ZFC SET THEORY AND THE BANACH–TARSKI PARADOX

BEN BABCOCK

1. Introduction

Few concepts can claim to be as fundamental to the modern practice of mathematics as sets. From its origins in the nineteenth century to its rigorous axiomatization in the 1900s, set theory has provided insight into the relationships among the integers, the natural numbers, and the reals. It has tamed the controversial concept of infinity. Mathematicians in every field use sets daily, both implicitly and explicitly. Modern axiomatic set theory as codified by Zermelo and Fraenkel is a system of logical axioms that avoids the paradoxes of earlier naive and axiomatic attempts at such a theory. Yet with the addition of the axiom of choice, set theory can lead to results that, while not contradictory, are counter-intuitive to our conceptions of geometry and physical space.

We will outline the development of set theory from its naive origins in Cantor to attempts at axiomatization by Frege and the successful axiomatization by Zermelo, Fraenkel, et al. Our particular focus will be on the axiom of choice. Understanding its unique place in set theory is crucial to the final section of this paper, where we will present the Banach–Tarski paradox. This paradox is an example of the widespread and immediate implications of the axiom of choice. For brevity, we will follow the conventions of abbreviating the axiom of choice to AC, using ZF to refer to ZermeloFraenkel set theory without AC, and using ZFC to refer to ZF with AC.

The author would like to thank his supervisor, Andrew J. Dean, for providing references, lecture notes, and support, and the honour seminar course coordinator, Adam Van Tuyl.

2. The Origins of Axiomatic Set Theory

At its most basic, a *set* is a collection of objects, which we refer to as *elements* of the set. We say two sets are *equal* if and only if they have the same elements. These two concepts provide the basic binary predicates of set theory: membership, which we denote as $x \in X$ for an element x in a given set X; and equality of any two sets A, B, which we denote by $A = B$. This much naive set theory and axiomatic set theory have in common. They also share the familiar operations of union, intersection, etc. Whereas naive set theory takes these operations, and indeed the existence of sets in general, as given, axiomatic set theory admits only those operations and sets that can be derived from one or more precisely-formulated axioms.

When using sets in other areas of mathematics, naive set theory often suffices. Yet this is acceptable only because mathematicians are aware that a more formal, rigorous treatment of sets exists and can be employed where necessary. Naive set theory did not enjoy this sort of tolerance in the nineteenth century. The reasons for this lie as much in the philosophical schools of mathematical thought during this period, particularly in Germany, as in the construction of naive set theory itself. Dedekind and Riemann favoured set theory because it appealed to their abstract, general approach in formulating problems. Dedekind and Riemann "consistently attempted to frame mathematical theories within the most general appropriate setting," and thus the general nature of sets appealed to them [5, p. 31]. A competing school, which drew proponents such as Weierstrass and Kronecker, insisted upon developing mathematics "from purely arithmetical notions" [5, p. 35]. This schism influenced Cantor's relationship with those two analysts as he began to investigate analysis and mathematical infinity using the abstract notions of sets.

A full treatment of naive set theory is, of course, beyond the scope of this paper. For a comprehensive history of set theory, see [5], which emphasizes the tendency of modern historians of mathematics to ascribe set theory's origins almost exclusively to Cantor. However, for our purpose, we will focus mostly on Cantor's contributions before we turn to Frege's axiomatic set theory.

2.1. **Cantor and Naive Set Theory.** In the 1870s, Georg Cantor found, in his study of analysis, that it was useful to regard the real numbers and natural numbers as infinite sets. Infinity is a concept that has long been controversial among mathematicians and philosophers. Hence, while finite sets were intuitive and comprehensible, infinite sets would, at first, be somewhat intractable in the framework of naive set theory. Cantor in fact demonstrated that among the infinite sets, some were "larger" in size than others. His work on the cardinality of infinite sets laid the groundwork that later motivated other mathematicians to pursue an axiomatic set theory.

Cantor investigated the cardinality of the real and natural numbers, specifically, whether the set of all reals could be placed in a one-to-one correspondence with the set of all naturals. He published two proofs that this was not possible: one in 1874 [2], and the other in 1891 [3], which is now known as his diagonalization argument. Although his work was not motivated directly by open problems in analysis, Cantor's discoveries are now seen as definitive moments in both analysis and the development of set theory [5]. In particular, Cantor established the notion of equipollence:

Definition 2.1. Two sets A and B are *equipollent* (or, equivalently, they have the same cardinality) if and only if there exists a bijection between them.

Thus, the reals and the naturals are not equipollent. In proving this result, Cantor showed that infinite sets have at least two different cardinalities: those sets that can be put on a one-to-one correspondence with the natural numbers are *countable* (Cantor called them *denumerable*), and those that cannot are *uncountable* (or *non-denumerable*). From this arises the question of whether there are any sets whose cardinality lies between the cardinality of these two sets. This is Cantor's *continuum hypothesis*, and his efforts to prove this conjecture motivated his study of transfinite numbers.

Conjecture 2.2 (Generalized Continuum Hypothesis (GCH)). *Given an infinite set A, there exists no set whose cardinality lies between that of A and that of its power set.*

Gödel proved that this stronger form of the continuum hypothesis is consistent with the axioms of ZF set theory; that is, it introduces no new contradictions. However, like the axiom of choice, GCH is also independent of ZF, so the axioms of ZF cannot be used to prove or disprove the hypothesis. In fact, Sierpinski proved that GCH implies the axiom of choice (although the converse is not true). Thus, there are no models of ZF set theory where the axiom of choice is false and GCH is true [11].

2.2. **Axiomatization of Set Theory.** As set theory rose to prominence in the 1880s and 1890s, it became apparent that the intuitive approach to describing sets leads to contradictions. Cantor's Theorem guarantees that the cardinality of any set's power set is strictly greater than the cardinality of the set itself. Cantor realized that this meant describing a set such as "the set of all infinite cardinals" was contradictory: given such a set A with cardinality α, its power set would have cardinality β, such that $\beta > \alpha$. Then we would have $\beta \in A$ and $\beta \notin A$. (This was Cantor's initial formulation of the paradox. See [5, p. 291] for a later, more rigorous argument.)

This paradox highlights why we cannot rely on verbal language to describe sets. Verbal language is imprecise, and while we can easily write the phrase "the set of all infinite cardinals," we quickly see that the set this phrase describes is paradoxical. Hence, relying on verbal arguments and definitions as a basis for set theory admits too much ambiguity and results in these contradictory constructions. Cantor and his contemporaries saw that a more formal approach to set theory was needed if it was to be relied upon as a foundation of mathematics. Modern set theory thus began with a return to the use of axioms, the use of which goes back at least to ancient Greece and Euclid.

Definition 2.3. An *axiom* is a statement assumed to be true, from which further statements are proved.

