

Mathematics Institute
University of Warwick

**Set Theory and the
Axiom of Choice**

Daniel Wood

Spring 2007

1 Introduction

The Axiom of Choice (**AC**) is a troublesome axiom. Simply stating that given any collection of non-empty sets there exists a function which takes exactly one element from each set, one might not immediately think it troublesome. Most textbooks accept **AC** without any comment and many theorems in various areas of mathematics require **AC** to be proved. The problem with **AC** is that it leads to counterintuitive results, some so counterintuitive as to be labelled ‘paradoxes’.

In this essay we will start by covering the historical development from naive set theory to axiomatic set theory and the formulation of **AC**. We will then prove the Banach–Tarski Paradox, a disturbing result implied by **AC**. Finally we shall look at two more unpleasant results, this time caused by the negation of **AC**.

2 Lost Innocence: The Downfall of Naive Set Theory

In this section we will explore the transition from naive set theory to axiomatic set in its historical context. We will finish by defining **AC**.

2.1 The Axiom of Abstraction and Russell’s Paradox

Axiom of Abstraction (AA). *For any given property there exists a set whose members are all and only those entities that have that property.*

We can rewrite this as:

Given some property A there exists a set $\{x : A(x)\}$.

This seemingly innocent axiom, although not explicitly stated until 1893 by Gottlob Frege, was used implicitly by Cantor when he developed set theory in the 1870s.¹ While Cantor’s ideas proved to be crucial in the formalisation of mathematics that occurred in the late 19th century, a problem arose when it was discovered that **AA** led to various paradoxes. (This is why Cantor’s set theory is today referred to as *naive*.) Perhaps the most well-known of these paradoxes is *Russell’s Paradox*, discovered in 1901 by Bertrand Russell, which showed **AA** to be false:

Paradox 2.1 (Russell). The set $z := \{x : x \notin x\}$ does not exist.

Proof. Clearly either $z \in z$ or $z \notin z$. First suppose that $z \in z$. Then, by the definition of the set z , we have that $z \notin z$ which is a contradiction. Now suppose that $z \notin z$. Then, again by the definition of z , we have that $z \in z$ which is a contradiction. Thus z cannot exist. \square

¹[Supp72], p. 5.

Despite Frege's dismay,² Russell's discovery meant that set theorists had to reassess their field; it appeared to have the potential to completely undermine the whole of set theory.

2.2 Zermelo to the rescue!

In 1908 Ernst Zermelo introduced what is usually called the *Axiom Schema of Separation* (or *Subsets*).

Axiom Schema of Separation (AS). *Let y be a set. Given any property A there exists a set $\{x : A(x) \wedge x \in y\}$.*³

The difference between **AS** and **AA** is not immediately clear. **AA** is an unconditional statement; we are allowed to take any property we so wish and construct a set whose elements are those entities which have that property. **AS** has the condition that we are first given an existing set from which we can construct a subset whose elements have some property. This is a very important difference which allows us to resolve Russell's Paradox:

Solution 2.2. In the proof of Paradox 2.1, we asserted the existence of the set $z := \{x : x \notin x\}$. But **AS** does not allow us this unconditional assertion of the existence of z ; instead we have to rewrite z as

$$z' := \{x : x \notin x \wedge x \in y\}$$

where y is some given set. We now suppose that $z' \notin z'$ (as in the proof of Paradox 2.1). Then

$$\neg(z' \notin z' \wedge z' \in y)$$

which, by De Morgan's laws, is equivalent to

$$z' \in z' \vee z' \notin y.$$

This is not a contradiction, since we simply take $z' \notin y$ to be true and $z' \in z'$ to be false. Thus we do not have a paradox.

AS was one of a set of axioms introduced by Zermelo in 1908. These axioms were later modified by A. Fraenkel and T. Skolem and today are known as the *Zermelo–Fraenkel Axioms*. Although there are alternative axioms for set theory, perhaps most notably the *von Neumann–Bernays–Gödel Axioms*, most of modern set theory is based on the Zermelo–Fraenkel Axioms.⁴ While **AS** resolved the paradoxes of naive set theory, one of Zermelo's axioms, the *Axiom of Choice*, would go on to cause even more trouble than **AA** had.⁵

²Ibid.

³Note that if there does not exist $x \in y$ such that $A(x)$, then $\{x : A(x) \wedge x \in y\} = \emptyset$.

