Set Theory (LAST):

**Definition 8.1.**

$$\forall T [ \emptyset \notin T \rightarrow \exists f [ \text{dom}(f) = T \wedge \forall y \in T ( f(y) \in y ) ] ]$$

What this definition is saying is that for every family  $T$  of non-empty sets, there exists a function  $f$  such that the domain of  $f$  is the family  $T$ , and for every set  $y$  in  $T$ ,  $f(y)$  is in  $y$ . So, the new set is formed by taking an element from each set  $y$  in the family  $T$  using the choice function  $f$ .

However, in certain special circumstances, we can prove the existence of this new set without invoking AC. We will describe some of these special circumstances where AC is not necessary in Section 8.1.

We will now provide some important definitions to help in our understanding of AC's different variations.

**Definition 8.2.** *Finite:* A set  $T$  is *finite* if  $T$  is empty or if there is a bijection from  $T$  onto the set  $\{1, 2, 3, 4, \dots, n\}$  for some positive natural number  $n$ . The sets  $\{1\}, \{1, 2\}, \{1, 2, 3\}, \dots$ ,

$\{1, 2, 3, 4, \dots, n\}$  for positive natural numbers  $n$ , are called the *initial segments* of the positive natural numbers. Thus, a set  $T$  is *finite* if  $T$  is empty or there is a bijection from  $T$  onto an initial segment of the positive natural numbers.

**Definition 8.3.** *Infinite:* A set  $T$  is *infinite* if it is not finite. This means that there is no bijection from  $T$  onto the set  $\{1, 2, 3, \dots, n\}$  for any positive natural number  $n$ . Thus, there is no bijection from  $T$  onto an initial segment of the positive natural numbers.

**Definition 8.4.** *Denumerable:* A set  $T$  is *denumerable* if there is a bijection from the positive natural numbers  $\mathbb{N}$  to  $T$ . This means that there exists a one-to-one and onto function from  $\mathbb{N}$  to  $T$ . Such a set is also called countably infinite.

**Definition 8.5.** *Countable:* A set  $T$  is *countable* if  $T$  is a finite set or  $T$  is a denumerable set. This suggests that the elements can be systematically listed or enumerated. Of course, in the infinite case, it is not possible to list all the elements, but it should be possible to describe a bijection from  $\mathbb{N}$  that would systematically list the elements in a way to show that eventually every element would appear on the list.

**Definition 8.6.** *Uncountable:* An infinite set that contains too many elements to be listed or enumerated is considered *uncountable*. This means we are unable to systematically list the elements in the set. Thus, the set  $T$  is *uncountable* if there is no mapping from the positive natural numbers onto  $T$ .

**Definition 8.7.** *Dedekind-Infinite:* A set  $A$  is *Dedekind-infinite* (DI) if there exists a bijection from  $A$  to a proper subset  $B$  of  $A$ .

**Example 8.1.** Given the sets  $A$  and  $B$ , let  $A = \mathbb{N}$  and  $B = 2\mathbb{N}$  (even natural numbers). Then  $A$  is *Dedekind-infinite* because there exists a bijection from  $\mathbb{N}$  to  $2\mathbb{N}$ . That is, we can map  $1 \rightarrow 2$ ,  $2 \rightarrow 4$ ,  $3 \rightarrow 6$ ,  $\dots$  in a one-to-one and onto way.

**Definition 8.8.** *Dedekind-Finite:* A set  $A$  is *Dedekind-finite* (DF) if there is no bijection from  $A$  to any proper subset  $B$  of  $A$ .

**Example 8.2.** Given the sets  $C$  and  $D$ , let  $C = \{1, 2, 3, 4, 5\}$  and  $D = \{1, 3, 4\}$ . Then  $C$  is *Dedekind-finite* because there does not exist a bijection from  $C$  to  $D$ . That is,  $1 \rightarrow 1$ ,

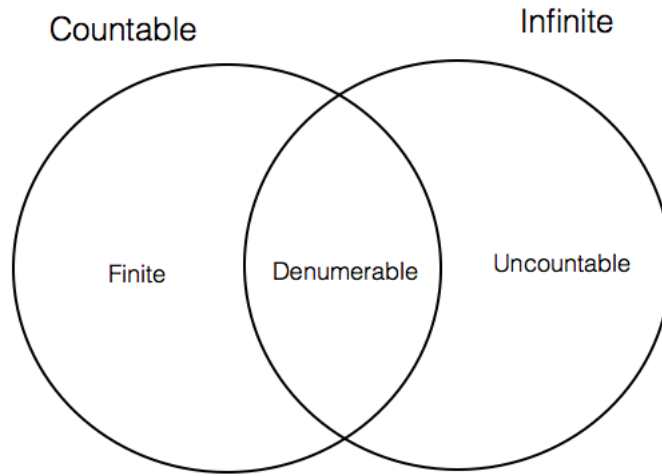


Figure 20: Venn diagram representation of above definitions.

$2 \rightarrow 3, 3 \rightarrow 4, 4 \rightarrow 1, 5 \rightarrow 3$ . Although a mapping from  $C$  to  $D$  may be onto  $D$ , no mapping from  $C$  to  $D$  can be one-to-one. Thus, it is not a bijection.

### 8.1 Versions We Do Not Need

There are two concerns about the sizes of objects:

1. **Size of Family:** The size of the family  $T$  itself (finite or infinite).
2. **Size of Sets:** The size of each of the sets in the family  $T$  (finite or infinite).

We are concerned about the sizes of objects (family or sets) because this is what tells us if we need to use the Axiom of Choice or not. If we do need to use it, the sizes of the sets in the family tell us what version of the Axiom to use. We remember that AC does not require that we have or need to find a specific function, it merely states that one exists. Thus, if a choice function is possible without invoking AC, then we do not need AC.

Here is an example to illustrate one concern about the sizes of objects, specifically, the **size of the family**  $T$  of sets:



Suppose that the size of our family of sets  $T$  is finite. This means the number of sets in the family  $T$  is finite. It does not matter how many elements are in each individual set within the family  $T$ . If the size of our family of sets  $T$  is finite, then a choice function exists without invoking AC. This is called *finite choice*. Thus, if the number of sets in our family,  $T$ , of sets is *finite*, we can, in ZF set theory, construct a new set consisting of an element from each of the finitely many sets in  $T$  without invoking AC.

**Example 8.3.** Suppose the set  $T = \{\{a, b, c\}, \{1, 2, 3\}, \{\square, \star, \heartsuit, \Delta\}\}$ . There is no need to assert the existence of a set consisting of one element from these finitely many sets because we can do this in ZF set theory.

We can show in finitely many statements of the Language of Set Theory (LAST), using the axioms of ZF, that not only there exists a new set consisting of one element from each of the sets in  $T$  but, we can, in this case, actually construct such a set in finitely many steps. Hence, there is no need for AC when the number of sets in the family  $T$  of sets is finite. Therefore, we will only consider cases when the family  $T$  of sets is infinite. Having established that the family  $T$  of sets is infinite, we now consider the sizes of the individual sets in this family.

Here are some examples to illustrate the concerns about the sizes of objects, specifically, the **size of the sets** within the family:

**Example 8.4.** If each set in the family consists of just one element  $\{v\}, \{w\}, \{y\}, \dots$  then one choice function is  $f(\{x\}) = x$  for every set in  $T$ . Thus, since there is already a choice function defined, so we do not need to use AC.

**Example 8.5.** Suppose each set  $A$  in the family  $T$  consists of two real numbers  $\{n, m\}$ . Let  $f(A) = \text{minimum of } \{n, m\}$ . Because we have a linear ordering of the real numbers, then the function,  $f(A) = \text{minimum of } \{n, m\}$ , allows us to create the new set without using AC. Since there is already a way (the linear ordering of  $\mathbb{R}$ ) to make choices, we do not need to apply AC.

**Example 8.6.** Using set theory, we define ordered pairs by  $(a, b) = \{\{a\}, \{a, b\}\}$ . Thus, if our sets are ordered pairs or ordered triples or, in general, ordered n-tuples, then again, a choice function can be defined without needing to apply AC.

“It is common enough, in mathematics, to make one arbitrary choice (we do this every time we say ‘let  $x$  be an arbitrary element of  $A$ ’), and experience confirms that we can make a finite succession of choices; but to make an infinite succession of choices is to carry an argument through an infinite number of steps – and nothing in our experience or in the logic we habitually use justifies such a process” (Pinter, 1971, 2014). Thus, we may need AC when we have an infinite family of sets. However, we do not always require the full version of AC to do the job.

## 8.2 Weaker Versions

### 8.2.1 Denumerable Choice

Often times, we do not need to use the full Axiom of Choice. We can use a weaker variant instead, Denumerable Choice or DAC.

**Definition 8.9.** *Denumerable Choice:* Every countable family of nonempty sets has a choice function (Jech, 1973, p. 20).

Jech refers to this as the Countable Axiom of Choice, but for the sake of consistency with the other texts we refer to, we shall call it the Denumerable Axiom of Choice or DAC for short. This version is a weaker form of AC because it does not include every possible family of sets, only the countable ones.

**Remark 6.** It is easy to see that  $AC \rightarrow DAC$ . If we have the full AC for any infinite family of sets, then we certainly have it for a countably infinite family of sets (denumerable sets). By  $AC \rightarrow DAC$ , we mean that in ZF Set Theory with AC, we can prove the DAC. Conversely, it turns out that  $DAC \not\rightarrow AC$ . If we assume DAC, which says for every countable family of sets there’s a choice function, it does not give us a choice function for every possible

family of sets, in particular, not for uncountable families of sets. By  $\text{DAC} \not\rightarrow \text{AC}$ , we mean that in ZF Set Theory with DAC, it is not possible to prove the full AC.

### 8.2.2 Consequences of DAC

Although DAC is a weaker version of AC, it still gives many remarkable consequences:

**Consequence 8.0.1.** Every infinite set has a denumerable subset.

**Remark 7.** Although it seems obvious that every infinite set has a denumerable subset, it is not possible to prove this inside of ZF. Thus, we require a version of AC to prove this consequence.