Axioms alone do not guarantee superior precision, for of course verbally-specified axioms are just as vulnerable to semantical paradoxes as an intuitive description of sets. Fortunately, mathematicians had a powerful new tool at their disposal in the form of mathematical logic. Through predicate logic it is possible to express statements with more precision than verbal language allows. Hilbert, Frege, and Russell were all proponents of reformulating mathematics in this way.

The central danger to any axiomatization of set theory then became the prospect of specifying axioms that *admitted too much.* Indeed, even with predicate logic available to them, this is exactly what happened. Gottlob Frege published an axiomatization of set theory. In part, this presented axioms that had already been understood intuitively as precise, logical propositions. One such axiom, the axiom of abstraction, would prove to be too generous.

Definition 2.4 (Axiom of Abstraction [11, p. 5])**.** Given any property, there exists a set whose members are just those entities having that property.

The axiom of abstraction makes sense to anyone accustomed to viewing sets intuitively. After all, sets are collections of objects, so it is reasonable to attempt grouping objects into sets based on a common property. Nevertheless, when expressed in this general way, we can specify properties that immediately lead to contradictions. In 1903 Russell published one such paradox that now bears his name and stands as the paradigmatical set-theoretic paradox [11]. Let the set R be "the set of all sets that are not members of themselves" [4], i.e.,

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

The contradiction occurs when we consider the membership of R. If R is not a member of itself, then by definition it *is* a member of itself. In terms of propositional logic, we can formulate the axiom of abstraction as:

$$(\exists y)(\forall x)\,(x \in y \iff \varphi(x))\,, \tag{2.1}$$

where $\varphi(x)$ is a formula that specifies our given property. To obtain our paradox, let $\varphi(x)$ be the statement

$$\neg(x \in x),$$

and now let $x = y$. Then we can simplify 2.1 to:

$$y \in y \iff \neg(y \in y),$$

which is a contradiction [11, p. 6].

Compared to the other paradoxes that preceded it, Russell's paradox is significant because it can be derived in this way using propositional logic. More than just a result of semantics, Russell's paradox revealed a fundamental flaw in Frege's formulation of axiomatic set theory, one which Frege was at a loss to correct [5, p. 308]. It was unclear whether a reformulation of set theory could resolve the problem or if, as Russell thought, logic itself needed reform [5, p. 330]. Above all, Russell's paradox underscored the delicacy required in the task of developing a formal axiomatic set theory. We will now examine the axiomatization, first published by Ernst Zermelo in 1908 and refined throughout the next three decades, that has now become the standard version of modern set theory. Even so, as we will see in later sections, ZermeloFraenkel set theory has its own unusual, though not contradictory, twists of logic.

3. ZermeloFraenkel Set Theory

As with many other presentations of ZF, we will present the axioms incrementally and in roughly chronological order. Some of the first axioms developed by Zermelo, notably the axiom schema of separation, were made obsolete by later axioms (in this case, the axiom schema of replacement). Nevertheless, presenting the axioms in this way emphasizes how ZF attempts, by the way it is constructed, to limit the existence of sets so as to curtail contradictions. Gödel's second incompleteness theorem implies that we cannot use ZF to prove that ZF is consistent unless it is actually inconsistent, and because ZF provides the foundation for so much of modern mathematics, a proof of its consistency is not likely to be forthcoming [5, p. 345]. However, ZF does preclude the paradoxes known to plague naive set theory, especially Russell's paradox.

We will not describe every axiom of ZF here, because our goal is not to provide a comprehensive development of set theory. For example, the axiom of extensionality merely reformulates as a logical proposition the idea that two sets are equal if and only if they have the same elements. In particular, we shall focus on the axiom schema of separation, the pairing axiom, the sum axiom, the axiom of regularity, and the axiom schema of replacement. These axioms exemplify the incremental approach to admitting the existence of sets and preventing paradoxes. For convenience, an appendix lists the axioms of ZF, including those omitted from our discussion.

The propositional forms of these axioms are from [11]. We shall follow the convention that uppercase letters denote sets unless otherwise stated, and lowercase letters denote variables that may or may not be sets. Sometimes all objects within the universe of discourse of ZF are considered to be sets, in which case the distinction is unnecessary; other models of ZF admit *ur-elements* (also called *atoms* or *individuals*), objects that are not themselves sets but can be elements of sets. Whether we admit ur-elements is not germane to the scope of this paper, as our discussion will apply to ZF and ZFC in general. See [7] for more on how the axiom of choice interacts with ZF with ur-elements.

Remark 3.1 (The Existence of the Empty Set)**.** Different formulations of ZF provide different justification for the existence of a set that contains no elements. Some specify an axiom of the empty set; others derive its existence from the axiom schemata of separation or replacement, etc. Once such a set exists, the axiom of extensionality guarantees its uniqueness, so we can call it *the empty set* and denote it with $\emptyset$. We will follow [11] and define a *set* to be anything which has elements or which is the empty set.

Definition 3.2 ([11, p. 19])**.** The object y is a *set* $\iff (\exists x)(x \in y \vee y = \emptyset)$.

3.1. **Axiom Schema of Separation.** Similar to the axiom of abstraction, the axiom schema of separation allows us to construct sets whose elements share a given property:

$$(\exists B)(\forall x)\left(x \in B \iff x \in A \wedge \varphi(x)\right).$$

Again, $\varphi(x)$ is a logical proposition equivalent to our property. We call this an axiom *schema* because $\varphi(x)$ is left unspecified. Thus, this actually defines for us a framework from which we construct infinitely-many axioms. Consequently, the axiom schema of separation is very powerful. Immediately it gives us the basic set operations of intersection and set-theoretic difference (though not union, which requires the pairing axiom and the sum axiom).

In addition to specifying $\varphi(x)$, notice that we also require the elements of B to be elements of a pre-existing set, A. This restriction prevents us from encountering the problems, like Russell's paradox, caused by the axiom of abstraction. Essentially, the axiom schema of separation allows us to construct subsets of a given set A. This follows the cautious approach to the development of sets in ZF; that is, we allow only those sets based upon pre-existing sets.

3.2. **The Pairing and Sum Axioms.** So far we can only create subsets. The pairing axiom says that, given any two variables x, y, there exists a set containing only x and y:

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

This axiom allows us to define ordered pairs, and hence relations and functions, within the framework of set theory.

Definition 3.3 (Due to Kuratowski [11, p. 32])**.** Given x and y, the *ordered pair* (x, y) is defined as the set

$$(x, y) = \big\{\{x\}, \{x, y\}\big\}.$$

Definition 3.4. A set R is a *relation* if and only if every element of R is an ordered pair.

Then given two sets A, B, a *function* f from A to B is a relation that maps every every element of A to exactly one element of B. Thus, f is simply set of ordered pairs (x, y), where $x \in A$, $y \in B$, and we write $f(x) = y$ if and only if $(x, y) \in f$.