⁴The details of the different axioms are not discussed here. See Chapter 7 of [Sto179] for a thorough discussion.

⁵Zermelo actually first stated **AC** in 1904.

2.3 The Axiom of Choice

Axiom of Choice (AC). For any collection of non-empty sets there exists a function that takes exactly one element from each set.

Notice that if we have a finite collection of sets then we do not need **AC**:

Proposition 2.3. Let X_1, X_2, \dots, X_n be a finite collection of non-empty sets. Then there exists a function which takes exactly one element from each set.

Proof. Let \mathcal{C} be the collection of sets X_1, X_2, \dots, X_n . Since each X_i is non-empty, there exist $x_1 \in X_1, x_2 \in X_2, \dots, x_n \in X_n$. Then the function

$$f(x) = \begin{cases} x_1 & \text{if } x \in X_1, \\ x_2 & \text{if } x \in X_2, \\ \vdots & \vdots \\ x_n & \text{if } x \in X_n \end{cases}$$

takes exactly one element from each set in \mathcal{C} . □

Notation 2.4. We will denote the Zermelo–Fraenkel Axioms *with AC* by **ZFC** and the Zermelo–Fraenkel Axioms *without AC* by **ZF**. We shall mark any result whose proof uses **AC** with an asterisk.

But when do we use **AC**? The following often-used theorem from elementary analysis has **AC** at the heart of its proof.

Theorem 2.5.* Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function which is sequentially continuous at some point $a \in \mathbb{R}$. Then f is continuous at a .

Proof. Suppose that f is not continuous at a . Then there exists $\varepsilon > 0$ such that for every $\delta > 0$ we have $|x - a| < \delta$ but $|f(x) - f(a)| > \varepsilon$. Consider such an ε . Then for some $\delta_1 > 0$ we take $x_1 \in \mathbb{R}$ such that $|x_1 - a| < \delta_1$. Let $\delta_2 = \frac{\delta_1}{2}$. We take $x_2 \in \mathbb{R}$ such that $|x_2 - a| < \delta_2$. We repeat this process, each time setting $\delta_{n+1} = \frac{\delta_n}{2}$ and taking $x_n \in \mathbb{R}$ such that $|x_n - a| < \delta_n$. Clearly $(\delta_n) \rightarrow 0$. Thus $(x_n) \rightarrow a$. Therefore, since f is sequentially continuous, $f(x_n) \rightarrow f(a)$ as $n \rightarrow \infty$. But this is a contradiction, since $|f(x_n) - f(a)| > \varepsilon$ for every x_n . Therefore f is continuous at a . □

We used **AC** in the above proof to pick each x_n , since we assumed that that there existed a function which took exactly one element from each of the sets $(\{x : |x - a| < \delta_n\})_{n=1}^{\infty}$.

3 The Banach–Tarski Paradox

Put simply, the Banach–Tarski Paradox (**BTP**) allows us to take apart a sphere in \mathbb{R}^3 into finitely many pieces and rearrange these pieces using only rigid motions to form two new spheres each with the same radius as the original sphere, i.e. we can ‘double the sphere’. At first glance the reader may think that there is some mistake, but it is a fact that if we accept **AC** then we also have to accept **BTP**.

Throughout this section we will suppose that **AC** is true, although we will still indicate which results use **AC** in their proofs.

3.1 Proving BTP

To prove this somewhat perturbing result we will need to do quite a bit of work. But do not fear! With a little perseverance, and perhaps some inspiration, we shall feel the sweet satisfaction from a hard theorem proved. We shall start with a definition of a free group, something that is crucial in showing **BTP**.

Definitions 3.1. Let G be a group with identity element e . G is a *free group* iff there exists a subset X of G such that every $g \in G \setminus \{e\}$ can be expressed in a unique way as a product

$$g = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$$

of finite length, where each x_i is in X , $x_i \neq x_{i+1}$, and each k_i is a non-zero integer. The elements of X are called the *free generators* of G .