*Proof.* Let  $T$  be an infinite set. For every natural number  $n$ , let  $F_n$  be the set of all one-to-one functions from  $\{1, 2, 3, 4, \dots, n\} \rightarrow T$ . Since  $T$  is infinite, for each  $n$ ,  $F_n$  is non-empty. Let  $R_n$  be the set of ranges of  $F_n$ . By DAC, there exists a choice function,  $f$ , on the set of ranges  $\{R_n\} : n \in \{1, 2, 3, \dots\}$ . Thus, choosing an element from each set in the collection, we have that  $f(R_n) = r_n$ . Thus we obtain a set  $\{r_1, r_2, r_3, \dots\}$ . It is technically messy but not theoretically difficult to delete repetitions of the elements chosen from  $R_n$  to ensure that they are distinct and thus the set  $\{r_1, r_2, r_3, \dots\}$  is denumerable.  $\square$

**Example 8.7.** Let  $T = \{\{a\}, \{a, c\}, \{b\}, \{a, c, d\}, \{m, n, p\}, \{g\}, \{r, s\}, \dots\}$ .

Informally define the one-to-one functions from the set containing 1 to  $T$  by:

$$F_1 = \{1 - \{a, c\}, 1 - \{b\}, 1 - \{a\}, 1 - \{m, n, p\}, 1 - \{r, s\}, \dots\}.$$

The notation  $1 - \{a, c\}$  means this is a function from 1 to  $\{a, c\}$ ; the notation  $1 - \{b\}$  means this is a function from 1 to  $\{b\}$ ; etc.

So,  $F_1$  is all the functions from  $\{1\}$  into  $T$ . The ranges of these functions are  $\{\{a, c\}, \{b\}, \{a\}, \{m, n, p\}, \{r, s\}, \dots\}$ .

Informally define the one-to-one functions from the set containing 1, 2 to  $T$  by:

$$F_2 = \{1 - \{a\} \ 2 - \{a, c, d\}, 1 - \{r, s\} \ 2 - \{b\}, 1 - \{a\} \ 2 - \{g\}, \dots\}.$$

So,  $F_2$  is all the functions from  $\{1, 2\}$  into  $T$ . The ranges of these functions are  $\{\{a\}, \{a, c, d\}, \{r, s\}, \{b\}, \{a\}, \{g\}, \dots\}$ .

Informally define the one-to-one functions from the set containing 1, 2, 3 to  $T$  by:

$$F_3 = \{1 - \{g\} \ 2 - \{a, c, d\} \ 3 - \{a\}, 1 - \{r, s\} \ 2 - \{b\} \ 3 - \{a, c\}, \dots\}.$$

So,  $F_3$  is all the functions from  $\{1, 2, 3\}$  into  $T$ . The ranges of these functions are  $\{\{g\}, \{a, c, d\}, \{a\}, \{r, s\}, \{b\}, \{a, c\}, \dots\}$ .

Now,  $f$  is our choice function, so  $f$  asserts the ability to somehow select one of the ranges in  $F_1, F_2, F_3, \dots$  for every  $n$ . For example, suppose  $f$  chose the range of the function  $1 - \{a, c\}$  which is the set  $\{\{a, c\}\}$ . Now suppose  $f$  chooses the range of  $1 - \{r, s\} \ 2 - \{b\}$ , which is the set  $\{\{r, s\}, \{b\}\}$ . We can then write the sequence  $\{a, c\}, \{r, s\}, \{b\}$ . Now suppose  $f$  chooses the range of the function  $1 - \{b\} \ 2 - \{a, c\} \ 3 - \{r, s\}$ , which is the set  $\{\{b\}, \{a, c\}, \{r, s\}\}$ . We can add these to our previous sequence  $\{a, c\}, \{r, s\}, \{b\}$  to obtain  $\{a, c\}, \{r, s\}, \{b\}, \{b\}, \{a, c\}, \{r, s\}$ . Crossing out any repetitions gives us the sequence  $\{a, c\}, \{r, s\}, \{b\}$ . Even though we did not add any new elements at this stage, eventually and continually, some new range will contain at least one more new element because the functions are one-to-one, so the sizes of the ranges are constantly increasing. So, the list will continue to grow. Now, suppose  $f$  chooses the range of the function  $1 - \{g\} \ 2 - \{a, c, d\} \ 3 - \{a\} \ 4 - \{m, n, p\}$  from  $F_4$ , which is the set  $\{\{g\}, \{a, c, d\}, \{a\}, \{m, n, p\}\}$ . Adding this to our sequence  $\{a, c\}, \{r, s\}, \{b\}$  gives  $\{a, c\}, \{r, s\}, \{b\}, \{g\}, \{a, c, d\}, \{a\}, \{m, n, p\}$ . Crossing out any repetitions gives us the sequence  $\{a, c\}, \{r, s\}, \{b\}, \{g\}, \{a, c, d\}, \{a\}, \{m, n, p\}$

(there were no repetitions this time, so the length of the sequence increased by four). Thus, DAC posits the existence of a denumerable subset from *any* infinite set.

Looking back on our definitions of Dedekind-infinite and Dedekind-finite, we can now give some other implications relating to DAC. Given DAC, the assertion that a set is finite (infinite) if and only if it is Dedekind-finite (Dedekind-infinite) is logically equivalent to the statement that every infinite set has a denumerable subset (consequence 3.0.1). However, the converse is not true. In other words, the statements: a set is finite (infinite) if and only if it is Dedekind-finite (Dedekind-infinite) and every infinite set has a denumerable subset being equivalent, does not imply DAC. We show these implications symbolically below:

$$DAC \rightarrow [(F \leftrightarrow DF) \equiv \text{consequence 8.0.1}]$$

however

$$[(F \leftrightarrow DF) \equiv \text{consequence 8.0.1}] \nrightarrow DAC.$$

**Consequence 8.0.2.** The union of a countable family of countable sets is countable.

*Proof.* Let  $T$  be a countable collection of countable sets,  $T = \{A_a, A_b, A_o, A_p, \dots\}$ . We will consider the denumerable case; if any sets in the family are finite, then the proof is much easier. We will assume without loss of generality, that each set is denumerable. Thus, for each set in the collection  $T$ , there exists infinitely many bijections from the natural numbers, but at least one. For example, we consider the set  $A_a = \{a_\Delta, a_\beta, a_\alpha, a_g, a_q, \dots\}$ . One of the bijections from the natural numbers onto  $A_a$  is

$$\begin{array}{r}
1 - a_\beta \\
2 - a_\Delta \\
3 - a_\alpha \\
4 - a_q \\
5 - a_g \\
\vdots \quad \vdots
\end{array} \tag{1}$$

Let  $B$  be the collection of all such bijections. Thus,  $B = \{B_a, B_b, B_o, B_p, \dots\}$ . Applying AC to this set  $B$ , we can *choose* a bijection from each set in  $B$  to construct a matrix. For example, if (1) is the bijection chosen from  $B_a$ , we can rename  $a_\beta$  as  $a_{11}$ ,  $a_\Delta$  as  $a_{12}$ , and so on. Therefore, we can construct the matrix:

$$\begin{array}{cccccc}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & \cdots \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & \cdots \\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & \cdots \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & \cdots \\
a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array} \tag{2}$$

Going down each counter-diagonal starting from top to bottom left, we are able to create a

bijection from the natural numbers to obtain:

$$\begin{aligned}
 &1 - a_{11} \\
 &2 - a_{12} \\
 &3 - a_{21} \\
 &4 - a_{13} \\
 &5 - a_{22} \\
 &6 - a_{31} \\
 &7 - a_{14} \\
 &8 - a_{23} \\
 &\vdots \quad \vdots
 \end{aligned} \tag{3}$$

We have that  $a_{jk}$  = the summation

$$\sum_{i=1}^{j+k-2} i + j. \tag{4}$$

Given any  $a_{jk}$  in the matrix, we are able to tell exactly what natural number it corresponds to in the bijection list. For example,  $a_{14}$  can be found by  $\sum_{i=1}^{(1+4)-2} i + 1 = \sum_{i=1}^3 i + 1 = 7$ . So, 7 maps to  $a_{14}$  in the matrix. Thus, we have demonstrated a bijection from  $\mathbb{N}$  to  $a_{jk}$ . Obviously, we can do this for finite sets. Therefore, we have shown the union of the countable sets in our countable collection  $T$  is countable, which is exactly what we wanted to prove.  $\square$

**Remark 8.** This proof is a generalization of Cantor’s diagonal argument showing the denumerability of the rational numbers.

**Remark 9.** In order for this summation to be considered one-to-one, we must remove any repetitions from the listing of  $T$ . This issue is similar to the problem in Cantor’s diagonal argument of the rational numbers (for example,  $\frac{2}{3} = \frac{4}{6}$ ).

It is somewhat ironic that the problem with showing this function is one-to-one, is that, in some sense, there are “too many” natural numbers (not too many elements in the two

dimensional array of elements). This is somewhat counter-intuitive because this is the opposite of what one might suspect. That is, we might guess that the number of the elements in a countably infinite collection of countable infinitely many objects would be “more” than the number of elements in just a single countable infinite list (the natural numbers) and if there were a problem in terms of one-to-one it would be because more than one element in the array was matched with the same natural number (not the other way). In any case, it is easy to fix this so that the function is, in fact, one-to-one and onto.

**Remark 10.** It turns out, there is a model of Zermelo-Fraenkel Set Theory (ZF) together with the negation of DAC, in which the set of all real numbers, although uncountable, is a countable union of countable sets (Moore, 1982, p. 9). This illustrates some of the friction that occurs if we do not accept AC.

**Consequence 8.0.3.** The set of all real numbers is not the union of a countable collection of countable sets (not a countable union of countable sets). This follows from Consequence 8.0.2 and  $\mathbb{R}$  being uncountable.

**Consequence 8.0.4.** Standard metric topological properties of the real line.

We have two kinds of definitions of some standard metric topological properties: analytic and sequential. In each of these properties below, the (i) analytic  $\varepsilon - \delta$  definitions are equivalent to the (ii) sequential definitions under the Denumerable Axiom of Choice.

1. **Closure:** of a set  $T$ .

- (i) A point  $x$  is in the closure of a set  $T$  if for every  $\varepsilon > 0$ , the  $\varepsilon$ -neighborhood of  $x$  intersects  $T$ . That is, no matter how small a ball we form around  $x$  (where  $\varepsilon$  is the radius of the ball) there is a point from  $T$  that is inside the ball. In other words, for every  $\varepsilon > 0$ , the distance between  $x$  and some point  $b \in T$  is  $< \varepsilon$  or, for every positive  $\varepsilon$ , there is a point  $b \in T$  such that  $|x - b| < \varepsilon$ .