Definition 3.5. Given two sets A, B, let f be a function from A to B. Then f is:

- *injective* $\iff (\forall x)(\forall y)(f(x) = f(y) \Rightarrow x = y)$,
- *surjective* $\iff (\forall b)\,(b \in B \implies (\exists a)(a \in A \wedge f(a) = b))$, and
- *bijective* if and only if f is both injective and surjective.

Referring back to Definition 2.1, we now have a well-defined idea of a bijection. Of course, our definition, framed as it is within ZF set theory and propositional logic, is anachronistic to how Cantor used the term, but for our purposes there is no difference. We will rely on the idea of a function as a set of ordered pairs when we examine the necessity of axiom of choice in section 4.

Given a family of sets, A, the sum axiom allows us to assert the existence of a union over that family:

$$(\exists C)(\forall x)\,(x \in C \iff (\exists B)(x \in B \wedge B \in A))\,.$$

We denote the *union* or *sum* of A by $\bigcup A$, and it is the set of all elements belonging to some element of A. Applying the sum axiom and our definitions, we can then derive the expected results, such as the fact that $\bigcup \emptyset = \emptyset$, and establish the corresponding idea of the intersection of a family of sets.

Remark 3.6. The sum axiom is also known as the *union axiom*. Some formulations of ZF include a different axiom under that name, which states that given any two sets A and B, there exists another set, denoted $A \cup B$, whose elements are all the elements of A and B, i.e.:

$$x \in A \cup B \iff (x \in A \vee x \in B).$$

However, this is just a special case of the sum axiom: $A \cup B = \bigcup\{A, B\}$, where the existence of $\{A, B\}$ is given by the pairing axiom.

The redundancy of this alternative axiom of union demonstrates the value in a judicious selection of the axioms that form the basis of our set theory. With the correct formulation, we eliminate extra axioms, which is more elegant from the standpoint of mathematical logic. Occasionally, one might only need to work with certain sets, not the full ZF set theory, and therefore find the union axiom sufficient. In general, the pairing and sum axioms are much more powerful. As we will see, the former is also redundant once Zermelo's theory meets Fraenkel's.

3.3. **Axiom of Regularity.** There remains one last axiom from Zermelo's formulation that we must not overlook. Russell's paradox proposes we consider the set of all sets that are not members of themselves. It is natural to ask, then, if there exists any set that *is* a member of itself. Almost every set that comes to mind cannot fit this definition, and potential concepts that do, such as the set of all sets, are objects we wish to avoid. Hence, we need an axiom that prohibits a set from having itself as an element.

We cannot simply use the statement

$$A \notin A,$$

for this does not prevent cycles of sets that have each other as elements, i.e.,

$$A \in B \wedge B \in C \wedge C \in A.$$

Thus, the axiom we adopt must be stronger than preventing the single contradictory example from Russell's paradox; it must prevent any cycle of arbitrary length. To ensure this, we require every non-empty set A to have an element x such that x and A have no elements in common:

$$A \neq \emptyset \Rightarrow (\exists x)\left[x \in A \wedge (\forall y)(y \in x \Rightarrow y \notin A)\right].$$

Von Neumann provided an equivalent definition that has the same result [1, 5]. The above formulation is due to Zermelo, who first incorporated it into ZF as the axiom of foundation in his 1930 paper [5, p. 374].

From a practical perspective, the axiom of regularity is not that useful. Since the paradoxical sets it prevents do not tend to arise in the application of set theory, "it has no single known consequence for actual mathematical work outside set theory" [5, p. 377]. Nevertheless, this axiom is worth mentioning because of its importance to the framework of ZF as a whole. It, along with the other axioms of ZF, forms an axiomatic set theory that prevents the known paradoxes which demolished previous formulations. Moreover, the axiom of regularity demonstrates that not every formulation of set theory need be equivalent or even similar in conception. For example, Quine's New Foundations allows the existence of a universal set, which has itself as a member and thus contradicts the axiom of regularity [6].

3.4. **Axiom Schema of Replacement.** Also known as the axiom of substitution, this axiom is commonly attributed to Fraenkel and cited as the reason Zermelo set theory is now ZermeloFraenkel set theory. However, it was proposed contemporaneously by both Skolem and von Neumann [1, p. 22]. Zermelo appended Fraenkel's name to his set theory in his 1930 paper ostensibly because of his incorporation of Fraenkel's axiom of replacement into the theory. However, Ferreirós postulates the credit had more to do with Fraenkel's consistent use of Zermelo's set theory even as other mathematicians continued to rely upon Russell's theory of types or explore competing alternatives. Indeed, even though Fraenkel developed the axiom of replacement, its restrictiveness made him reluctant to recommend its adoption in general set theory. It was Von Neumann who, in 1928, refined Fraenkel's formulation of the axiom and proposed it should be incorporated into the rest of ZF [5, p. 372].

Von Neumann was drawn to the axiom of replacement while studying his eponymous ordinals. His formal proof that every well-ordered set has a unique corresponding ordinal requires this axiom [5]. Similarly, Fraenkel was interested in the existence of sets of the form

$$P = \{A, \mathcal{P}(A), \mathcal{P}(\mathcal{P}(A)), \dots\},$$

where A is an arbitrary infinite set and $\mathcal{P}(A)$ is its power set. In 1921, he found that Zermelo's original axioms were insufficient for proving that such sets existed [1, p. 21] and in fact required the axiom of replacement:

$$(\forall x)(\forall y)(\forall z)\,(x \in A \wedge \varphi(x,y) \wedge \varphi(x,z) \Rightarrow y = z)$$
$$\Rightarrow (\exists B)(\forall y)\,(y \in B \iff (\exists x)\,(x \in A \wedge \varphi(x,y)))\,.$$

As with the axiom of separation, the axiom of replacement is actually an axiom *schema*. Given a set A, we define a formula φ such that, for every $x \in A$, we guarantee the existence of at most one y satisfying the formula. If that is the case, then we assert the existence of a set of these y's, which we denote B. We also require φ to be *functional in* x; i.e., the set $F = \{(x, y) \mid \varphi(x, y)\}$ is a function. For example, let A be a collection of sets, and let $\varphi(X, Y)$ be "$Y = \mathcal{P}(X)$." Using the axiom schema of replacement, we *replace* each element of A with its power set Y, thereby obtaining the set B [4].

Thus, this replacement quality makes the axiom schema "stronger" than the axiom schema of separation in the sense that the latter only allows us to construct subsets of existing sets, whereas with the former we can use a given set A to construct a corresponding set B whose elements meet the conditions above. Indeed, once this axiom is incorporated into ZF, both the axiom schema of separation and the pairing axiom become redundant. We shall prove the former; for a proof of the latter see [11, p. 237].