Theorem 3.2. *The free group with n free generators is unique up to isomorphism.*

Proof. Let G and H be free groups with sets of free generators $X := \{x_1, x_2, \dots, x_n\}$ and $Y := \{y_1, y_2, \dots, y_n\}$ respectively. Let $\varphi : X \rightarrow Y$ be the map defined by $\varphi(x_i) = y_i$. Then the map $\varphi' : G \rightarrow H$ defined by

$$\varphi'(x_a^{k_a} x_b^{k_b} \dots x_c^{k_c}) = \varphi(x_a)^{k_a} \varphi(x_b)^{k_b} \dots \varphi(x_c)^{k_c}$$

is an isomorphism. □

We denote the free group with n free generators by F_n .

Definitions 3.3. A *word* in a set X is a finite product

$$x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \tag{1}$$

where each x_i is in X and each k_i is an integer. A word is *reduced* iff $x_i \neq x_{i+1}$ and each k_i is a non-zero integer. The word which is a product of no elements of X is the *empty word*.

Remark 3.4. The empty word is often said to be *trivially* reduced. Thus to avoid any confusion, when we say that a word is reduced we shall take it to be that the word in question is *not* the empty word (unless otherwise stated).

As we say with words in English, the *length* of a word is the number of ‘letters’ in the word from the ‘alphabet’ X . So for example, the length of the word in (1) is n and the length of the empty word is 0. Again as with English, we say that a word *begins* with $x \in X$ iff x is the first element (from left to right) in the word. So in (1), for example, the word begins with x_1 .

We now go through an example of how we reduce words:

Example 3.5. Let $X := \{a, b, c\}$. Consider the word

$$w := ca^0b^{-2}b^5ac^4c^{-4}a^{-2}a^3a^{-7}b^2$$

in X . Then

$$\begin{aligned} w &= cb^3ac^0a^{-6}b^2 \\ &= cb^3aa^{-6}b^2 \\ &= cb^3a^{-5}b^2 \end{aligned}$$

which is now reduced.

We shall not prove it here, but every word can be simplified to only one reduced word (see pp. 167–168 of [Arms88] for a proof).

We see that given a group X we can generate a group from the reduced words in X . Whether or not such a group is free is the subject of the following lemma. But before we state and prove this lemma we shall go over an example to explain the terms *distinct* and *equal*:

Example 3.6. Let G be the group generated by the words in the set $X := \{a, b, c, d\}$. Consider the two words

$$\nu := a^1a^{-3}c^4d \text{ and } v := a^{-2}c^4d$$

in X . While ν and v are clearly equal in G , as *words* they are distinct since they are different products of elements in X . Now, in this example v is simply the reduced form of ν , but to assume that two distinct reduced words are equal in G because they are reduced is incorrect.

Lemma 3.7. *Let G be the group generated by the reduced words in a set X , where $|X| = n \in \mathbb{N}$. Then $G \cong F_n$ iff the identity element in G is not equal to any reduced word and, if $n \geq 2$, G is non-abelian.⁶*

Proof. (\Rightarrow) Let $n \geq 2$ and suppose that F_n is abelian. Consider $g = ab \in F_2$

⁶ $H \cong F_n$ means that H is isomorphic to F_n .

where a, b are free generators. Then, since F_n is abelian, $g = ab = ba$ and so g can be expressed as two distinct reduced words in the set of free generators. But by Definitions 3.1 this is a contradiction (since every element in F_n can be *uniquely* expressed as a reduced word).

Let e be the identity element of F_n . Suppose that we have a reduced word ν such that $\nu = e$. Take some $x \in X$ such that ν does not begin with x . Then

$$\begin{aligned} x\nu x^{-1} &= xex^{-1} \\ &= x^0 \\ &= e. \end{aligned}$$

Since ν does not begin with x , we see that $x\nu x^{-1}$ can be reduced to a word v which is distinct from ν . Thus we have two distinct reduced words ν and v with $\nu = v$ which contradicts our definition of a free group.

