Naive Comprehension, Paradoxes, and the Foundations of Set Theory

Esther Choe, Maya Figelman, Alicia Pité

May 2026

Contents

Abstract

This paper examines how modern set theory developed from naive comprehension to the axiomatic framework of Zermelo–Fraenkel set theory with Choice (ZFC). We first explain George Cantor's naive comprehension principle and show how it leads to Russell's, Cantor's, and Burali-Forti paradoxes. We then describe how ZFC avoids these contradictions by restricting set formation through axioms such as Separation, Power Set, Replacement, and Foundation. Finally, we discuss why the axiomatic approach matters and how modern set theory has evolved since the introduction of ZFC.

1 The origins of set theory

First, let's introduce the definition of a set.

Definition 1. *A* set *is a collection of objects that are referred to as elements.*

Example 1. *The collection of numbers 1,2,3 is written as the set* $\{1, 2, 3\}$.

If S is a set and x is an element of the set S, we write $x \in S$. Inversely, if the element x is not in the set S, we write that $x \notin S$.

Example 2. *The set* $\{1, 2, 3\} = \{2, 3, 1\}$ *because both sets contain the same elements. In other words, here the order of the elements in a set does not matter. However,* $\{1, 2, 3\} \neq \{4, 5, 6\}$, *as they contain different elements, even though they have 3 elements each.*

Definition 2. Naive comprehension *is the principle that every property determines a set. In symbols, for any property* $P(x)$, *naive set theory assumes that the collection* $\{x \mid P(x)\}$ *is a set.*

This notation means: the set of all objects x such that $P(x)$ is true. At first, this principle seems natural. For example, the property "x is an even natural number" gives the set

$$\{x \in \mathbb{N} \mid x \text{ is even}\}$$

The problem is that naive comprehension allows properties to range over all objects, without first restricting them to an already existing set. This unrestricted version is too powerful and leads to contradictions.

2 The foundational crisis in set theory

2.1 Historical background

Modern set theory originated with the work of George Cantor in the 1874 paper: "On a Property of the Collection of All Real Algebraic Numbers," where he introduced a set as a defined collection of objects – this is referred to as *Naive Set Theory*. Cantor used it to manipulate sets to discover interesting properties of collections of numbers and other objects. Notably, he developed a way to compare the sizes of sets by using bijections.

Definition 3. *Bijection A function is called a bijection if it is both injective and surjective.*

- **Injective (one to one):** A function ($f : A \to B$) is injective if for all $(a_1, a_2) \in A, f(a_1) = f(a_2) \implies a_1 = a_2$. This means that two different inputs never equal the same output.
- **Surjective (onto):** A function ($f : A \to B$) is surjective if for every element in B, there exists at least one element in A such that

  $$f(a) = b$$

  So every element in the function's output is accounted for.

  Bijection: A function is a bijection if every element of A maps onto exactly one element of B, and every element of B is matched with only one element of A

By using bijections between sets, Cantor proved that some infinities are countable, such as the set of all natural numbers and the set of all integers. In contrast, the set of all real numbers is an uncountable infinity. These ideas were revolutionary at the time. However, other mathematicians noticed certain flaws within this "naive" definition of a set. Notably, Leopold Kronecker, a prominent mathematician at the time, rejected Cantor's work on infinity and argued that math should be approached from a finite perspective. One of his famous quotes is "Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk" ("God made the integers, all else is the work of man").

In the early 20th century, set theory faced increasing criticism because mathematicians kept uncovering foundational principles that led to paradoxes, such as those posited by Russell, Cantor, and Burali-Forti, which the next section will cover. In response, mathematicians sought to establish rules. The first attempt at axiomatization was by Zermelo in 1908, and it underlined the basic principles of set theory. Zermelo's work was later expanded upon by Abraham Fraenkel and Thoralf Skolem, resulting in what is now known as Zermelo-Fraenkel set theory (ZF), which forms the basis of most modern set theory. Later, the Axiom of Choice was added to Zermelo-Fraenkel, forming ZFC, which remains, to this day, the agreed-upon foundation of mathematics.