(ii) A point  $x$  is in the closure of a set  $T$  if for some sequence  $x_n$  of points in  $T$ ,  

$$\lim_{n \rightarrow \infty} x_n = x.$$

*Proof. (ii)  $\rightarrow$  (i):* Let  $T$  be a set and let  $x$  be a point in the closure of  $T$ . Show that if  $x$  is in the closure by definition (ii), then  $x$  is in the closure by definition (i). We have the closure of  $T$  according to definition (ii). We must prove that this also satisfies definition (i). So, for any point  $x$  in the closure of  $T$ , there is a sequence of points  $x_n$  in  $T$ , such that  $\lim_{n \rightarrow \infty} x_n = x$ . We need to show, for every  $\varepsilon > 0$ , the  $\varepsilon$ -neighborhood of  $x$  intersects  $T$ . So, let  $\varepsilon > 0$ . By contradiction, if there exists an  $\varepsilon$ -neighborhood of  $x$  that does not intersect  $T$ , no sequence converges to  $x$ .  $\square$

*Proof. (i)  $\rightarrow$  (ii):* Now, we must show if  $x$  is in the closure by definition (i), then  $x$  is in the closure by definition (ii). So, we assume for every  $\varepsilon > 0$ , the  $\varepsilon$ -neighborhood of  $x$  intersects  $T$ . We need to show that for any point  $x$  in the closure of  $T$  there is a sequence of points  $x_n$  in  $T$ , such that  $\lim_{n \rightarrow \infty} x_n = x$ . Here is where we will need AC.  $\forall n \in \{1, 2, 3, 4, \dots, n, \dots\}$  let  $\varepsilon = \frac{1}{n}$ . Each of these  $\varepsilon$ -neighborhoods intersects  $T$  by definition (i), so each  $\varepsilon$ -neighborhood is non-empty. Let  $s_n$  be the set of points of  $T$  in the  $\varepsilon_n = \frac{1}{n}$ -neighborhood of  $x$ . This gives us a family of sets  $\{s_n : n = 1, 2, 3, 4, \dots, n, \dots\}$ . Using DAC, we select one element,  $x_n$ , out of each of these sets to form a sequence  $x_1, x_2, x_3, x_4, \dots$  which converges to  $x$ . Hence, there exists a sequence  $x_n$  of points in  $T$ , such that  $\lim_{n \rightarrow \infty} x_n = x$ .

$\square$

So, we see that using denumerable choice, definition (ii) is equivalent to definition (i) for closed sets.

## 2. **Continuous:** at a point $c$ .

(i)  $f$  is continuous at a point  $c$  if for every  $\varepsilon > 0$ , there exists a  $\delta > 0$  such that when  
 $|x - c| < \delta$  then  $|f(x) - f(c)| < \varepsilon$ .

(ii)  $f$  is continuous at a point  $c$  if  $\lim_{n \rightarrow \infty} (x_n) = c$  then  $\lim_{n \rightarrow \infty} (f(x_n)) = f(c)$ .

*Proof.* (ii)  $\rightarrow$  (i)

We assume that  $f$  is continuous at a point  $c$  by definition (ii). We must prove that  $f$  is continuous at a point  $c$  by definition (i). We assume definition (ii). We let  $\varepsilon > 0$  and need to show that there exists a  $\delta > 0$ . By definition (ii), we have that  $\lim_{n \rightarrow \infty} (f(x_n)) = f(c)$ . Thus, for some  $m$ ,  $|f(x_m) - f(c)| < \varepsilon$ . Otherwise, we do not have sequential convergence.  $\square$

*Proof.* (i)  $\rightarrow$  (ii)

Let  $\varepsilon = \frac{1}{n}$  for  $n = \{1, 2, 3, \dots\}$ . This gives us the sequence  $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$ . By definition (i), for each time we pick a  $\varepsilon > 0$ , there exists a  $\delta$  such that whenever  $|x - c| < \delta$  then  $|f(x) - f(c)| < \varepsilon$ . Thus, as we sequentially choose  $x_n$ 's closer and closer to  $c$  (using DAC), the  $f(x_n)$ 's become closer and closer to  $f(c)$ . Hence, DAC tells us there exists a sequence  $x_n$  such that  $\lim_{n \rightarrow \infty} x_n = c$  implies  $\lim_{n \rightarrow \infty} f(x_n) = f(c)$ . Therefore, we have shown definition (i)  $\rightarrow$  (ii).  $\square$

### 3. Compactness: of a set $A$ .

- (i) A set  $A$  is compact if  $A$  is closed and bounded.
- (ii)  $A$  is compact if every sequence  $x_n$  of points of  $A$  has a convergent subsequence with  $\lim \in A$ .

*Proof.* Similarly, the analytic and sequential definitions of compactness have the property that that are equivalent under AC.  $\square$

**Consequence 8.0.5.** A set is Dedekind-infinite if and only if it is infinite.

*Proof.* ( $\rightarrow$ ) *If a set is Dedekind-infinite then it is infinite.* This is the easier direction that does not require AC. Suppose  $S$  is a Dedekind-infinite set. So,  $S$  is a set which has a proper subset  $T$  and there is a bijection from  $T$  to  $S$ . Thus,  $S$  and  $T$  are the same size,  $|S| = |T|$ . See Figure 21 for reference. By way of contradiction, assume  $S$  is finite. Since  $S$  is finite, by definition, there exists a natural number  $n$  such that there is a bijection,  $f$ , from  $S$  to

$\{1, 2, 3, 4, \dots, n\}$ .  $T$  is a proper subset, so  $S$  is not empty, hence  $n > 0$ . Let  $W = f(T)$ . Since  $|S| = n$ , where  $n$  is a finite number, and because  $f$  restricted to  $T$ ,  $f \upharpoonright T$ , is still a bijection, then  $|W| = |f(T)| = |T| < n$ . This is a contradiction since  $|S| = |T| < n$  but  $|S| = n$ . See Figure 22 for reference. Therefore, in ZF, if a set is Dedekind-infinite, then it is infinite.

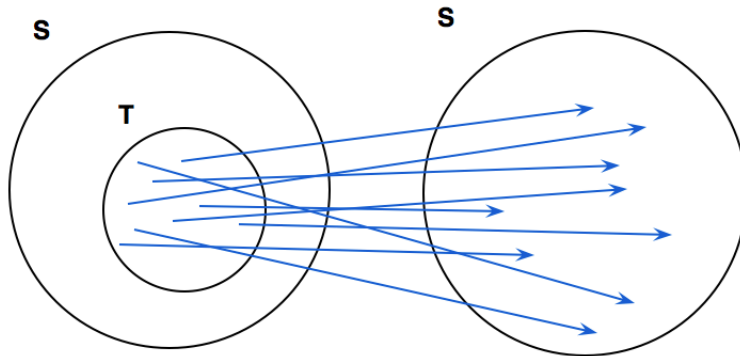


Figure 21: Bijection from  $T$  to  $S$  showing that  $|T| = |S|$ .

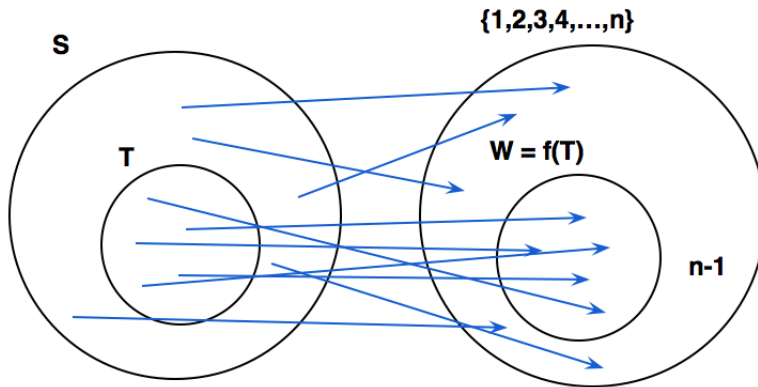


Figure 22: Bijection from  $S$  to  $\{1, 2, 3, 4, \dots, n\}$  and  $T$  to  $W$ .

( $\leftarrow$ ) To show that *if a set is infinite then it is Dedekind-infinite*, requires AC. Let  $S$  be an infinite set. We want to find a proper subset  $T$  of  $S$  such that there is a bijection from  $T$  to  $S$ . By a previous result, we know that  $S$  has a denumerable subset. This is where we use DAC. This means there is a bijection  $g$  from  $\omega = \{0, 1, 2, 3, 4, \dots\} = \mathbb{N}^0$  onto a proper subset  $T$  of  $S$ . This bijection enumerates the elements of  $T$ .

$$0 \rightarrow g(0) = t_1 \in T$$

$$1 \rightarrow g(1) = t_2 \in T$$

$$2 \rightarrow g(2) = t_3 \in T$$

$$n \rightarrow g(n) = t_{n+1} \in T$$

where  $T = \{t_1, t_2, t_3, \dots, t_n, \dots\} \neq S$ . Note that some of the elements of  $S$  will be the output of the function  $g$ , that is, for  $v \in S$ , it may happen that  $m \rightarrow g(m) = v$  and some elements of  $S$  will not be the output of the function  $g$ , that is, for  $r \in S$ , it may happen that  $r \neq g(m)$  for any  $m \in \omega = \mathbb{N}^0$ . We define a new function,  $f$ , such that  $f : S \rightarrow S$  by

$$f(x) = x,$$

if  $x \neq g(n)$ , for any  $n \in \mathbb{N}^0$  (this just says  $x \in S$  is not in the range of  $g$ ) and

$$f(x) = g(n + 1)$$

if  $x = g(n)$ , for some  $n \in \mathbb{N}^0$  (this takes care of the case when  $x$  is in the range of  $g$ ).