(\Leftarrow) It suffices to show that no two distinct words are equal. Suppose we have two distinct reduced words ν and v such that $\nu = v$. Then we have

$$\nu^{-1}\nu = \nu^{-1}v$$

which is equivalent to

$$e = \nu^{-1}v. \tag{2}$$

Since ν and v are reduced words, we can write them as $\nu = x_1^{k_1} \dots x_n^{k_n}$ and $v = y_1^{l_1} \dots y_m^{l_m}$, where each x_i and y_i are in X ; $x_i \neq x_{i+1}$ and $y_i \neq y_{i+1}$; and each k_i and l_i are non-zero integers. We see that $\nu^{-1} = x_n^{-k_n} \dots x_1^{-k_1}$ and so we can rewrite (2) as

$$e = x_n^{-k_n} \dots x_1^{-k_1} y_1^{l_1} \dots y_m^{l_m}. \tag{3}$$

Let $\tau := x_n^{-k_n} \dots x_1^{-k_1} y_1^{l_1} \dots y_m^{l_m}$ be the word in (2). Since ν and v are distinct, for at least one i we have that $x_i \neq y_i$ or $k_i \neq l_i$. Therefore we can reduce τ and so (3) implies that we have a reduced word equal to e which is a contradiction. \square

The reader can be forgiven for being a little confused at this point (the author certainly was when he first covered this topic). To help clarify what a free group actually is, we shall go through some examples:

Examples 3.8. (i) Let ρ be the clockwise rotation of $\pi/2$ about the origin in \mathbb{R}^2 . Then the reduced words in the set $X := \{\rho\}$ generate a group; we can write this group as $G := \{\rho^k : k \in \mathbb{Z}\}$. This group is *not* free, since ρ^4 is equal to the identity and so

$$\rho = \rho^5 = \rho^9 = \dots,$$

contradicting the criterion that every $g \in G$ can be expressed in a unique way as ρ^k for some integer k .

(ii) Let φ be the clockwise rotation of $\sqrt{2}\pi$ about the origin in \mathbb{R}^2 . Then the reduced words in the set $X := \{\varphi\}$ generate a group; we can write this group as $G := \{\varphi^k : k \in \mathbb{Z}\}$. Since $\sqrt{2}$ is irrational, we see that no integers m, n satisfy the equation

$$m\sqrt{2}\pi = 2n\pi.$$

Hence there does not exist $k \in \mathbb{Z}$ such that φ^k is equal to the identity. Thus $i \neq j \Rightarrow \varphi^i \neq \varphi^j$ and so G is free.

(iii) Let $\mathbb{Z}^3 := \{(z_1, z_2, z_3) : z_1, z_2, z_3 \in \mathbb{Z}\}$. We see that every $g \in \mathbb{Z}^3$ can be written as

$$g = i(1, 0, 0) + j(0, 1, 0) + k(0, 0, 1)$$

where $i, j, k \in \mathbb{Z}$. We can rewrite this in multiplicative notation as

$$g = a^i b^j c^k$$

where $a := (1, 0, 0), b := (0, 1, 0), c := (0, 0, 1)$. Because \mathbb{Z}^3 is abelian, we see that

$$g = a^i b^j c^k = b^j a^i c^k.$$

Thus $g \in \mathbb{Z}^3$ cannot be written in a unique way as a reduced word in $X := \{a, b, c\}$ and so \mathbb{Z}^3 is *not* free. Notice, however, that no reduced word in $X := \{a, b, c\}$ is equal to the identity.

Note: \mathbb{Z}^3 is an example of a *free abelian group*. The reader should not confuse free groups and free *abelian* groups. For a good introduction to the topic of free abelian groups, see Professor Derek Holt's lecture notes for the course *MA251 Algebra I: Advanced Linear Algebra* (written September 2006).

It should now be clear that F_1 is the infinite cyclic group; and that for $n \geq 2$, F_n is a non-abelian group with every element having infinite order.

We shall now get into the business of actually proving **BTP**.

Definition 3.9. Let G be a group with identity element e and let X be a set. G acts on X iff there exists a map

$$\mu : G \times X \rightarrow X,$$

denoted $\mu(g, x) = g(x)$, such that

- (i) $e(x) = x$ for every $x \in X$; and
- (ii) $(gh)(x) = g(h(x))$ for every $g, h \in G$ and $x \in X$.

Please note that most of the following working up to and including the proof of Theorem 3.20 is taken from [Su90] and Chapters 1–3 of [Wag86]. While this author has re-worded and attempted to clarify the definitions, theorems, and proofs, he makes no claims of originality to the key ideas used.

Definition 3.10. Let G be a group acting on a set X and suppose that $E \subseteq X$. E is G -paradoxical iff for some $m, n \in \mathbb{N}$ there exist $g_1, \dots, g_m, h_1, \dots, h_n \in G$ and pairwise disjoint $A_1, \dots, A_m, B_1, \dots, B_n \subset E$ such that $E = \bigcup g_i(A_i) = \bigcup h_i(B_i)$.