Proposition 3.7. *The axiom schema of replacement implies the axiom schema of separation.*

Proof. Let $\varphi(x, y)$ be "$x = y \wedge \psi(y)$," where $\psi(y)$ is a proposition of the type satisfying the axiom schema of separation. Then, by the axiom schema of replacement, we have:

$$(\forall x)(\forall y)(\forall z)\,(x \in A \wedge (x = y \wedge \psi(y)) \wedge (x = z \wedge \psi(z)) \Rightarrow y = z)$$
$$\Rightarrow (\exists B)(\forall y)\,(y \in B \iff (\exists x)\,(x \in A \wedge (x = y \wedge \psi(y))))\,.$$

Since $x = y \wedge y = z$ implies $y = z$, the antecedent is true, so we have

$$(\exists B)(\forall y)\,(y \in B \iff (\exists x)\,(x \in A \wedge (x = y \wedge \psi(y))))\,.$$

If no such x exists, then obviously $y \notin A$. We need only consider, then, those cases where such an $x \in A$ exists. Then since $x = y$, the above is equivalent to

$$(\exists B)(\forall y)\,(y \in B \iff y \in A \wedge \psi(y))\,,$$

which is the axiom schema of separation. □

Despite its redundancy, the axiom schema of separation is still useful and interesting. Of course, it is essential to any development of ZF from an historical perspective. Also, we can develop a great deal of ZF without resorting to the axiom schema of replacement, and so depending on one's purposes, it might be more desirable to use the simpler of the two schemata.

ZermeloFraenkel set theory as developed thus far (including those axioms we have omitted from our discussion here) is sufficient to prevent the known paradoxes to which previous systems were vulnerable and to prove the existence of most sets of interest to us. However, there is one additional axiom notorious enough that it is often regarded separately. We now examine this axiom of choice in the context of its history and controversy.

4. The Axiom of Choice

There are many variations on the formulation of this axiom. Indeed, entire books have been devoted to cataloging those statements equivalent to the axiom of choice (see [9] for one). The fact that the axiom of choice is equivalent to, or implies, so many other important results in mathematics is one of the reasons its controversial nature has made it so notorious: some mathematicians do not like using it, but they use it anyway, because not using it would mean rejecting those useful results.

Definition 4.1 (Axiom of Choice [11, p. 239])**.** For any set A there is a function f such that for any non-empty subset B of A, $f(B) \in B$.

We call f the *choice function for A*. The axiom of choice guarantees the existence of a choice function for any set. However, it does not offer any guidance in constructing the choice function; we shall see that this non-constructivity is one of the reasons for the controversy about the axiom.

Proposition 4.2. *Let M be a finite set. Then we can construct a choice function for M using only the axioms of ZF.*

Proof. If $M = \{m\}$, then the only non-empty subset of M is M itself. Then the choice function is obvious:

$$f = \{(M, m)\},$$

where we have written f as a set of ordered pairs as defined in section 3.2. Suppose that this is true for sets of up to n elements. Let X be a set of $n + 1$ elements. Then $X = Y \cup \{x\}$, where Y is a set of n elements. By induction, Y has a choice function, g. Similarly, as a singleton set, $\{x\}$ also has a choice function, $\{(\{x\}, x)\}$. So we can define the choice function for X by

$$f = g \cup \{(\{x\}, x)\}.$$

□

This proof cannot be extended to the infinite case, since we cannot do induction over uncountably infinite sets. If one only needs the guarantee of a choice function for *countable* sets, one can take the weaker "countable axiom of choice" [7, p. 20]. Proposition 4.2 should not be confused with the axiom of choice of finite sets, which restricts AC to collections of finite sets [7, p. 107].

Firstly, we will explore the adoption of AC in an historical context. Secondly, we will mention some results that are equivalent to AC. Finally, we will discuss AC's consistency with and independence from the rest of ZF and why AC has historically been so controversial. This will provide us with the knowledge required for an examination of the Banach–Tarski paradox.

4.1. **ZF and AC Prior to ZFC.**

Historically the formal introduction of AC into mathematics is attributed to Zermelo, who used it in his 1904 proof of the well-ordering theorem. He acknowledged that his proof relied upon assuming AC to be true. However, he was not the first to make use of the AC, for mathematicians had been using it implicitly for decades, and Zermelo acknowledged this fact, using the axiom's apparent self-evidence to justify his assumption. Indeed, use of AC pervaded mathematics, but this did not prevent other mathematicians from questioning its legitimacy. Ferreirós frames the controversy around Zermelo's well-ordered theorem and AC as a symptom of the larger question of abstractness in mathematics [5, p. 313].

Since Hilbert's presentation of his famous list of problems at the 1900 International Congress of Mathematicians, the continuum hypothesis and, by extension, the question of well-ordering was once again a popular topic of study and debate. Notably, at the 1904 Congress, prior to Zermelo's publication of his proof of the well-ordering theorem, König presented a proof that the continuum hypothesis was false. The proof was later shown to rely upon a flawed lemma, but its reception vis-à-vis Zermelo's proof reveals the philosophical divide present among mathematicians at the turn of the century. Many, Cantor reportedly among them, expressed approval when König first presented his proof at the International Congress. His proof purported to establish a contradiction when one assumed the reals could be well-ordered. When Zermelo's proof of well-orderering for *any* set, including the reals, was published later that year, it was met with criticism and hostility.

Definition 4.3. A set S is *well-ordered* if there exists a relation $\prec$ on S such that

(1) For any $a, b \in S$, exactly one of the following statements is true: $a \prec b, b \prec a, a = b$.
(2) For any $a, b, c \in S$, if $a \prec b$ and $b \prec c$, then $a \prec c$.
(3) For any non-empty subset $T \subseteq S$, there exists a $t \in T$ such that for all $x \in T$, either $t = x$ or $t \prec x$.

We say that $\prec$ is a *well-order on* S or that $\prec$ well-orders S.

Theorem 4.4 (Well-Ordering of Sets [11, p. 242])**.** *Every set can be well-ordered; that is, for every set A there is a relation R such that R well-orders A.*

Although well-ordering is a nice property to have on a set, no actual example of such an ordering for $\mathbb{R}$ is evident. Hence, some mathematicians balked at accepting a well-ordering theorem for any set. Since Zermelo had introduced AC in the context of proving this theorem, it became associated with the controversy surrounding well-ordering, despite the fact that, as Zermelo claimed in his paper, mathematicians had unconsciously used AC elsewhere. However, as we shall see in section 4.2, AC is actually equivalent to well-ordering, so its involvement in this controversy was unavoidable.

The axiom of choice became a major point of contention between constructivists, such as Borel and Lebesgue, and formalists, particularly Hilbert. These two schools of mathematical thought were divided at the most fundamental philosophical level. The former held that "a mathematical notion . . . does not truly *exist* unless it has been finitely defined by means of characteristic properties." The latter, on the other hand, viewed "'existence' [as] non-contradictoriness: whenever an axiom system is consistent, we are entitled to regard the set of objects it describes as existing" [5, p. 315-16]. The well-ordering theorem and AC are inherently non-constructive: if we take them to be true, then a well-ordering of $\mathbb{R}$ "exists," but we cannot say it exists in the sense of the constructivists.