For example, let  $S = \{0, 1, 2, 3, 4, 5, 6, \dots\}$  and let  $g : \mathbb{N}^0 \rightarrow S$  by  $g(n) = 3n$ .  $g(0) = 0, g(1) = 3, g(2) = 6, g(3) = 9, g(4) = 12, g(5) = 15, \dots$ . Clearly, this is a bijection from  $\mathbb{N}^0$  onto the proper subset  $\{0, 3, 6, 9, 12, \dots\}$  of  $\{0, 1, 2, 3, 4, 5, \dots\}$ . Now,  $f(0) = g(1) = 3$ , since  $0 = g(0)$  and  $g(0 + 1) = g(1) = 3$ .  $f(1) = 1$ , since  $1 \neq g(n)$  for any  $n \in \mathbb{N}^0$ .  $f(2) = 2$ ,

since  $2 \neq g(n)$  for any  $n \in \mathbb{N}^0$ .  $f(3) = g(2) = 6$ , since  $3 = g(1)$  and  $g(1+1) = g(2) = 6$ .  
 $f(4) = 4$ , since  $4 \neq g(n)$ , for any  $n \in \mathbb{N}^0$ .  $f(5) = 5$ , since  $5 \neq g(n)$ , for any  $n \in \mathbb{N}^0$ .  
 $f(6) = g(3) = 9$ , since  $6 = g(2)$  and  $g(2+1) = g(3) = 9$ . Continuing in this way, we see the following mapping:

$$0 \rightarrow 3$$

$$1 \rightarrow 1$$

$$2 \rightarrow 2$$

$$3 \rightarrow 6$$

$$4 \rightarrow 4$$

$$5 \rightarrow 5$$

$$6 \rightarrow 9$$

$$7 \rightarrow 7$$

$$8 \rightarrow 8$$

$$9 \rightarrow 12$$

$$10 \rightarrow 10$$

$$11 \rightarrow 11$$

$$12 \rightarrow 15$$

$$\vdots$$

$$n \rightarrow n, n \neq 3m$$

$$n \rightarrow n + 3, n = 3m$$

We note that nothing maps to 0. So,  $f(S)$  is a proper subset of  $S$ . In fact, this is always

true. So, this will allow us to always construct a 1 – 1 mapping from  $S$  onto a proper subset  $f(S) = T$ . Every element of  $S$  gets mapped to a unique element of the proper subset  $f(S)$ . Now we go back to the general proof: Let  $T = f(S)$ . Note  $f(S) \subseteq S$ , since  $f$  maps to either  $x \in S$  or to  $g(n + 1) \in S$ . Show  $f$  is 1 – 1. Then, show  $T$  is a proper subset of  $S$ , by showing that  $g(0) = z$  does not get anything mapped to it by  $f$ , thus  $T$  does not contain the element of  $S$  represented by  $z$ . To show that nothing in  $S$  maps to  $z$  under the function  $f$ , we consider two cases in the definition. Case 1 is: could  $f$  map  $z$  to  $z$ ? The answer to this is no because by definition, the only way  $z$  could map to itself is if  $z$  was not an element of the output of function  $g$ . That is, if  $x \neq g(n)$ , for any  $n \in \mathbb{N}^0 = \{0, 1, 2, 3, 4, \dots\}$ , but  $z = g(0)$ . The other case is: could  $f$  map something else to  $z$ ? The answer to this is also no because suppose  $f(v) = z$ . By definition, the only way this could happen is if  $v = g(0)$ , but if that happens, then the function  $f$  says  $v = g(n + 1)$ . This is a contradiction, since  $f(v)$  can not be both equal to  $g(n)$  and  $g(n + 1)$  since  $g$  is a bijection. Now we will show that  $f$  is 1-1. Suppose for  $a, b$  in  $S$  that  $f(a) = f(b)$ . We must prove that  $a = b$ . Are  $a$  and  $b$  in the range of  $g$ ? There are three cases:

1. If both  $a, b$  are not in the range of  $g$ , then  $a = f(a) = f(b) = b$ . So,  $a = b$ .
2. If both  $a, b$  are in the range of  $g$ , then  $a = g(n)$  for some  $n$  and  $f(a) = g(n + 1)$  and  $b = g(m)$  for some  $m$  and  $f(b) = g(m + 1)$ . But,  $f(a) = f(b)$ , so  $g(n + 1) = g(m + 1)$  and  $g$  is a bijection so  $n = m$  which means  $a = g(n) = g(m) = b$ . Hence again  $a = b$ .
3. Without loss of generality, suppose  $f(a) = a$  and  $f(b) = g(m + 1)$ .  $f(a) = a$  means that  $a \neq g(n)$ , for any  $n \in \mathbb{N}^0$  and  $f(b) = g(m + 1)$  means  $b = g(m)$  for some  $m \in \mathbb{N}^0$ . However,  $f(a) = f(b)$ , so  $a = f(a) = f(b) = g(m + 1)$  and the result that  $a$  is equal to  $g(m + 1)$  contradicts the result that  $a \neq g(n)$  for any  $n \in \mathbb{N}^0$ . Thus, this case never occurs when  $f(a) = f(b)$ .

Thus, the mapping  $f$  is a 1 – 1 mapping from  $S$  into  $S$  but not onto, hence  $f(S) = T$  is a proper subset of  $S$ . □

### 8.2.3 Consequences of DAC Not Discussed

**Consequence 8.0.6.** Every subspace of a separable metric space is separable.

**Consequence 8.0.7.** The first uncountable ordinal  $\omega_1$  is not a limit of a countable increasing sequence of ordinals.

### 8.2.4 Dependent Choice

Another weaker variant of AC is Dependent Choice or DPC. We will first provide some more necessary definitions to help us understand DPC.

**Definition 8.10.** *Dependent Choice (DPC):* If  $S$  is a relation on a set  $A$ , such that for every  $x$  in  $A$ , there exists some  $y$  in  $A$  with  $x S y$ , then there is a sequence  $a_1, a_2, a_3, \dots$  such that for every natural number  $n$ ,  $a_n$  is in  $A$  and  $a_n S a_{n+1}$  holds. That is, we can construct a sequence  $a_1, a_2, a_3, \dots$  where the choice of each new element,  $a_m$ , in the sequence is *dependent* on the previous choices  $a_1, a_2, a_3, \dots, a_{m-1}$ .

**Remark 11.** DPC is related to AC and DAC in the following way:

$$AC \rightarrow DPC \rightarrow DAC$$

It is obvious that  $AC \rightarrow DPC$  is true. So, we will only prove  $DPC \rightarrow DAC$ .

*Proof.*  $DPC \rightarrow DAC$

Let  $T = \{S_0, S_1, S_2, S_3, S_4, \dots\}$  be a countable family of nonempty sets. For example, let  $S_0 = \{h, k, a, 4, 2, e, \dots\}$ , let  $S_1 = \{\text{real numbers}\}$ , let  $S_2 = \{\text{@, !, *, \%, \#, \dots}\}$ , let  $S_3 = \{y, w, r, 5, o, 1, z, \dots\}$ , and let  $S_4 = \{\text{names of trees}\}$ . To find a choice function, let  $A$  be the set of all finite sequences

$$\langle x_0, x_1, x_2, x_3, x_4, \dots, x_n \rangle$$

where  $x_m \in S_m$ . For example, some of the elements of  $A$  are  $A = \{\langle a \rangle, \langle h \rangle, \langle 2 \rangle, \dots, \langle k, \sqrt{2} \rangle, \langle 4, 105 \rangle, \langle e, -27 \rangle, \dots, \langle h, 88, ! \rangle, \langle a, \pi, * \rangle, \langle 2, 600, @ \rangle, \dots, \langle e, 43, \%, 1 \rangle,$

$\langle 4, -999, \# \rangle, \dots\}$ . Define a relation  $R$  on these sequences by, a sequence of  $n$  elements is related to a sequence of  $n + 1$  elements if and only if the two sequences agree on the first  $n$  terms. For example,  $\langle k \rangle R \langle k, \sqrt{2} \rangle$ , but  $\langle k \rangle$  is not related to  $\langle a, \sqrt{2} \rangle$ . We now have the set-up for dependent choice: we have a relation  $R$  on a set  $A$  such that for each sequence  $x$  in  $A$ , there is a sequence  $y$  in  $A$  with  $xRy$ . For the sequence  $\langle e, 55, @, w \rangle$  there is a sequence  $\langle e, 55, @, w, \text{pine} \rangle$ , that is,  $\langle e, 55, @, w \rangle R \langle e, 55, @, w, \text{pine} \rangle$ . Applying Dependent Choice tells us there is a sequence which “chooses” an element from the first set  $S_0$ , and then “chooses” a second element from the set  $S_1$ , depending on the first selection. We call these  $x_0, x_1, x_2, \dots$ , but this is just a set with one element “chosen” from each of the sets  $S_j$  in our countable family of sets  $T$ . Thus,  $\text{DPC} \rightarrow \text{DAC}$ .  $\square$

## 8.3 Equivalent Variations

### 8.3.1 Well-Ordering Theorem

The Well-Ordering Theorem (WOT) was first known as the Well-Ordering Principle. In 1883, mathematician Georg Cantor approached the Well-Ordering Principle by declaring it as a self-evident logical law. After a decade, Cantor attempted to deduce this principle, which turned it into the Well-Ordering Problem. Cantor repeatedly realized that it was not able to be proved under his assumptions. It was then the mathematician Ernst Zermelo, giving a much clearer description of the problem than Cantor, showed that it had many amazing consequences in mathematics. This transformed it into the Well-Ordering Theorem. Zermelo came up with the familiar variation of AC and used it to prove WOT, later realizing the two were equivalent.

**Theorem 8.1.** *Well-Ordering Theorem (WOT):* Every set can be well-ordered.

Transfinite recursion is used in a proof involving WOT. Thus, we define transfinite recursion first for a deeper understanding.

**Definition 8.11.** *Transfinite Recursion:* Transfinite recursion is a way to build up a function or some other mathematical object. It is a technique to construct a function from a



base case, then using a recursive technique to define new elements from previously defined ones. We usually define these functions (or objects) on the ordinal numbers. The goal is to assign a value to the function  $f$  for each ordinal up to some  $\alpha$  and that value must be in the given well-ordered set. The recursive part is that we define the next value of the function in terms of the previously defined function values.