We can refer to paradoxical groups by letting G act on itself by left multiplication. When this is the case we shall say that G is paradoxical (rather than G is G -paradoxical).

Theorem 3.11. F_2 is paradoxical.

Proof. Let a and b be the generators of F_2 and let e be the identity element of F_2 . Consider the sets $B(\rho) := \{\rho\varphi : \varphi \text{ is a reduced word in } F_2\}$, where $\rho \in \{a, a^{-1}, b, b^{-1}\}$. Then $F_2 = \{e\} \cup B(a) \cup B(a^{-1}) \cup B(b) \cup B(b^{-1})$. But $F = B(a) \cup aB(a^{-1}) = B(b) \cup bB(b^{-1})$. Thus F_2 is paradoxical. \square

Theorem 3.12.* Let G be a paradoxical group acting on a set X such that only the identity element of G fixes any points of X . Then X is G -paradoxical.

Proof. Let $A_i, B_j \subseteq G$ and $g_i, h_j \in G$ be the sets and elements witnessing that G is paradoxical.⁷

By **AC** we can take an element from each of the sets $(\{gx : g \in G\})_{x \in X}$; let M be the set of these elements. We see that $gM := \{g(M) : g \in G\}$ partitions X because only the identity fixes any points of X .

Let $A'_i = \bigcup \{g(M) : g \in A_i\}$ and $B'_j = \bigcup \{g(M) : g \in B_j\}$. Then $\{A'_i\} \cup \{B'_j\}$ is a pairwise disjoint collection of subsets of X because $\{A_i\} \cup \{B_j\}$ is pairwise disjoint and gM partitions X . Thus $X = \bigcup g_i(A'_i) = \bigcup h_i(B'_j)$. \square

Notation 3.13. We denote the group of orthogonal 3×3 matrices with determinant equal to $+1$ by SO_3 .⁸

It turns out that every element of SO_3 represents a rotation in \mathbb{R}^3 which fixes the origin and, conversely, that every rotation in \mathbb{R}^3 which fixes the origin is represented by an element of SO_3 (see p. 47 of [Arm88] for a proof).

Theorem 3.14. There exists a subgroup H of SO_3 such that $H \cong F_2$.

Proof. Let φ and ρ be anticlockwise rotations through the angle $\arccos(3/5)$ about the z -axis and x -axis respectively. Then

⁷For an existential statement $(\exists x \in X)(A(x))$ we say that $\varphi \in X$ witnesses the statement iff $A(\varphi)$ is true.

⁸ SO stands for 'Special Orthogonal'.

$$\varphi^{\pm 1} = \begin{pmatrix} 3/5 & \mp 4/5 & 0 \\ \pm 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \rho^{\pm 1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & \mp 4/5 \\ 0 & \pm 4/5 & 3/5 \end{pmatrix}$$

Let H be the group generated by the reduced words in $\{\varphi, \rho\}$. Clearly H is a non-abelian subgroup of SO_3 . We aim to show that no reduced word is equal to the identity, since if this is so, then by Lemma 3.7 we have that $H \cong F_2$.

Let e be the identity and let w be some reduced word. We claim that $w(0, 1, 0)$ is of the form $(a, b, c)/5^n$ where a, b, c are integers, b is not divisible by 5, n is the length of w . If this is true then $w(0, 1, 0) \neq (0, 1, 0)$ and so $w \neq e$.

We will prove the claim by induction. If $n = 1$, then either $w = \varphi^{\pm 1}$ or $w = \rho^{\pm 1}$. Since $\varphi^{\pm 1}(0, 1, 0) = (\mp 4, 3, 0)/5$ and $\rho^{\pm 1}(0, 1, 0) = (0, 3, \pm 4)/5$, the case for $n = 1$ is true.