2.2 Russell's Paradox

Russell's Paradox is one of the most notable paradoxes that contradicted Cantor's naive set theory. In 1901, in a letter to fellow mathematician Gottlob

Frege, Bertrand Russell noted an inconsistency within the axiomatic system of set theory: it is possible to form the set of all things that satisfy a specific condition.

$$\{x \mid \varphi(x)\}$$

For example, the set of all sets; better known as the universal set, would be

$$U = \{x \mid x = x\}.$$

Theorem 1. *This led Russell to consider what occurred when the condition is exclusivity, meaning R does not contain itself. By making R, the set of all non-self-membered sets or*

$$R = \{x \mid x \notin x\}.$$

The contradiction with this idea of the condition being exclusive arises when we ask whether R is a member of itself. If $R \in R$, then by definition of R, it must be that $R \notin R$. On the other hand, if $R \notin R$, then it satisfies the property that defines R, meaning $R \in R$. This is certainly a contradiction. An easier way to think about this is through a nice analogy, colloquially called the Barber Paradox. Say there is a town where everyone needs to get shaved, and there is a barber who shaves those who do not shave themselves. Then who shaves the barber? Does the barber shave himself? If he does, then by definition he shouldn't, and if he doesn't, then by definition he should. This paradox is one of many that influenced the creation of more structured axiomatic systems.

2.3 Cantor's Paradox

Definition 4. *The power set of a set S is the set of all subsets, denoted by $\mathcal{P}(S)$.*

Example 3. *In the set $\{1, 2, 3\}$ the power set would be:*

$$\{\varnothing, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}.$$

If there are n elements in a set, the power set contains 2^n elements, because for each element of the original set, we have two choices when building a subset (either choose the element or not). So, using the example above, since the set $\{1, 2, 3\}$ contains 3 elements, its power set has 2^3 elements, or 8. The Power Set is always larger than the set itself.

George Cantor dedicated much of his time to set theory and the idea of transfinite numbers. He used one-to-one correspondence, also known as a bijection, to help show that there are different sizes of infinity in sets. Cantor's Paradox was discovered in 1899 and published in 1932. It is motivated by Cantor's theorem:

Theorem 2 (Cantor)**.** *For every set A,*

$$|A| < |\mathcal{P}(A)|.$$

Meaning that the size of set A will always be smaller than the size of its power set.

Proof. The mapping on A into $\mathcal{P}(A)$, which takes a in A into $\{a\}$ in $\mathcal{P}(A)$, is a one-to-one mapping on A into $\mathcal{P}(A)$. Thus, $A \preceq \mathcal{P}(A)$. To prove that $A \prec \mathcal{P}(A)$, we show that the assumption $A \approx \mathcal{P}(A)$ yields a contradiction. Let $f : A \to \mathcal{P}(A)$ demonstrate the assumed similarity of A and $\mathcal{P}(A)$ and consider $A_0 = \{a \in A \mid a \notin f(a)\}$. Since $A_0 \in \mathcal{P}(A)$, there exists a_0 in A such that $f(a_0) = A_0$. Now either $a_0 \in A_0$ or $a_0 \notin A_0$. If $a_0 \in A_0$, then $a_0 \notin f(a_0)$ and hence $a_0 \notin A_0$, which yields a contradiction. Similarly, $a_0 \notin A_0$ implies that $a_0 \in f(a_0)$ or $a_0 \in A_0$, and again a contradiction results. Thus, we have proved that $A \preceq \mathcal{P}(A)$ and not $A \approx \mathcal{P}(A)$. This gives the desired conclusion. □

Thus, the power set is always strictly larger than the original set.
Cantor's Paradox arises when naive comprehension is used to form a universal set, $V = \{x \mid x \text{ is a set}\}$. If such a set existed, then its power set $\mathcal{P}(V)$ would also be a set. Since V is supposed to contain every set, every element of $\mathcal{P}(V)$ would also belong to V. Thus $\mathcal{P}(V) \subseteq V$, which means the Power set of V is a subset of or equal to V, so $|\mathcal{P}(V)| \leq |V|$. However, Cantor's theorem applied to V gives $|V| < |\mathcal{P}(V)|$, which contradicts $|\mathcal{P}(V)| \leq |V|$. Therefore, the assumption that there is a set of all sets is impossible.