Transfinite recursion can also be roughly described in this way: Given a well-ordered set, an ordinal  $\alpha$ , and a map  $g$  which takes the set of maps from  $\beta$  to  $A$ , into  $A$ , with  $\beta < \alpha$ , then we can construct a function  $f : \alpha \rightarrow A$  such that  $\forall \beta < \alpha, f(\beta) = (f \upharpoonright \beta)$ . The notation  $(f \upharpoonright \beta)$  is called the “restriction of  $f$  to  $\beta$ ”. This means  $g$  will tell us how to extend the function which maps smaller ordinals into  $A$ , to a function which defines the value of  $f$  further. The function  $g$  tells us how to define the value at a successor ordinal  $\zeta + 1$ , when we have defined the values up to some ordinal  $\zeta$ . This will extend the function to  $\zeta + 1$ . The function  $g$  tells us how to define the value at a limit ordinal  $\eta$ , when we have defined all the values  $< \eta$ . This will extend the function to  $\eta$ . When we are done we will have defined a function,  $f$ , from  $\alpha$  to  $A$ . This allows us to construct functions on well-ordered sets. In other words, having constructed the function  $f$  from 0 to all the ordinals up to but not including  $\beta$ , we use this information and the function  $g$  to define the next value of the function,  $f(\beta)$ .

**Example 8.8.** Let  $\alpha$  be the ordinal  $\omega^2$  and let  $A = \omega$ . We will build a map from  $\alpha$  to  $A$  by transfinite recursion,  $f(\beta) = g(f \upharpoonright \beta)$ . We define the function  $g : \alpha \rightarrow A$  where  $(\beta < \alpha)$  by: let  $g(\beta) = 0$  if  $\beta$  is a limit ordinal, and  $g(\beta) = g(\beta - 1) + 1$  if  $\beta$  is a successor ordinal.

We use recursion:  $f(\beta) = g(f \upharpoonright \beta)$

0 is a limit ordinal, so  $f(0) = g(0) = 0$ ,

1 is a successor ordinal, so  $f(1) = g(f \upharpoonright 1) = g(1) = g(0) + 1 = 0 + 1 = 1$ ,

2 is a successor ordinal, so  $f(2) = g(f \upharpoonright 2) = g(2) = g(1) + 1 = 1 + 1 = 2$ ,

3 is a successor ordinal, so  $f(3) = g(f \upharpoonright 3) = g(3) = g(2) + 1 = 2 + 1 = 3$ ,

⋮

$n + 1$  is a successor ordinal, so  $f(n + 1) = g(f \upharpoonright n + 1) = g(n + 1) = g(n) + 1 = n + 1$ ,

$\omega$  is a limit ordinal, so  $f(\omega) = 0$ ,

$\omega + 1$  is a successor ordinal, so  $f(\omega + 1) = g(f \uparrow \omega + 1) = g(\omega + 1) = g(\omega) + 1 = 0 + 1 = 1$ ,

$\vdots$

$\omega + n$  is a successor ordinal, so  $f(\omega + n) = g(f \uparrow \omega + n) = g(\omega + n) = g(\omega + n - 1) + 1 = 0 + n - 1 + 1 = n$ ,

$\vdots$

$\omega^2$  is a limit ordinal, so  $f(\omega^2) = 0$ .

Thus, we have defined a function from  $\omega^2$  to the set  $A = \omega$ .

**Theorem 8.2.**  $AC \leftrightarrow WOT$

*Proof.* ( $\rightarrow$ ) Let  $S$  be a set. The goal is to prove that a well-ordering of  $S$  exists using AC. To start, we let  $T = \mathcal{P}(S) - \emptyset$ . So,  $T$  is the collection of all non-empty subsets of  $S$ . To apply AC on  $T$ , we assert the existence of a function  $f$ , which chooses an element from each of the non-empty subsets of  $S$ . Doing so by transfinite recursion, we define a sequence  $\langle a_\alpha : \alpha < \theta \rangle$  that enumerates  $S$ . We let for every  $\alpha$

$$a_\alpha = f(S - \{a_\xi : \xi < \alpha\}).$$

When  $S - \{a_\xi : \xi < \alpha\}$  is empty, we are done. Let  $\theta$  be the least ordinal such that  $S = \{a_\xi : \xi < \theta\}$ . Clearly,  $\langle a_\alpha : \alpha < \theta \rangle$  enumerates  $S$ .

For example, let  $S =$  the set of real numbers,  $\mathbb{R}$ , and let  $T = \mathcal{P}(\mathbb{R}) - \emptyset$ . Applying AC, we assert the existence of a function,  $f$ , which selects a real number out of each subset of real numbers. Say  $f(\mathbb{R}) = 17$ . We call  $a_0 = 17$ . Now, we consider  $f(\mathbb{R} - \{17\})$ . Say  $f(\mathbb{R} - \{17\}) = e$ . We call  $a_1 = e$ . Therefore, we have enumerated the first two numbers. Continuing in this way we enumerate the first countably many real numbers and eventually get to  $a_\omega$ . Then,  $a_\omega = f(\mathbb{R} - a_\xi : \xi < \omega) = f(\mathbb{R} - \{a_0, a_1, a_2, \dots, a_n, \dots\})$ . Then  $a_{\omega+1} = f(\mathbb{R} - a_\xi : \xi < \omega + 1) = f(\mathbb{R} - \{a_0, a_1, a_2, \dots, a_\omega\})$ . This continues until we order all the real numbers.

( $\leftarrow$ ) If every set can be well-ordered, then every family  $T$  of nonempty sets has a choice function. To see this, well-order  $\bigcup T = \bigcup_{x \in A} A$  where  $A \in T$  and let  $f(A)$  be the least element of  $A$  for every  $A \in T$ . We can do this because every subset of  $\bigcup T$  (where each  $A$  is such a subset) has a least element by well-ordering.  $\square$

### 8.3.2 Zorn's Lemma

Zorn's Lemma, also known as Maximal Principle, was published in 1935 by mathematician Max Zorn. Max Zorn was the first to regard ZL not as a theorem, but as an axiom. When Zorn published ZL in a paper, mathematicians quickly put it to use. Zorn himself used it to demonstrate many known theorems in topology and algebra to avoid using the WOT. Until Zorn illustrated how useful ZL could be, algebraists remained ignorant of it because it had been formulated previously in set theory.

“Our next axiom equivalent to AC is perhaps the one most familiar to the working mathematician outside of set theory. For historical reasons it is known as ‘lemma’, but it is indeed just another formulation of the Axiom of Choice” (Devlin, 1993, p. 60). Before we discuss this equivalent, some more definitions and examples are provided to help with our understanding.

**Definition 8.12.** *Maximal Element:* Let  $(P, \preceq)$  be a poset. An element  $a$  of  $P$  is said to be *maximal* in  $P$  if and only if there is no  $b$  in  $P$  such that  $a < b$ . A poset can have many maximal elements.

**Example 8.9.** When we reverse definition 4.9 of a minimal element, Figure 8 becomes inverted. Thus, our minimal elements are now maximal elements. We see some of the maximal elements of  $(E, |)$  are 2 from subset  $\{2, 8, 24\}$ , 3 from subset  $\{3, 6, 12, 24\}$ , and 6 from subset  $\{6, 12, 24\}$ .

**Definition 8.13.** *Upper Bound:* Given a partially ordered set  $(P, \preceq)$ , an element  $u$  is called an *upper bound* of the subset  $S$  if for every element  $s \in S$ ,  $s \preceq u$ .

**Definition 8.14.** *Lower Bound:* Given a partially ordered set  $(P, \preceq)$ , an element  $w$  is called a *lower bound* of the subset  $S$  if for every element  $s \in S$ ,  $w \preceq s$ .

**Definition 8.15.** *Bounded Above:* Given a partially ordered set  $(P, \preceq)$ , a subset  $S \in P$  is said to be *bounded above* if there exists an upper bound in the set.

**Definition 8.16.** *Bounded Below:* Given a partially ordered set  $(P, \preceq)$ , a subset  $S \in P$  is said to be *bounded below* if there exists a lower bound in the set.

**Definition 8.17.** *Bounded:* A set is *bounded* if it is bounded both above and below. A set is *unbounded* if it is not bounded above or not bounded below.

**Definition 8.18.** *Supremum:* Given a partially ordered set  $(P, \preceq)$ , if a subset  $S \in P$  is bounded above, then a number  $u$  is said to be a *supremum* of  $S$  if:

1.  $u$  is an upper bound of  $S$
2.  $v$  is any upper bound of  $S$ , then  $u \preceq v$ .

**Definition 8.19.** *Infimum:* Given a partially ordered set  $(P, \preceq)$ , if a subset  $S \in P$  is bounded below, then a number  $w$  is said to be an *infimum* of  $S$  if:

1.  $w$  is a lower bound of  $S$
2.  $t$  is any lower bound of  $S$ , then  $t \preceq w$ .



Figure 23: Graphical representation of supremum, infimum, upper bounds, and lower bounds on the open interval  $(a, b)$ .

**Definition 8.20.** *Transfinite Sequence:* A *transfinite sequence* is a function whose domain is a transfinite ordinal. If  $f$  is a sequence and  $\text{dom}(f) = \alpha$ , we say  $f$  is an  $\alpha$ -sequence. If  $f(\xi) = x_\xi$  for all  $\xi < \alpha$ , we often write  $\langle x_\xi : \xi < \alpha \rangle$  in place of  $f$ . Then, for  $\beta < \alpha$ ,  $\langle x_\xi : \xi < \beta \rangle$  denotes  $f(\beta)$ . This clearly gives a precise meaning to what we generally think of as a (transfinite, perhaps) sequence. The ‘sequences’ of elementary analysis are just the special case of  $\omega$ -sequences, of course; so,  $\{a_n\}_{n=0}^\infty = \langle a_n : n < \omega \rangle$ .

**Definition 8.21.** *Chain:* A subset  $C$  of a partially ordered set  $(P, \preceq)$  is a *chain* in  $P$  if  $C$  is linearly ordered by  $\preceq$ . We recall that all subsets of tosets are chains.

**Theorem 8.3.** *Zorn’s Lemma (ZL):* Let  $(P, \preceq)$  be a nonempty partially ordered set and let every chain in  $P$  have an upper bound. Then  $P$  has a maximal element.