Now suppose that the case for $n = k - 1$ is true. Then for any reduced word w' of length $k - 1$ we have $w'(0, 1, 0) = (a', b', c')/5^{k-1}$ where a', b', c' are integers and b' is not divisible by 5. Now, any reduced word w of length k is either of the form $w = \varphi^{\pm 1}w'$ or $w = \rho^{\pm 1}w'$. If $w = \varphi^{\pm 1}w'$, then $w(0, 1, 0) = (a, b, c)/5^k$ where

$$a = 3a' \mp 4b', \quad b = 3b' \pm 4a', \quad \text{and } c = 5c'.$$

If $w = \rho^{\pm 1}w'$, then $w(0, 1, 0) = (a, b, c)/5^k$ where

$$a = 5a', \quad b = 3b' \mp 4c', \quad \text{and } c = 3a' \pm 4b'.$$

Thus a, b, c are integers. To show that b is not divisible by 5 isn't too tricky, but it is somewhat laborious and so we do not show it here (see pp. 13–14 of [Su90] for a proof).⁹ \square

Before we prove the following corollary, recall that $S^2 := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ is the unit sphere about the origin in \mathbb{R}^3 .

Corollary 3.15 (Hausdorff Paradox).* *There exists a countable set D such that $S^2 \setminus D$ is SO_3 -paradoxical.*

Proof. Consider the free subgroup H from the proof of Theorem 3.14. Now, every $h \in H$ fixes two points on S^2 . Let D be the set of all points on the sphere fixed by some $h \in H$. Then D is countable. We apply Theorems 3.11 & 3.12 to obtain the desired result. \square

Definitions 3.16. Let G be a group acting on a set X and let $A, B \subseteq X$. A and B are G -equidecomposable, denoted $A \sim_G B$, iff for some $n \in \mathbb{N}$ we can partition A and B into A_1, \dots, A_n and B_1, \dots, B_n respectively such that each A_i is G -congruent to B_i . A_i is G -congruent to B_i , denoted $A_i \approx_G B_i$, iff there exists $g_i \in G$ such that $g_i(A_i) = B_i$.

⁹Please note that for Theorem 3.13 both Wagon and Su also use proof by laziness.

Lemma 3.17.* S^2 and $S^2 \setminus D$ are SO_3 -equidecomposable.¹⁰

Proof. Since D is countable, we can take a line l passing through the origin which intersects S^2 at two points both not in D . Let A be the set of angles θ such that for some integer $n > 0$ and some $P \in D$ we have that $\rho_\theta^n(P) \in D$, where ρ_θ is the rotation about l through angle θ . We claim that A is countable (we will prove this later). Thus we can take some angle $\varphi \notin A$; let ψ be the corresponding rotation about l . Thus for distinct integers $m, n \geq 0$ we have that $\phi^m(D) \cap \psi^n(D) = \emptyset$. Let $\Psi := \{\psi^n(D) : n \in \mathbb{N}\}$ and $B := S^2 \setminus \Psi$. Then $S^2 \setminus D = \Psi \cup B \sim_{SO_3} B \cup \psi^{-1}(\Psi) = S^2$.

We now prove our claim that A is countable. Consider some $\theta \in A$. Since ρ_θ^n is equal to the identity for some integer $n > 0$, we see that $n\theta = 2m\pi$ for some integers $n, m > 0$. So θ can be written as $\frac{2m\pi}{n}$. Thus $|A| \leq |\mathbb{Q}| = \aleph_0$.¹¹ \square

Lemma 3.18. Let G be a group acting on a set X . Suppose that E, E' are G -equidecomposable subsets of X . If E is G -paradoxical, then so is E' .

Proof. The result follows immediately from Definitions 3.10 and 3.16. \square

Notation 3.19. $B^3 := \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ is the unit ball in \mathbb{R}^3 about the origin.

Theorem 3.20 (BTP).* S^2 is SO_3 -paradoxical. Moreover, B^3 is SO_3 -paradoxical.

Proof. By Corollary 3.15 and Lemmas 3.17 & 3.18, we have that S^2 is SO_3 -paradoxical.

Since S^2 is SO_3 -paradoxical, for some $m, n \in \mathbb{N}$ there exists $g_1, \dots, g_m, h_1, \dots, h_n \in SO_3$ and pairwise disjoint $A_1, \dots, A_m, B_1, \dots, B_n \subset S^2$ such that $S^2 = \bigcup g_i(A_i) = \bigcup h_i(B_i)$. Consider the map $f : S^2 \rightarrow \mathcal{P}(B^3)$ defined by $f(P) = \{\alpha P : 0 < \alpha \leq 1\}$. Then $(\{\alpha P : 0 < \alpha \leq 1\})_{P \in S^2}$ partitions $B^3 \setminus \{\mathbf{0}\}$.¹² Thus $B^3 \setminus \{\mathbf{0}\} = \bigcup g_i(f(A_i)) = \bigcup h_i(f(B_i))$ and so $B^3 \setminus \{\mathbf{0}\}$ is SO_3 -paradoxical.