2.4 Burali-Forti Paradox

Definition 5. *A set α is* transitive *if whenever $x \in \alpha$, $x \subseteq \alpha$ holds.*

Definition 6. *An* ordinal *is a transitive set that is well-ordered by the membership relation $\in$.*

Ordinal numbers extend on natural numbers by describing the order structure of well-ordered sets.

Example 4. *We write*

$$0 = \varnothing, \qquad 1 = \{0\}, \qquad 2 = \{0, 1\}, \qquad 3 = \{0, 1, 2\}$$

Each ordinal contains all smaller ordinals as its elements. If α is an ordinal, its successor is defined by

$$\alpha + 1 = \alpha \cup \{\alpha\},$$

which is again an ordinal and, by definition, strictly larger than α.

Lemma 1. *Every element of an ordinal is an ordinal.*

Proof. Let α be an ordinal and $\beta \in \alpha$. As α is transitive, every element of β is also an element of α. Thus $\beta \subseteq \alpha$. Because α is well-ordered by $\in$, every subset of α, including β, is also well-ordered by $\in$. Moreover, β is transitive because it forms an initial segment of α, which is an ordinal. Therefore, β is a transitive set well-ordered by $\in$, so β is an ordinal. □

Every nonempty set of ordinals has an $\in$-least element.

Proof. Let A be a nonempty set of ordinals, and choose some $\alpha \in A$. If $\alpha \cap A = \varnothing$, then no element of A belongs to α, so α is the $\in$-least element of A.
If $\alpha \cap A \neq \varnothing$, then $\alpha \cap A$ is a nonempty subset of the ordinal α. Since α is well-ordered by $\in$, there exists some

$$\beta \in \alpha \cap A$$

such that β is $\in$-least in $\alpha \cap A$. We claim that β is $\in$-least in all A. If we suppose otherwise, then there exists $\gamma \in A$ such that $\gamma \in \beta$. Since $\beta \in \alpha$ and α is transitive, we have $\beta \subseteq \alpha$, so $\gamma \in \alpha$. Therefore $\gamma \in \alpha \cap A$, contradicting the statement that β is the least element of $\alpha \cap A$. Hence β is the $\in$-least element of A. □

Theorem 3 (Burali–Forti Paradox)**.** *There is no set of all ordinals.*

Proof. Suppose that there exists a set

$$\Omega = \{\alpha \mid \alpha \text{ is an ordinal}\}.$$

We first show that Ω is well-ordered by $\in$. Say $A \subseteq \Omega$ is nonempty. Since every element of A is an ordinal, A is a nonempty set of ordinals. By the previous lemma, A has an $\in$-least element. Therefore every nonempty subset of Ω has an $\in$-least element, so Ω is well-ordered by $\in$.
Next, we show that Ω is transitive. Let $\alpha \in \Omega$. Then α is an ordinal. By the first lemma, every element of α is also an ordinal. Thus every element of α belongs to Ω, so

$$\alpha \subseteq \Omega.$$

Thus Ω is transitive.
Since Ω is both transitive and well-ordered by $\in$, by definition, Ω is an ordinal. And as Ω was assumed to contain every ordinal, $\Omega \in \Omega$.
However, This is is not possible, because no ordinal can contain itself. If $\Omega \in \Omega$, then $\Omega < \Omega$, which contradicts the strictness of the well-ordered set. Since Ω is an ordinal,

$$\Omega + 1 = \Omega \cup \{\Omega\}$$

is also an ordinal and satisfies $\Omega \in \Omega + 1$, so $\Omega + 1$ is strictly larger than Ω. But this contradicts the assumption that Ω already contains all ordinals.
Therefore, the assumption that there exists a set of all ordinals is false. The collection of all ordinals cannot be a set. □