In the subsequent years, mathematicians on both sides, as well as those from competing, third party philosophies, would use their prestige and influence to lobby for their philosophy's interpretation of mathematics. That AC is commonly accepted as true today is as much a testament to the influence of Hilbert and other abstract mathematicians as it is an endorsement of the axiom's objective truth, which, of course, cannot be established. Consequently, both constructivism and formalism continue to play roles in mathematics today, with AC the dividing line in the sand.

4.2. **Equivalents of AC.** The undecidability of AC, which we will discuss further in section 4.3, means that it is possible to construct models of ZF in which the axiom of choice is false. For example, Mycielski introduced the axiom of determinacy in the context of its contradiction of AC [7]. The study of such models is interesting from a set-theoretic perspective, because it provides insight into how we rely upon AC. Mathematicians were using AC long before Zermelo formulated it and began advocating for its adoption as an axiom; contemporary mathematicians operating outside the scope of set theory often use AC implicitly, with good reason. For an extensive list of equivalents, see [9]. We will now state some of the most best-known equivalents.

Definition 4.5. Given a non-empty set X with a partial order $\leq$, a subset $A \subseteq X$ is a *chain* in X if and only if it is totally ordered under $\leq$. A chain is *maximal* if and only if for all chains S in X, $A \not\subset S$.

The axiom of choice is equivalent to the following:

(1) The well-ordering theorem (Theorem 4.4).
(2) The Cartesian product of any collection of non-empty sets is non-empty.
(3) **Numeration Theorem.** For any set A there is an ordinal α such that A is equipollent with α.
(4) **Hausdorff's Maximal Principle.** If X is a non-empty set with a partial order $\leq$, then every chain in X is contained in some maximal chain.
(5) **Zorn's Lemma.** If X is a non-empty set with a partial order $\leq$ such that every chain in X is bounded above, then X contains a maximal element.
(6) **Law of Trichotomy.** For all sets A, B, the cardinality of A and the cardinality of B are comparable. That is, exactly one of the following holds:
    - There exists a bijection between A and B (A and B are equipollent).
    - There exists an injection $f : A \to B$, and there is no injection from B to A (A is equipollent with a proper subset of B).
    - There exists an injection $g : B \to A$, and there is no injection from A to B (B is equipollent with a proper subset of A).

Mathematicians use these equivalents throught mathematics, in ring theory, topology, linear algebra, etc. That these major theorems are equivalent to AC should demonstrate its widespread influence beyond set theory. In particular, the Trichotomy Law is often taken for granted; owing to our intuitive understanding of finite sets, it is a statement that seems obvious. However, without the axiom of choice (or equivalent), it is not at all obvious for infinite sets.

Theorem 4.6. *The well-ordering theorem implies AC.*

Proof. Given a set A with a non-empty subset B, the well-ordering theorem implies that there exists a well-order on A such that B has a least element, b. Then we can define a choice function $f : \mathcal{P}(A) \setminus \{\emptyset\} \to A$ by $f(B) = b$. That is, we send every non-empty subset to its least element. □

For proofs of the equivalence of AC to the law of trichotomy and Zorn's Lemma, see [4, 7, 11]. It is sufficient for our purposes to observe that AC is equivalent to many propositions that we tend to assume for their usefulness in other areas of mathematics. Refusing to use AC means rejecting those propositions as well. Hoping to find evidence for the objective truth or falsity of AC within the framework of ZF itself, many mathematicians turned to the consistency and independence of AC following its introduction by Zermelo. As we will now see, the results of those investigations have only confirmed AC's status as apart from ZF.

4.3. **Consistency and Independence of AC.** In logic, a theory is *consistent* when it does not contain a contradiction. When we say that the axiom of choice is *consistent with ZF*, we mean that if ZF is consistent, then ZFC is also consistent. Thus, using AC does not lead to any more contradictions than using ZF without AC would. This is essential to our goal of creating a formally axiomatized system that does not admit the existence of paradoxical sets. We do not know whether ZF itself admits any such sets, of course, as one consequence of Gödel's second incompleteness theorem is that we cannot, within ZF, prove its own consistency. However, ZF does prevent the known paradoxes to which other systems were vulnerable. It has now been used as a foundation of most of mathematics for a century, and so far no one has been able to create a contradiction with the axioms of ZF. So we do not know, with a rigorous certainty, that ZF is consistent, but it is the most reliable system available to us at present.

Gödel proved that AC is consistent with ZF in 1938 [7]. This was a landmark result in mathematical logic, for it further diminished worries that ZF and comparable systems of set theory were not as "safe" to use as foundations for further theories as competing logical systems, such as Russell's type theory [5, p. 382]. Moreover, Gödel's approach to the problem would be replicated to settle future consistency proofs (his proof of the consistency of AC was in fact a sequel to his consistency proof for the continuum hypothesis). Gödel, influenced by Russell's type theory, defined a "universe" of so-called *constructible* sets and showed that they satisfied the axioms of ZF, making it a model for ZF. He then proved that AC holds in the constructible universe, thus showing that ZF and AC are consistent. For a more comprehensive explanation of how to develop the constructible universe, refer to [7, ch. 3].

The question of independence remained open considerably longer; Cohen answered it in the affirmative in 1964 [7]. When we say that the axiom of choice is *independent* of ZF, we mean that we cannot prove that AC follows from the axioms of ZF alone. Prior to Cohen's proof, the unusual nature of the axiom of choice and decades of experience using it had led most mathematicians to believe that AC was independent. Cohen's proof was an important confirmation, and the techniques he developed to achieve it have become valuable to the pursuit of consistency and independence proofs in general [5, p. 385]. Just as Gödel constructed a model of ZF in which AC holds, Cohen used his technique of "forcing" to construct a model of ZF in which AC fails. Hence, there exists both models of ZF where AC is true and models of ZF where AC fails, so neither AC nor its negation follows from the axioms of ZF alone.

Historical overviews of ZFC tend to end after Cohen's proof of the independence of AC. This is generally considered a culmination in the story of the development of axiomatic set theory, though it is by no means the conclusion or even the story in its entirety. However, the independence of AC was the last big open question that had ramifications for mathematics outside of set theory. We will conclude by examining an older implication of the axiom of choice, one from 1924, which demonstrates how AC can produce counter-intuitive results.