**Theorem 8.4.** *WOT  $\leftrightarrow$  ZL:*

*Proof.* ( $\rightarrow$ ) Let  $(P, \preceq)$  be a nonempty poset and assume that every chain in  $P$  has an upper bound. We find a maximal element of  $P$ . By assumption, the set  $P$  can be well-ordered, i.e., there is an enumeration

$$P = \{p_0, p_1, \dots, p_\xi, \dots\}$$

where  $(\xi < \alpha)$  for some ordinal number  $\alpha$ . By transfinite recursion, let

$$c_0 = p_0$$

and

$$c_\xi = p_\gamma,$$

where  $\gamma$  is the least ordinal such that  $p_\gamma$  is an upper bound of the chain  $C = \{c_\eta : \eta < \xi\}$  and  $p_\gamma \notin C$ . Note that  $\{c_\eta : \eta < \xi\}$  is always a chain and that  $p_\gamma$  exists unless  $c_{\xi-1}$  is a maximal element of  $P$ . Eventually, the construction comes to a halt, and we obtain a maximal element of  $P$ .  $\square$

( $\leftarrow$ ) It can be shown that ZL  $\rightarrow$  WOT.

“ZL could be just taken as an axiom of set theory. However, it is not psychologically satisfactory as an axiom, because its statement is too involved, and one does not visualize easily the existence of the maximal element asserted in the statement” (Lang, 2002, p. 881).

### 8.3.3 Equivalences Not Discussed

The following are equivalences we will not discuss or prove due to lack of tools and time.

**Theorem 8.5.** *Tukey’s Lemma:* Let  $F$  be a nonempty family of sets. If  $F$  has a finite character, then  $F$  has a maximal element (maximal with respect to inclusion  $\subseteq$ ).

**Theorem 8.6.** *Bases for Vector Spaces:* Every vector space has a basis.

## 8.4 Friction with AC

Constructivists do not accept certain parts of LAST, in particular, the law of excluded middle. The law of excluded middle is the tautology  $P \vee \neg P$  in propositional logic. This is not accepted by constructivists because we can obtain  $P$  by way of contradiction. The proof by contradiction goes as follows: Let us assume  $\neg P$  is true. This assumption leads us to a contradiction. Therefore, by law of excluded middle, since  $\neg P$  did not happen, we can conclude  $P$  did happen. So, the fact that  $\neg P$  is false implies that  $P$  is true. Thus, we have gotten  $P$  without actually constructing it. In general, constructivists are unsatisfied with AC because it asserts the existence of a set without giving us that set. In other words, we are not constructing anything when we discuss AC.

For many years, mathematicians criticized AC for its non-constructive character and some even believed that it lead to a contradiction. It was not until the year 1914 that there was evidence of this, when a seemingly unnatural consequence emerged, Hausdorff’s paradox. Hausdorff’s paradox says that one half of a sphere is congruent to one third of the same sphere. The first to respond to this paradox, Borel, felt as though the offender of this paradox was AC. Borel eventually found the contradiction, relating it to AC, and believed

that his years of resistance to AC had been justified. But, other mathematicians did not agree with him. Thus, there was still disagreements about AC. Some mathematicians were convinced of its greatness, while others were not.

*“The Axiom of Choice is used (in classical mathematics) to extract elements from equivalence classes where they should never have been put in the first place”* (Bishop, 2012, p. 9).

*“The Axiom of Choice is obviously true, the Well-Ordering Principle is obviously false, and who can tell about Zorn’s Lemma?”* (Krantz, 2002, p.121).

## 9 Continuum Hypothesis

Throughout this thesis, we have discussed different axioms and arithmetic having to do with set theory. As seen previously, things begin to get messy when we talk about sets in the transfinite. This remains true when we compare the sizes of infinite sets. We know from Cantor’s theorem that  $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$ . We saw previously that  $|\mathbb{N}| = \aleph_0$  and we know that  $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$ . Some mathematicians believe the notion that  $|\mathbb{R}| = \aleph_1$ , the next  $\aleph$  after  $\aleph_0$ . But, the question is, is there an infinite set that has a size in between  $\aleph_0$  and  $\aleph_1$  or is  $|\mathbb{R}| = \aleph_1$ ? The **Continuum Hypothesis** (CH) asserts that there is no infinity between the natural numbers and the real numbers, thus  $|\mathbb{R}| = \aleph_1$ . This question has puzzled mathematicians for decades; some mathematicians believe this hypothesis to be true, while others believe it to be false.

After CH was announced and conjectured, mathematicians took it further to ask is  $2^{\aleph_1} = \aleph_2$ ? The more general statement was that for all alephs,  $2^{\aleph_\alpha} = \aleph_{\alpha+1}$  for each ordinal number  $\alpha$ . This statement is known to be the **Generalized Continuum Hypothesis** (GCH). It was shown by mathematician Kurt Gödel that GCH is consistent with the axioms in ZFC. Equally remarkable, mathematician Paul Cohen showed that GCH is independent from the axioms of ZFC. It turns out, under the assumption of GCH, some interesting events happen. One of those interesting events is that an extreme simplification of cardinal exponentiation occurs. Unlike the axioms in ZF and the Axiom of Choice (AC), most mathematicians tend

not to assume GCH as an additional axiom in set theory because they have no idea whether it is true or false. CH and GCH have been helpful for proving several important theorems, even before they were proved in ZFC or turned out not to be provable in ZFC.

The first mathematician to tackle CH was David Hilbert who gave a proof in 1925. The proof was incorrect but contained some important ideas. Then, Gödel came into the picture with his incompleteness theorems, crushing Hilbert's idea that every mathematical question can be solved, although, Gödel became an advocate for being able to solve CH. In 1937, Gödel proved that with the current mathematical methods, we can not prove that CH is false in ZFC. Later, Paul Cohen invented the forcing method, which is a way to add new reals to a model of the mathematical universe. Together the work of Gödel and Cohen combined to demonstrate that CH could neither be proved nor disproved using the axioms of ZFC. In other words, CH is independent of ZFC. So, this means that new axioms would have to be added to our system in order to prove CH. There are two main contenders for additions to ZFC: the forcing axioms and the inner model axiom  $V = L$ , where  $V$  represents our hierarchy of sets. For the inner model axiom, if  $V = L$ , then CH holds. So, what do we really even know about infinities? Some argue that we know a lot, while others assert that we know only as much as 1,2, 3.

The great mathematician Henri Poincaré who lived at the turn of the last century (1900), wrote that Cantor's "*set theory was a malady, a perverse disease from which mathematics would someday recover*". But perhaps the greatest mathematician of his time, David Hilbert, advocating for the richness of logic, set theory and the axiomatic method he pursued for more than three decades, wrote that "*No one can expel us from the paradise that Cantor has created for us*".



## 10 Appendix 1: Other Axiomatic Schemes

- Peano Postulates: (Smullyan, 2017, p. 255)

**Axiom 10.1.** *0 is a natural number.*

**Axiom 10.2.** *If  $n$  is a natural number, so is  $n^+$ .*

**Axiom 10.3.** *There is no natural number such that  $n^+ = 0$ .*

**Axiom 10.4.** *If  $n^+ = m^+$ , then  $n = m$ .*

**Axiom 10.5.** *Principle of Mathematical Induction: Every inductive set contains all the natural numbers.*

- Euclid's Axioms: (Joyce, 1996)

**Axiom 10.6.** *To draw a straight line from any point to any point.*

**Axiom 10.7.** *To produce a finite straight line continuously in a straight line.*

**Axiom 10.8.** *To describe a circle with any center and radius.*

**Axiom 10.9.** *That all right angles equal one another.*

**Axiom 10.10.** *That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than the two right angles.*

- Hilbert's Axioms: (Hilbert, 1980)

Incidence:

**Axiom 10.11.** *Given any two points, there exists a line containing both of them.*

**Axiom 10.12.** *Given any two points, there exists no more than one line containing both points. I.e., the line described in Axiom 10.11 is unique.*

**Axiom 10.13.** *A line contains at least two points, and given any line, there exists at least one point not on it.*

**Axiom 10.14.** *Given any three points not contained in one line, there exists a plane containing all three points. Every plane contains at least one point.*

**Axiom 10.15.** *Given any three points not contained in one line, there exists only one plane containing all three points.*

**Axiom 10.16.** *If two points contained in line  $m$  lie in some plane  $\alpha$ , then  $\alpha$  contains every point in  $m$ .*

**Axiom 10.17.** *If the planes  $\alpha$  and  $\beta$  both contain the point  $A$ , then  $\alpha$  and  $\beta$  both contain at least one other point.*

**Axiom 10.18.** *There exist at least four points not all contained in the same plane.*

Order:

**Axiom 10.19.** *If a point  $B$  is between points  $A$  and  $C$ ,  $B$  is also between  $C$  and  $A$ , and there exists a line containing the points  $A, B, C$ .*

**Axiom 10.20.** *Given two points  $A$  and  $C$ , there exists a point  $B$  on the line  $AC$  such that  $C$  lies between  $A$  and  $B$ .*

**Axiom 10.21.** *Given any three points contained in one line, one and only one of the three points is between the other two.*

**Axiom 10.22.** *Axiom of Pasch: Given three points  $A, B, C$  not contained in one line, and given a line  $m$  contained in the plane  $ABC$  but not containing any of  $A, B, C$ : if  $m$  contains a point on the segment  $AB$ , then  $m$  also contains a point on the segment  $AC$  or on the segment  $BC$ .*

Congruence:

**Axiom 10.23.** *Given two points  $A, B$ , and a point  $A'$  on line  $m$ , there exist two and only two points  $C$  and  $D$ , such that  $A'$  is between  $C$  and  $D$ , and  $AB \cong A'C$  and  $AB \cong A'D$ .*

**Axiom 10.24.** *If  $CD \cong AB$  and  $EF \cong AB$ , then  $CD \cong EF$ .*

**Axiom 10.25.** Let line  $m$  include the segments  $AB$  and  $BC$  whose only common point is  $B$ , and let line  $m$  or  $m'$  include the segments  $A'B'$  and  $B'C'$  whose only common point is  $B'$ . If  $AB \cong A'B'$  and  $BC \cong B'C'$  then  $AC \cong A'C'$ .

**Axiom 10.26.** Given the angle  $\angle ABC$  and ray  $B'C'$ , there exist two and only two rays,  $B'D$  and  $B'E$ , such that  $\angle DB'C' \cong \angle ABC$  and  $\angle EB'C' \cong \angle ABC$ .