3 Resolving the crisis via construction of ZFC

3.1 The Axioms of ZFC

All these constructions finally led to Zermelo-Fraenkel set theory, which establishes some Axioms - rules that are unprovable to establish some guidelines for how we approach set theory. The axioms are as follows:

1. **Axiom of Extensionality**: If X and Y have the same elements, then X = Y:

$$\forall X \forall Y \left[(\forall u (u \in X \iff u \in Y)) \implies X = Y \right]$$

2. **Axiom of Pairing**: For all a and b there is a set (a,b) that contains only a and b:

$$\forall a \forall b \exists c \forall x (x \in c \equiv (x = a \lor x = b))$$

3. **Axiom of Separation**: If φ is a property, then for any X there is a set Y = u $\in$ X: φ(u,p) that has all those u $\in$ X that have the property φ

$$\forall X \forall p \exists Y \forall u (u \in Y \equiv (u \in X \land \varphi(u,p)))$$

4. **Axiom of Union**: For any X there exists a set Y = $\bigcup$X, the union of all elements of X.

$$\forall X \exists Y \forall u (u \in Y \equiv \exists z (z \in X \land u \in z)).$$

5. **Axiom of the Power Set**: For any X, there exists a set Y = P(x), the set of all subsets of X.

$$\forall X \exists Y \forall u (u \in Y \equiv u \subseteq X).$$

6. **Axiom of Infinity**: There is an infinite set.

$$\exists S [\varnothing \in S \land \forall x \in S (x \cup \{x\} \in S)]$$

7. **Axiom of Replacement:** If F is a function, the for all X there exists a set $Y = F[x] = (F(x) : x \in X$

$$\forall x \forall y \forall z [\varphi(x,y,p) \land \varphi(x,z,p) \implies y = z] \implies \forall X \exists Y \forall y [y \in Y \equiv (\exists x \in X) \varphi(x,y,p)]$$

8. **Axiom of Foundation:** A set contains a smallest element.

$$\forall x (x \neq \varnothing \implies \exists y \in x (y \cap x = \varnothing))$$

Those first 8 axioms are known as Zermelo-Fraenkel set theory or ZF. To reach ZFC, another axiom, the axiom of choice, was added by Ernest Zermelo in 1904. It essentially states that, given non-empty sets, it is possible to construct a new set by choosing one element from each set.

$$\forall x \in a \exists A(x, y) \implies \exists y \forall x \in a A(x, y(x))$$

However, this axiom resolves and also raises questions about set theory. It is widely accepted because it is incredibly useful in mathematics, as it asserts the existence of objects without providing a way to construct them. In addition, it cannot be proved or disproved with the older axioms of ZF.

3.2 How ZFC Resolves Each Paradox

3.2.1 Naive Comprehension vs. ZFC

Naive comprehension has the core idea that $\{x \mid P(x)\}$ always exists, which means that for any condition $P(x)$, a set can be formed.

3.2.2 Resolving Russell's Paradox

Russell's Paradox depends on that assumption of unrestricted comprehension. As previously stated, Russell applied this idea to the property $x \notin x$, forming

$$R = \{x \mid x \notin x\}.$$

Through the *Axiom of Separation*, creating a set like this is impossible, so it is avoided. Rather than allowing sets to be formed from all objects, Separation only permits subsets of an already existing set. Given a set A and a property $P(x)$, a set can only be created in this manner

$$\{x \in A \mid P(x)\}.$$

As a result, the collection R cannot be constructed in ZFC, since there is no universal set containing every set from which a subset could be separated. The paradox is therefore not considered because it is not an object or set within ZFC. This does not mean it does not exist at all but just not within the axiomatic approach.