5. The Banach–Tarski Paradox

In 1924, Banach and Tarski proved, using the axiom of choice, that it is possible to obtain "a paradoxical decomposition of the sphere." That is, they showed that, given a sphere in $\mathbb{R}^3$, we can decompose it into finitely many disjoint pieces and, through only rotations and translations, reassemble it into two spheres, each with the same volume as the original. We call this result paradoxical not because it contradicts any of the axioms of ZFC but rather because it contradicts our own intuitive understanding of three-dimensional space: we begin with a single sphere, and we obtain two identical spheres by doing nothing but disassembling the original and reassembling its pieces. Moreover, we do this with only *finitely many* pieces. Finally, in their original paper, Banach and Tarski also proved that this paradox holds for *any* bounded subset of $\mathbb{R}^3$ with a non-empty interior [12, p. 33]. Indeed, the paradox can be generalized to any dimension $n \geq 3$. For $n < 3$, the free non-Abelian subgroups we require in the construction of the paradox are not available [12, p. 114].

The Banach–Tarski paradox is far from unique in portraying the counter-intuitive implications of the axiom of choice. Nevertheless, its geometrical nature and its easily-understood premise make it a paradigm example: it demonstrates how AC has implications outside the scope of pure set theory, and even people with a limited understanding of mathematics can comprehend the paradox, if not the proof behind it. Indeed, the proof uses AC in a very obvious way, but it is not immediately obvious *why* this usage allows us to decompose the sphere into two identical copies of itself. In fact, the axiom of choice implies the existence of sets that are not Lebesgue measurable (Vitali proved this in 1905 [7]). When we decompose the sphere, the resulting components will be non-measurable sets, i.e., they do not have "volume," and it is this trick that allows us to duplicate the volume of the original sphere.

We begin by defining some terms that will be used extensively throughout the proof. Unless otherwise stated, the sets here are subsets of $\mathbb{R}^3$.

Definition 5.1. A *rigid motion* is a mapping from $\mathbb{R}^3$ to $\mathbb{R}^3$ that is a translation, rotation, or combination of the two (but not a reflection).

Thus, if we apply a rigid motion to a set X, we receive back an image set Y that is a copy of X translated by some vector and rotated by some angle. For any sets X, Y, we say X and Y are *congruent* if and only if Y is the image of X under some rigid motion.

Definition 5.2. A *partition* of a set X is a collection of pairwise disjoint sets $\{X_i\}_{i=1}^n$, for some $n \in \mathbb{N}$, such that $X = \bigcup_{i=1}^n \{X_i\}$.

We may also refer to partitions as *finite decompositions*, since in the geometrical sense we have split X into distinct components.

Definition 5.3. Two sets X, Y are *equidecomposable*, denoted $X \approx Y$, if and only if there exist partitions $\{X_i\}_{i=1}^n$, $\{Y_i\}_{i=1}^n$ such that $X = \bigcup_{i=1}^n \{X_i\}$, $Y = \bigcup_{i=1}^n \{Y_i\}$, and X_i is congruent to Y_i for each $i = 1, \ldots, n$.

Lemma 5.4 ([7, p. 5])**.** *The following hold:*

(a) *equidecomposability is an equivalence relation.*
(b) *If X and Y are disjoint unions of X_1, X_2 and Y_1, Y_2, respectively, and if $X_i \approx Y_i$ for each $i = 1, 2$, then $X \approx Y$.*
(c) *If $X_1 \subseteq Y \subseteq X$ with $X \approx X_1$, then $X \approx Y$.*

Proof. (a) Obviously a set is equidecomposable with itself, so $\approx$ is reflexive. Similarly, if $X \approx Y$, then the same partitions show that $Y \approx X$, so we have symmetry. Finally, if $X \approx Y$ and $Y \approx Z$, then there exist partitions $\{X_i\}_{i=1}^n, \{Y_i\}_{i=1}^n, \{Z_i\}_{i=1}^n$ of X, Y, Z, respectively, such that each X_1 is congruent to Y_i and each Y_i is congruent to Z_i for $i = 1, \ldots, n$. Since congruence is transitive, each X_i is congruent to Z_i, so $X \approx Z$. Therefore, $\approx$ is an equivalence relation.

(b) Let $X = X_1 \cup X_2, Y = Y_1 \cup Y_2$, where $X_1 \cap X_2 = \emptyset = Y_1 \cap Y_2$, such that $X_1 \approx Y_1$ and $X_2 \approx Y_2$. Then

$$X_1 = \bigcup_{i=1}^n \{X_{1i}\} \text{ and } Y_1 = \bigcup_{i=1}^n \{Y_{1i}\},$$

where X_{1_i} is congruent to Y_{1_i} for $i = 1, \ldots, n$. Similarly,

$$X_2 = \bigcup_{i=1}^m \{X_{2i}\} \text{ and } Y_2 = \bigcup_{i=1}^m \{Y_{2i}\}.$$

So

$$X = \left(X_{1_1} \cup \ldots \cup X_{1_n}\right) \cup \left(X_{2_1} \cup \ldots \cup X_{2_m}\right)$$

and

$$Y = \left(Y_{1_1} \cup \ldots \cup Y_{1_n}\right) \cup \left(Y_{2_1} \cup \ldots \cup Y_{2_m}\right),$$

and it follows that $X \approx Y$.

(c) Since $X \approx X_1$, let $X = X^1 \cup \ldots \cup X^n$ and $X_1 = {X_1}^1 \cup \ldots \cup {X_1}^n$ such that X^i is congruent to ${X_1}^i$ for each $i = 1, \ldots, n$, and let $f_i : X^i \to {X_1}^i$ denote a congruence mapping. Let $f : X \longrightarrow X_1$ such that, for each $x \in X_i \subset X$, $f(x) = f_i(x)$. Then f is injective. Suppose $f(x) = f(y)$ for some $x, y \in X$. Since the ${X_1}^i$ are pairwise disjoint, we have $f_i(x) = f_i(y)$. But each f_i is a congruence mapping, which is injective, so $x = y$.

Now we define two sequences of sets recursively. Recall that $Y \subseteq X$. Let $X_0 = X$ and $Y_0 = Y$, and for $n \geq 1$, define

$$X_n = f(X_{n-1}), \quad Y_n = f(Y_{n-1})$$

and let

$$Z = \bigcup_{n=0}^{\infty} (X_n \setminus Y_n).$$

For any $n \geq 0$, suppose there exists $x \in X_n \setminus Y_n$ such that $f(x) \in Y_{n+1} = f(Y_n)$. Then $x \in Y_n$, because f is injective, which is a contradiction. Thus, for all n,

$$f(X_n \setminus Y_n) \subseteq X_{n+1} \setminus Y_{n+1},$$

so $f(Z) \subseteq Z$. So $f(Z) \approx Z$ and $f(Z)$ is disjoint from $X \setminus Z$. If we write

$$X = Z \cup (X \setminus Z), \quad Y = f(Z) \cup (X \setminus Z),$$

then by (b), it follows that $X \approx Y$.

□

Our proof of the Banach–Tarski paradox will mostly follow Jech's [7] in form, although we will augment it with elements from Stromberg's [10] where appropriate. We will first prove a result due to Hausdorff, wherein we decompose the sphere into three congruent sets and a countable set. This is now the conventional way of developing the Banach–Tarski paradox, though Banach and Tarski did not rely upon Hausdorff's theorem in their original proof [12, p. 33].