**Axiom 10.27.** Given two triangles  $\triangle ABC$  and  $\triangle A'B'C'$  such that  $AB \cong A'B'$ ,  $AC \cong A'C'$ , and  $\angle BAC \cong \angle B'A'C'$ , then  $\triangle ABC \cong \triangle A'B'C'$ .

Parallels:

**Axiom 10.28.** Playfair's postulate: Given a line  $m$ , a point  $A$  not on  $m$ , and a plane containing both  $m$  and  $A$ : in that plane, there is at most one line containing  $A$  and not containing any point on  $m$ .

Continuity:

**Axiom 10.29.** Axiom of Archimedes: Given the line segment  $CD$  and the ray  $AB$ , there exist  $n$  points  $A_1, \dots, A_n$  on  $AB$ , such that  $A_j A_{j+1} \cong CD$ ,  $1 \leq j < n$ . Moreover,  $B$  is between  $A_1$  and  $A_n$ .

**Axiom 10.30.** Line completeness. Adding points to a line results in an object that violates one or more of the following axioms: 10.11, 10.12, 10.19, 10.20, 10.23, 10.24, 10.29.

- Nuemann-Bernays-Gödel Axioms (Stoll, 2006)

**Axiom 10.31.** Axiom of extension: If  $A$  and  $B$  are classes and if, for all sets  $x$ ,  $x \in A$  if and only if  $x \in B$ , then  $A = B$ .

**Axiom 10.32.** Axiom of the empty set: There exists a set  $A$  such that, for all  $x$ , it is false that  $x \in A$ .

**Axiom 10.33.** Axiom schema for class formation: If  $S(x)$  is a condition on  $x$  in which only set variables are introduced by the phrase "for all" or "for some" and in which  $B$  is not free, then there exists a class  $B$  such that  $x \in B$  if and only if  $S(x)$ .

**Axiom 10.34.** *Axiom of pairing: If  $A$  and  $B$  are sets, there exists a set (symbolized  $\{A, B\}$  and called the unordered pair of  $A$  and  $B$ ) having  $A$  and  $B$  as its sole members.*

**Axiom 10.35.** *Axiom of union: If  $C$  is a set, there exists a set  $A$  such that  $x \in A$  if and only if  $x \in B$  for some member  $B$  of  $C$ .*

**Axiom 10.36.** *Axiom of power set: If  $A$  is a set, there exists a set  $B$ , called its power set, such that  $x \in B$  if and only if  $x \subseteq A$ .*

**Axiom 10.37.** *Axiom of infinity: There exists a set  $A$  such that  $\emptyset \in A$  and, if  $x \in A$ , then  $(x \cup \{x\}) \in A$ , in which  $x \cup \{x\}$  is the set  $x$  with  $x$  adjoined as the further member.*

**Axiom 10.38.** *Axiom of choice: If  $A$  is a set the elements of which are nonempty sets, then there exists a function  $f$  with domain  $A$  such that, for each member  $B$  of  $A$ ,  $f(B) \in B$ .*

**Axiom 10.39.** *Axiom of replacement: If (the class)  $X$  is a function and  $A$  is a set, then there exists a set  $B$  such that  $y \in B$  if and only if for some  $x$ ,  $(x, y) \in X$  and  $x \in A$ ; i.e., the range of the restriction of a function  $X$  to a domain that is a set is also a set.*

**Axiom 10.40.** *Axiom of restriction (foundation axioms): Every nonempty class  $A$  contains an element  $B$  such that  $A \cap B = \emptyset$ .*

## 11 Appendix 2: Definitions

- **Anti-Symmetric:** Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is anti-symmetric if and only if  $(\forall a, b \in X)[(aRb \wedge bRa) \rightarrow a = b]$ .
- **Axiom:** In mathematics, an *axiom* is an unproven rule that is accepted as true because it is self-evident.
- **Axiom of Choice:** Given any family  $T$  of non-empty sets, there is a function  $f$  which assigns to each member  $A$  of  $T$  an element  $f(A)$  of  $A$ . Such an  $f$  is now called a choice function for  $T$ ;  $\forall T [\emptyset \notin T \rightarrow \exists f [ \text{dom}(f) = T \wedge \forall y \in T (f(y) \in y) ]]$ .
- **Axiomatic System:** An *axiomatic system* is any set of axioms which can be used together to derive theorems.
- **Bijection:** Given two sets  $A$  and  $B$ , a function is defined to be a *bijection* if it is both an injection and a surjection.
- **Binary Relation:** A binary relation over sets  $X$  and  $Y$  is a subset  $R$  of  $X \times Y$ . If  $X = Y$ , we say  $R$  is a binary relation on  $Y$ . In general, an  $n$ -ary relation over  $A_1, \dots, A_n$  is a subset of  $A_1 \times \dots \times A_n = \{(a_1, \dots, a_n) \mid a_k \in A_k \text{ for each } k = 1, \dots, n\}$ . We say  $aRb$  if  $(a, b) \in R$ .
- **Bounded:** A set is *bounded* if it is bounded both above and below. A set is *unbounded* if it is not bounded above or not bounded below.
- **Bounded Above:** Given a partially ordered set  $(P, \preceq)$ , a subset  $S \in P$  is said to be *bounded above* if there exists an upper bound in the set.
- **Bounded Below:** Given a partially ordered set  $(P, \preceq)$ , a subset  $S \in P$  is said to be *bounded below* if there exists a lower bound in the set.
- **Cardinal Addition:** Let  $\mu$  and  $\nu$  be two arbitrary cardinal numbers. The sum of  $\mu$  and  $\nu$ , denoted  $\mu + \nu$ , is the cardinal number of  $B \cup C$  where  $B$  and  $C$  are any disjoint sets having cardinals  $\mu$  and  $\nu$  respectively. So, let  $\mu = |B|$  and  $\nu = |C|$  with  $B \cap C = \emptyset$ . Then,  $\mu + \nu = |(B \cup C)|$ .

- **Cardinal Addition (index sets):** Let  $T$  be an index set and  $C_t$  a set for each  $t \in T$  such that  $\gamma_t = |C_t|$  and  $C_t \cap C_r = \emptyset$  if  $t \neq r$ . Then, the sum of the cardinal numbers  $\gamma_t$ , denoted  $\sum_{t \in T} \gamma_t$ , is defined by  $\sum_{t \in T} \gamma_t = \left| \bigcup_{t \in T} C_t \right|$ .
- **Cardinal Exponentiation:** Let  $\mu$  and  $\nu$  be cardinal numbers. If  $\mu = |A|$  and  $\nu = |B|$ , then  $\mu^\nu = |B \nearrow A|$ , where  $B \nearrow A$  is the set of all functions from  $B$  to  $A$ .
- **Cardinal Multiplication:** Let  $\mu$  and  $\nu$  be arbitrary cardinal numbers with sets  $A$  and  $B$  such that  $\mu = |A|$  and  $\nu = |B|$ . Then, the product of  $\mu$  and  $\nu$ , denoted  $\mu \cdot \nu$  is defined by  $\mu \cdot \nu = |A \times B|$ .
- **Cardinal Number:** A *cardinal number* is an ordinal  $\alpha$  such that for no  $\beta < \alpha$  does there exist a bijection  $f : \beta \leftrightarrow \alpha$ .
- **Cartesian Product:** A Cartesian product of sets  $X$  and  $Y$  is the set of all ordered pairs which is defined to be the set  $X \times Y = \{(a, b) \mid a \in X \wedge b \in Y\}$ .
- **Chain:** A subset  $C$  of a partially ordered set  $(P, \preceq)$  is a *chain* in  $P$  if  $C$  is linearly ordered by  $\preceq$ . We recall that all subsets of tosets are chains.
- **Connected:** Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is connected if and only if  $(\forall a, b \in X)(a \neq b) \rightarrow [(aRb) \vee (bRa)]$ .
- **Countable:** A set  $T$  is *countable* if  $T$  is a finite set or  $T$  is a denumerable set. This suggests that the elements can be systematically listed or enumerated. Of course, in the infinite case, it is not possible to list all the elements, but it should be possible to describe an algorithm (bijection from  $\mathbb{N}$ ) that would systematically list the elements in a way to show that eventually every element would appear on the list.
- **Dedekind-Finite:** A set  $A$  is *Dedekind-finite* (DF) if there is no bijection from  $A$  onto any proper subset  $B$  of  $A$ .
- **Dedekind-Infinite:** A set  $A$  is *Dedekind-infinite* (DI) if there exists a bijection from  $A$  to a proper subset  $B$  of  $A$ .

- **Denumerable:** A set  $T$  is *denumerable* if there is a bijection from the positive natural numbers  $\mathbb{N}$  to  $T$ . This means that there exists a one-to-one and onto function from  $\mathbb{N}$  to  $T$ . Such a set is also called countably infinite.
- **Denumerable Choice:** Every countable family of nonempty sets has a choice function.
- **Dependent Choice:** If  $S$  is a relation on a set  $A$ , such that for every  $x$  in  $A$ , there exists some  $y$  in  $A$  with  $x S y$ , then there is a sequence  $a_1, a_2, a_3, \dots$  such that for every natural number  $n$ ,  $a_n$  is in  $A$  and  $a_n S a_{n+1}$  holds. That is, we can construct a sequence  $a_1, a_2, a_3, \dots$  where the choice of each new element,  $a_m$ , in the sequence is *dependent* on the previous choices  $a_1, a_2, a_3, \dots, a_{m-1}$ .
- **Finite:** A set  $T$  is *finite* if  $T$  is empty or if there is a bijection from  $T$  onto the set  $\{1, 2, 3, 4, \dots, n\}$  for some positive natural number  $n$ . The sets  $\{1\}, \{1, 2\}, \{1, 2, 3\}, \dots, \{1, 2, 3, 4, \dots, n\}$  for positive natural numbers  $n$ , are called the *initial segments* of the positive natural numbers. Thus, a set  $T$  is *finite* if  $T$  is empty or there is a bijection from  $T$  onto an initial segment of the positive natural numbers.
- **Function:** A *function*  $f : A \rightarrow B$  is a subset of  $A \times B$  such that if  $(a, b), (a, b') \in f$ , then  $b = b'$ .
- **Infimum:** Given a partially ordered set  $(P, \preceq)$ , if a subset  $S \in P$  is bounded below, then a number  $w$  is said to be an *infimum* of  $S$  if:
  1.  $w$  is a lower bound of  $S$
  2.  $t$  is any lower bound of  $S$ , then  $t \preceq w$ .
- **Infinite:** A set  $T$  is *infinite* if it is not finite. This means that there is no bijection from  $T$  onto the set  $\{1, 2, 3, \dots, n\}$  for any positive natural number  $n$ . Thus, there is no bijection from  $T$  onto an initial segment of the positive natural numbers.
- **Injection:** Given two sets  $A$  and  $B$ , a function  $f : A \rightarrow B$  is defined as an *injection* (one-to-one) if every input of  $A$  has its own unique output in  $B$ . In other words,