3.2.3 Resolving Cantor's Paradox

The reason why Cantor's Theorem does not hold true in restricted comprehension/ZFC is because one cannot construct the set of all sets. Cantor's theorem remains true in ZFC:

$$|A| < |\mathcal{P}(A)|$$

for every set A. What ZFC rejects is the naive step that:

$$V = \{x | x \text{ is a set}\}$$

Because there is no set of all sets, Cantor's theorem cannot be applied to a universal set V.

3.2.4 Resolving the Burali–Forti Paradox

The naive assumption that any collection that one can define with a certain property, in this case, the set of all ordinals, leads to a contradiction. Similarly to Russell's Paradox, ZFC does not allow the Burali-Forti Paradox to be considered.

The *Axiom of Separation*, as previously mentioned, does not allow any arbitrary property to determine a set, only allowing the formation of subsets from an already existing set. Therefore, within ZFC, the Burali-Forti Paradox no longer arises by treating the collection of all ordinals as a set and then asking whether it is a member of itself.

Another important axiom contributing to this resolution is the *Axiom of Foundation* (Regularity), which ensures that no set is a member of itself. While the contradiction in the Burali-Forti Paradox does not rely solely on self-membership, Foundation reinforces the well-founded nature of the set-theoretic universe and aligns with the fact that ordinals are strictly well-ordered by $\in$.

Finally, the cumulative hierarchy of sets in ZFC provides a structural explanation to resolve the paradox. In summary, ZFC resolves the Burali-Forti Paradox by restricting set formation and eliminating the possibility of a universal set. The collection of all ordinals is not a set, and thus the contradiction derived in naive set theory is avoided entirely within ZFC.

4 Summary and Extensions of Set Theory

4.1 Summary

We explain the concepts of naive set theory and naive comprehension. In naive understanding, it is assumed that for any condition $P(x)$, a set of the form $\{x \mid P(x)\}$ always exists.

We then examined the contradictions of naive set theory through Russell's Paradox, Cantor's Paradox, and Burali-Forti's Paradox.

To resolve these issues, ZFC set theory was introduced, in which the formation of sets is restricted by axioms. These restrictions prevent the construction of problematic sets such as the set of all sets or the set of all ordinals.

The transition from naive set theory to axiomatic set theory shows that restricting the formation of sets is essential to maintain consistency in mathematics.

4.2 The Relevance of Axiomatic Set Theory

By constructing axioms and taking an axiomatic approach in set theory, we can prevent paradoxes by avoiding inconsistencies. Some ways in which inconsistencies can be prevented is by clarifying what it means for something to "exist" in math. It also lays the foundation for comparisons between mathematical worlds, depending on the axiom systems, such as Euclidean Geometry.

Thanks to the axioms, it turns set theory from a vague language of "collections" into a disciplined framework for understanding the theorems and assumptions of mathematics.

4.3 Extensions and Critiques of ZFC, Towards NBG

Although ZFC remains the standard foundation of most modern math, mathematicians have still continued to study other systems and extensions of set theory. One reason for this is that ZFC does not treat collections that are too large to be sets as formal objects within the theory, such as the set of all sets, or the set of all ordinals. This led mathematicians Von Neumann, Bernays, and Gödel to create the aptly named Von Neumann–Bernays-Gödel set theory or NBG, which is an extension of ZFC that includes classes into it formally. In NBG, there is a distinction between sets and proper classes. Sets are objects that may belong to other collections, while proper classes are collections too large to be members of anything. This allows the concept itself to be preserved while avoiding the logical contradictions that arose in Naive Set Theory.

References

[1] Joan Bagaria. Set theory. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2023 edition, 2023.

[2] Robert R. Stoll. *Set Theory and Logic*. W. H. Freeman and Company, San Francisco, 1963.

[3] Robert R. Stoll. *Set Theory and Logic*. Dover Publications, New York, 1979.

[4] Eric W. Weisstein. Zermelo-fraenkel axioms. In *MathWorld—A Wolfram Web Resource*. Wolfram Research.

[5] Wikipedia contributors. Controversy over Cantor's theory — Wikipedia, The Free Encyclopedia, 2026.