Theorem 5.5 ([7, p. 3])**.** *A sphere S can be decomposed into disjoint sets*

$$S = A \cup B \cup C \cup Q$$

such that:

1. *the sets A, B, C are congruent to each other;*
2. *the set $B \cup C$ is congruent to each of the sets A, B, C;*
3. *Q is countable.*

Let G be a free product of the groups $\{1, \varphi\}$ and $\{1, \psi, \psi^2\}$, where $\varphi^2 = \psi^3 = 1$, the identity element of G. We can in fact think of G as a group of matrices under matrix multiplication, where

$$\varphi = \begin{bmatrix} -\cos\theta & 0 & \sin\theta \\ 0 & -1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$$

and

$$\psi = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, φ and ψ correspond to rotations by 180° about an axis a_φ and 120° about an axis a_ψ, respectively. The angle $\theta \in \mathbb{R}$ is the angle between these two axes of rotation.

We can write elements of G as expressions of the form

$$\alpha = \varphi\psi^{\pm 1} \cdots \varphi\psi^{\pm 1},$$

since $\psi^2 = \psi^{-1}$. We call these expressions *words* in the letters φ, ψ, and ψ^2. An element's expression as a word may not be unique. For example, $\varphi\psi\varphi\varphi\psi = \varphi\psi^2$. We call the latter word a *reduced word* because it has been simplified to its most basic expression.

Nevertheless, our definition of φ and ψ mean that an element might be expressed as two different reduced words, depending on our choice of θ. We require a θ such that any two elements of G are distinct rotations. Stromberg uses the matrix definitions of φ and ψ and solves a system of polynomials to determine that any θ such that $\cos\theta$ is transcendental will suffice [10, Theorem A]. Jech does not go into as much detail, merely noting that "equations for orthogonal transformations and some elementary trigonometry" will guarantee that only countably many values of θ are undesirable [7, Lemma 1.3]. Therefore, we fix a θ from among the uncountably many values that remain acceptable.

Lemma 5.6 ([7, p. 4])**.** *There exists a partition* $\{G_1, G_2, G_3\}$ *of* G *such that*

(1) $G_1 \cdot \varphi = G_2 \cup G_3$,
(2) $G_1 \cdot \psi = G_2$, *and*
(3) $G_1 \cdot \psi^2 = G_3$.

Sketch of the proof. This is also Theorem B in [10]. We construct the partition of G by recursion on the lengths of elements of G. Begin by assigning

$$1 \in G_1, \varphi, \psi \in G_2, \psi^2 \in G_3.$$

Let $\alpha \in G$. If α ends with ψ or ψ^2, then we assign

$$\alpha\varphi \in G_2 \text{ if } \alpha \in G_1$$

and

$$\alpha\varphi \in G_1 \text{ if } \alpha \in G_2 \cup G_3.$$

If α ends with φ, then we assign

$$\alpha\psi \in G_2, \alpha\psi^2 \in G_3 \text{ if } \alpha \in G_1,$$

$$\alpha\psi \in G_3, \alpha\psi^2 \in G_1 \text{ if } \alpha \in G_2,$$

and

$$\alpha\psi \in G_1, \alpha\psi^2 \in G_2 \text{ if } \alpha \in G_3.$$

Thus, for any element of G, we can recursively determine in which element of the partition it lies. The rules above guarantee that the three conditions in the lemma are satisfied. For example, take $\alpha = \psi^2\varphi\psi$. Then we have

$$\psi^2 \in G_3 \Rightarrow \psi^2\varphi \in G_1 \Rightarrow \psi^2\varphi\psi \in G_2,$$

so $\alpha \in G_2$. Then $\alpha\psi = \psi^2\varphi\psi^2 \in G_3$, because $\psi^2\varphi \in G_1$. This is consistent with (iii) from above. □

Proof of Theorem 5.5. We will use this partition of G to establish the partition $\{A, B, C, Q\}$ of the sphere S. Let Q be the set of all points on S fixed by some $\alpha \in G$, where $\alpha \neq 1$. Since G is countable and each α has two fixed points that correspond to the poles of its axis of rotation. Thus, Q is a countable set, as we require.

We define the *orbit* of a point $x \in S$ by

$$P_x = \{x\alpha \mid \alpha \in G\}$$

and claim that if $x \in S \setminus Q$, then $P_x \subseteq S \setminus Q$, and furthermore, all such P_x form a partition of $S \setminus Q$. Suppose $x\alpha \in Q$. Then, by definition, there exists a $\rho \in G$ such that $\rho \neq 1$

and $x\alpha\rho = x\alpha$. Thus, $x\alpha\rho\alpha^{-1} = x$, and since $\alpha\rho\alpha^{-1} \neq 1$, we have $x \in Q$, which is a contradiction.

Now take $x, y \in S \setminus Q$. We will show that $P_x \cap P_y = \emptyset$ or $P_x = P_y$. Suppose P_x and P_y are not disjoint. Then there exist some $\alpha, \beta \in G$ such that $x\alpha = y\beta$. For any $x\sigma \in P_x$,

$$x\sigma = x\alpha\alpha^{-1}\sigma = y\beta\alpha^{-1}\sigma \in P_y,$$

and hence, $P_x = P_y$. Therefore, the P_x are disjoint for distinct $x \in S \setminus Q$.

We will now choose exactly one point from each $P_x, x \in S \setminus Q$ and denote the set of these points by M. This requires the axiom of choice, because we cannot otherwise assume M exists. Let $A = M \cdot G_1, B = M \cdot G_2, C = M \cdot G_3$. By Lemma 5.6, these sets are disjoint from each other. This is a partition for $S \setminus Q$, because for each point in M, applying the rotations in G_1, G_2, G_3 recover all the points in that orbit. Finally, owing to the way we constructed the partition of G, $A \cdot \psi = B$, $A \cdot \psi^2 = C$, and $A \cdot \varphi = B \cup C$, so A, B, C are congruent to each other and to $B \cup C$. Finally,

$$S = A \cup B \cup C \cup Q. \tag{5.1}$$

□

Recall that the above is a partition of the *surface* of a sphere, whereas the Banach–Tarski paradox claims we can decompose an entire closed ball, surface and interior. We will now prove this using our surface decomposition as a starting point.