$$(\forall a, b \in A) f(a) = f(b) \rightarrow a = b.$$

- **Limit Ordinal:** A *limit ordinal* is an ordinal that will have no greatest member; it is not the successor of any ordinal. Thus,  $\omega$  is a limit ordinal. Also, zero is a limit ordinal because it is not the successor of any ordinal.
- **Lower Bound:** Given a partially ordered set  $(P, \preceq)$ , an element  $w$  is called a *lower bound* of the subset  $S$  if for every element  $s \in S$ ,  $w \preceq s$ .
- **Maximal Element:** Let  $(P, \preceq)$  be a poset. An element  $a$  of  $P$  is said to be *maximal* in  $P$  if and only if there is no  $b$  in  $P$  such that  $a < b$ . A poset can have many maximal elements.
- **Minimal Element:** Let  $(X, \preceq)$  be a poset and let  $Y \subseteq X$ . An element  $a$  of  $Y$  is a minimal element of  $Y$  if and only if there is no  $b$  in  $Y$  such that  $b \preceq a$ . That is,  $a$  is minimal in  $Y$  with respect to the given ordering.
- **Ordered Pair:** We define an ordered pair  $(a, b)$  to be the set  $\{\{a\}, \{a, b\}\}$ .
- **Ordinal Addition:** First, we set  $A = (\alpha \times \{0\}) \cup (\beta \times \{1\})$ . Then, we define a well-ordering of  $A$  by  $(\alpha, a) <_A (\beta, b) \leftrightarrow (a < b) \vee (a = b \wedge \alpha < \beta)$  which states that ordinal addition is ordered by first comparing the second position of the ordered pairs.
- **Ordinal Exponentiation:** *Ordinal exponentiation*, denoted by  $\alpha^\beta$ , is defined by the recursion:
  - $\alpha^0 = 1$
  - $\alpha^\beta = \alpha^{\beta-1} \cdot \alpha$  if  $\beta$  is a successor ordinal.
  - $\alpha^\beta = \lim_{\gamma < \beta} \alpha^\gamma = \sup\{\alpha^\gamma : \gamma < \beta\}$  if  $\beta$  is a limit ordinal. See Definition 8.18 for reference on supremum.
- **Ordinal Multiplication:** We define ordinal multiplication as iterated addition, that is  $\alpha \cdot \beta = \sum_{\xi < \beta} \alpha$  where  $\alpha \cdot \beta$  denotes  $\beta$  copies of  $\alpha$ .



- **Ordinal Number:** Let  $(X, \preceq)$  be a woset. For  $a \in X$ , let the segment determined by  $a$  be the set  $X_a = \{x \in X : x < a\}$ . An *ordinal* is defined to be a woset  $(X, \preceq)$  such that  $X_a = a$  for  $a \in X$ .
- **Partial Ordering:** A partial ordering of a set  $X$ , denoted by  $\preceq$ , is a binary relation on that set which is reflexive, anti-symmetric, and transitive.
- **Partially Ordered Set (poset):** A partially ordered set is an ordered pair  $(X, \preceq)$ , consisting of a set  $X$  together with a partial ordering  $\preceq$  of  $X$ .
- **Positive Natural Numbers:** The *positive natural numbers*, denoted by  $\mathbb{N}$ , is the set  $\{1, 2, 3, 4, \dots, n, \dots\}$ . When discussing the positive natural numbers, we will start the set at 1. When discussing the natural numbers,  $\mathbb{N}^0$ , we will adopt the usual convention that logicians and computer scientists start the set at 0.
- **Power Set:** For any set  $S$ , the *power set* of  $S$ , denoted by  $\mathcal{P}(S)$  is the set of all subsets of  $S$ . Axiom 5.1 allows us to work with the power set as a set.
- **Reflexive:** Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is reflexive if and only if  $(\forall a \in X)(aRa)$ .
- **Relation:** A *relation* is any subset of the Cartesian Product.
- **Successor Ordinal:** A *successor ordinal* is an ordinal that comes directly after some ordinal in the ordering. For example,  $\beta$  is a successor ordinal if  $\beta = \alpha + 1$  for an ordinal  $\alpha$ .
- **Supremum:** Given a partially ordered set  $(P, \preceq)$ , if a subset  $S \in P$  is bounded above, then a number  $u$  is said to be a *supremum* of  $S$  if:
  1.  $u$  is an upper bound of  $S$
  2.  $v$  is any upper bound of  $S$ , then  $u \preceq v$ .
- **Surjection:** Given two sets  $A$  and  $B$ , a function  $f : A \rightarrow B$  is defined as a *surjection* (onto) if for every element in  $B$ , there is an element in  $A$  that maps to it. In other words,  $(\forall b \in B)(\exists a \in A) f(a) = b$ .

- **Total Ordering (linear ordering):** A total ordering ( $\preceq$ ) of a set  $X$  is a partial ordering of  $X$  whose elements are comparable.
- **Totally Ordered Set (toset):** A pair  $(X, \preceq)$  is a totally ordered set if  $\preceq$  is a total ordering of the set  $X$ .
- **Transfinite Recursion:** *Transfinite recursion* is a way to build up a function or some other mathematical object. It is a technique to construct a function from a base case, then using a recursive technique to define new elements from previously defined ones. We usually define these functions (or objects) on the ordinal numbers. The goal is to assign a value to the function  $f$  for each ordinal up to some  $\alpha$  and that value must be in the given well-ordered set. The recursive part is that we define the next value of the function in terms of the previously defined function values.
- **Transfinite Sequence:** A *transfinite sequence* is a function whose domain is a transfinite ordinal. If  $f$  is a sequence and  $\text{dom}(f) = \alpha$ , we say  $f$  is an  $\alpha$ -sequence. If  $f(\xi) = x_\xi$  for all  $\xi < \alpha$ , we often write  $\langle x_\xi : \xi < \alpha \rangle$  in place of  $f$ . Then, for  $\beta < \alpha$ ,  $\langle x_\xi : \xi < \beta \rangle$  denotes  $f(\beta)$ . This clearly gives a precise meaning to what we generally think of as a (transfinite, perhaps) sequence. The ‘sequences’ of elementary analysis are just the special case of  $\omega$ -sequences, of course; so,  $\{a_n\}_{n=0}^\infty = \langle a_n : n < \omega \rangle$ .
- **Transitive:** Let  $R$  denote any binary relation on a set  $X$ . Then,  $R$  is transitive if and only if  $(\forall a, b, c \in X)[(aRb \wedge bRc) \rightarrow (aRc)]$ .
- **Uncountable:** An infinite set that contains too many elements to be listed or enumerated is considered *uncountable*. This means we are unable to systematically list the elements in the set. Thus, the set  $T$  is *uncountable* if there is no mapping from the positive natural numbers onto  $T$ .
- **Upper Bound:** Given a partially ordered set  $(P, \preceq)$ , an element  $u$  is called an *upper bound* of the subset  $S$  if for every element  $s \in S$ ,  $s \preceq u$ .
- **Well-Founded:** A partially ordered set is well-founded if every nonempty subset has a minimal element.

- **Well-Ordering:** A well-ordering of a set  $X$  is a well-founded, total ordering ( $\preceq$ ) of  $x$ .
- **Well-Ordered Set (woset):** A pair  $(X, \preceq)$  is a *well-ordered set* when  $X$  is a well-founded set together with a total ordering ( $\preceq$ ).

## References

- Bartle, R. G., & Sherbert, D. R. (2011). *Introduction to Real Analysis*. John Wiley and Sons, Inc.
- Bishop, E. (2012). *Foundations of Constructive Analysis*. Ishi Press.
- Cohen, P. (1966). *Set Theory and the Continuum Hypothesis*. NY, NY: W. A. Benjamin, Inc.
- Devlin, K. (1993). *The Joy of Sets, Fundamentals of Contemporary Set Theory*. New York, NY: Springer-Verlag New York, Inc.
- Halmos, P. (2015). *Naïve Set Theory*. Oxford: Benedictan Classics.
- Hilbert, D. (1980). *The Foundations of Geometry, 2nd ed.* Chicago, IL: Open Court.
- Jech, T. (1973). *The Axiom of Choice*. Amsterdam: North-Holland Publishing Company.
- Jech, T. (2003). *Set Theory*. Springer-Verlag Berlin Heidelberg.
- Joyce, D. E. (1996). *Euclid's Elements, Book 1*.  
<https://mathcs.clarku.edu/~djoyce/elements/bookI/bookI.html>.
- Koellner, P. (2010). *Independence and Large Cardinals*. The Stanford Encyclopedia of Philosophy.
- Krantz, S. (2002). *Handbook of Logic and Proof Techniques for Computer Science*. Springer New York.
- Lang, S. (2002). *Algebra*. 175 Fifth Avenue, New York, NY 10010: Springer-Verlag New York, Inc.
- Levy, A. (1979). *Basic Set Theory*. Mineola, NY: Dover Publications, Inc.
- Moore, G. H. (1982). *Zermelo's Axiom of Choice*. New York, NY: Springer.
- Nagel, E., & Newman, J. R. (2001). *Gödel's Proof*. New York: New York University Press.
- Pinter, C. C. (1971, 2014). *A Book of Set Theory*. Dover Press.
- Smullyan, R. (2017). *A Beginner's Further Guide to Mathematical Logic*. 5 Toh Tuck Link, Singapore 596224: World Scientific Publishing Co. Pte. Ltd.
- Stoll, R. R. (2006). *Set Theory*. <https://www.britannica.com/science/set-theory>.
- Zehna, P. W., & Johnson, R. L. (1972). *Elements of Set Theory*. Allyn and Bacon, Inc.