Theorem 5.7 (Banach–Tarski Paradox [7, p. 3])**.** *Let U be a closed ball. Then there exists a partition $\{X, Y\}$ of U such that $U \approx X$ and $U \approx Y$.*

Proof. We shall denote the centre of the ball by c. Let

$$S = A \cup B \cup C \cup Q$$

be the decomposition of the surface of U from (5.1). For each $X \subset S$, let $\bar{X}$ denote the set of all $x \in U \setminus \{c\}$ such that the projection of x onto the surface is in X. Then according to our decomposition of S, we have

$$U = \bar{A} \cup \bar{B} \cup \bar{C} \cup \bar{Q} \cup \{c\},$$

and

$$\bar{A} \approx \bar{B} \approx \bar{C} \approx \bar{B} \cup \bar{C}. \tag{5.2}$$

Let $X = \bar{A} \cup \bar{Q} \cup \{c\}$ and $Y = U \setminus X$. Obviously $\{X, Y\}$ partitions U. We will show that $U \approx X$ and then that $U \approx Y$.

We can write $\bar{A} = \bar{A} \cup \bar{A}$, and then by (5.2) and Lemma 5.4(b), we have

$$\bar{A} \approx \bar{A} \cup \bar{B} \cup \bar{C}.$$

Consequently, from the definition of X and Lemma 5.4(b) again, $X \approx U$.

Take some rotation $\tau \notin G$ such that $Q \cap Q \cdot \tau = \emptyset$, i.e., the set of fixed points of S for all rotations in G is disjoint from its rotation under τ. Hence,

$$Q \cdot \tau \subseteq A \cup B \cup C,$$

and (5.2) implies

$$\bar{C} \approx \bar{A} \cup \bar{B} \cup \bar{C},$$

so there exists some $T \subseteq C$ such that $\bar{T} \approx \bar{Q}$.

Let $p \in \bar{C} \setminus \bar{T}$. Then since $\bar{A} \approx \bar{B}$ and $\bar{Q} \approx \bar{T}$, Lemma 5.4(b) implies

$$\bar{A} \cup \bar{Q} \cup \{c\} \approx \bar{B} \cup \bar{S} \cup \{p\}. \tag{5.3}$$

Since $X \approx U$, we can use the fact that

$$\bar{B} \cup \bar{S} \cup \{p\} \subseteq Y \subseteq U$$

and Lemma 5.4(c) to obtain $Y \approx U$.

Therefore, we have partitioned the closed ball U into two closed balls X and Y such that $U \approx X$ and $U \approx Y$. □

6. Conclusion

The Banach–Tarski paradox is notable not only for duplicating a ball but doing so using only finitely many pieces. Von Neumann, Sierpinski, and others have investigated the number of pieces required for such a decomposition. Stromberg [10] uses 40 pieces. The proof of Theorem 5.7, from [7], uses only five pieces, one of which contains only the centre of the ball. R.M. Robinson proved in 1947 that five pieces are the minimum [8].

The axiom of choice's role in the paradox is essential. Without it, we cannot choose a point from each of uncountably many sets. With it, we have the existence of such a set but no way of constructing it explicitly. The Banach–Tarski paradox's counter-intuitive result demonstrates how AC can sometimes have strange consequences. Furthermore, the paradox is geometrical in nature. Its result, albeit not the procedure leading up to it, is easily visualized. Hence, AC and its implications are not confined to discussions of axiomatic set theory. The axiom of choice has implications throughout mathematics.

Although most mathematicians use AC when necessary, it has not historically enjoyed the same acceptance as do the axioms of ZF. From its introduction by Zermelo to Cohen's independence proof, AC has been regarded separately, as something different. In part this is due to AC's non-constructive nature. However, it is also the result of larger philosophical debates among mathematicians, especially during the early part of the twentieth century. Similarly, axiomatic set theory arose amid discussions of the appropriate foundations for all of mathematics. It emerged from the program of logicism begun by Dedekind and Frege and continued into the twentieth century by Russell. ZermeloFraenkel set theory is formulated using predicate logic, and despite the limitations placed upon it by Gödel's incompleteness theorem, it remains the de facto foundation for mathematics.

Appendix: Axioms of ZFC

The following axioms, from [11], form the basis of ZFC. Axioms marked with an asterisk are redundant. We proved that the axiom schema of separation is a special case of the axiom schema of replacement in Proposition 3.7. The axiom schema of replacement and the power axiom together imply the pairing axiom.

Axiom of Extensionality:

$$(\forall x)(x \in A \iff x \in B) \Rightarrow A = B.$$

Axiom Schema of Separation:*

$$(\exists B)(\forall x)\,(x \in B \iff x \in A \wedge \varphi(x))\,.$$

Pairing Axiom:*

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

Sum Axiom:

$$(\exists C)(\forall x)\,(x \in C \iff (\exists B)(x \in B \wedge B \in A))\,.$$

Power Set Axiom:

$$(\exists B)(\forall C)(C \in B \iff C \subseteq A).$$

Axiom of Regularity:

$$A \neq \emptyset \Rightarrow (\exists x)\,(x \in A \wedge (\forall y)(y \in x \Rightarrow y \notin A))\,.$$

Axiom of Infinity:

$$(\exists A)\,(\emptyset \in A \wedge (\forall B)(B \in A \Rightarrow B \cup \{B\} \in A))\,.$$

Axiom Schema of Replacement:

$$(\forall x)(\forall y)(\forall z)\,(x \in A \wedge \varphi(x,y) \wedge \varphi(x,z) \Rightarrow y = z)$$
$$\Rightarrow (\exists B)(\forall y)\,(y \in B \iff (\exists x)\,(x \in A \wedge \varphi(x,y)))\,.$$

Axiom of Choice:

For any set A there is a function f such that for any non-empty subset B of A, $f(B) \in B$.

References

[1] P. Bernays, A. A. Fraenkel, *Axiomatic Set Theory*, 2nd ed, North-Holland, Amsterdam, 1968.

[2] G. Cantor, *Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen*, Journal für die reine und angewandte Mathematik 77 (1874), 258-62.

[3] G. Cantor, *Über eine elementare Frage der Mannigfaltigkeitslehre*, Jahresbericht der Deutschen Mathematiker-Vereinigung 1 (1892), 75-78.

[4] A. J. Dean, *Set Theory* (Print out)

[5] J. Ferreirós, *Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics*, 2nd ed, Birkhäuser, Basel, Switzerland, 2007.

[6] T. Forster, "Quine's New Foundations," *Stanford Encyclopedia of Philosophy* (2006). Available at: http://plato.stanford.edu/entries/quine-nf/

[7] T. J. Jech, *The Axiom of Choice*, North-Holland, Amsterdam, 1973.

[8] R. M. Robinson, *On the decomposition of sphere*, Fund. Math. 34 (1947), 246-260.

[9] H. Rubin, J. Rubin, *Equivalents of the Axiom of Choice*, North-Holland, Amsterdam, 1970.

[10] K. Stromberg, *The Banach–Tarski paradox*, Amer. Math. Monthly 86 (1976), 151-161.

[11] P. Suppes, *Axiomatic Set Theory*, D. Van Nostrand, Princeton, NJ, 1960.

[12] S. Wagon, *The Banach–Tarski Paradox*, Cambridge Univ. Press, New York, 1985.