Consider an line through the point $(0, 0, \frac{1}{2})$ which does not meet the origin. Let ρ be a rotation about this axis such that for no integer $n > 0$ does ρ^n equal the identity. Let $D := \{\rho^n(\mathbf{0}) : n \geq 0\}$. Then $\rho(D) = D \setminus \{\mathbf{0}\}$ and so $D \approx_{SO_3} D \setminus \{\mathbf{0}\}$. Thus $B^3 \sim_{SO_3} B^3 \setminus \{\mathbf{0}\}$. Therefore, by Lemma 3.18, B^3 is SO_3 -paradoxical. \square

Remark 3.21. Notice that the above proof easily lifts to balls of any radius centred anywhere in \mathbb{R}^3 : we translate the ball so its centre moves to the origin and let $\alpha \in (0, r]$, where r is the radius of the ball.

This is a very counterintuitive result, so much so that it is used by many

¹⁰ D is the countable set from Corollary 3.15.

¹¹Recall that $\aleph_0 := |\mathbb{N}|$.

¹² $\mathcal{P}(X)$, the power set of X , is the set of all subsets of X

mathematicians to argue that **AC** must be false. However, the non-constructive¹³ nature of the proof is seen by some mathematicians to show that **BTP** is not a paradox at all, but rather a false conclusion drawn from incorrect mathematical methods. It is left to the reader to decide whether they accept non-constructive proofs.

But exactly how many pieces do we need to carry out this paradoxical decomposition of B^3 ? Well, in 1947 Raphael Robinson showed that as few as five pieces suffice, i.e. there exist $g_1, g_2, g_3, h_1, h_2 \in G_3$ and pairwise disjoint $A_1, A_2, A_3, B_1, B_2 \subset B^3$ such that $B^3 = g_1(A_1) \cup g_2(A_2) \cup g_3(A_3) = h_1(B_1) \cup h_2(B_2)$. Moreover, he showed that five is in fact the smallest possible number of pieces with which we can carry out **BTP**.¹⁴

3.2 Is AC to Blame?

How do we know that **AC** is the cause of **BTP**? Are similar paradoxical decompositions possible in **ZF**? The trick is to use model theory and measure theory. We will start by going over some (simplified) model theory.

Definitions 3.22. A *structure* is a non-empty set D (called the *domain* of the structure) along with a set of functions and relations well-defined in D .

Let's go through an example of a structure:

Example 3.23. Consider the function $f(x, y) = x + y \pmod{n}$ and the relation $R(x, y)$ defined by $x \equiv y \pmod{n}$. We see that these are both well-defined in \mathbb{Z} . Thus with f and R we can construct a structure \mathcal{S} with domain \mathbb{Z} . We can write \mathcal{S} as $\mathcal{S} := \{\mathbb{Z}; +_n, \equiv_n\}$.

Definition 3.24. Let \mathcal{S} be structure and let ν be a statement which is defined in D . \mathcal{S} *models* ν (or \mathcal{S} is a *model* of ν) if ν is true in \mathcal{S} .

At first glance, the above definition is a little confusing. We shall consider an example of an model:

Example 3.25. Let $\mathcal{S} := \{\mathbb{Z}; +_n, \equiv_n\}$ be the structure from Example 3.23. Consider the following statements:

- (i) $(\forall x \in \mathbb{Z})(x + 0 \equiv x \pmod{n})$.
- (ii) $(\exists x \in \mathbb{Z})(\forall y \in \mathbb{Z})(x + y \equiv 1 \pmod{n})$.
- (iii) $(\forall x \in \mathbb{Z})(\exists y \in \mathbb{Z})(x + y \equiv 1 \pmod{n})$.

We can see that (i) and (iii) are true in \mathcal{S} while (ii) is false in \mathcal{S} . So \mathcal{S} models (i) and (iii).

¹³A proof is *non-constructive* iff it proves an existential statement $(\exists x \in X)(A(x))$ without finding a witness to the statement.

¹⁴[Rob47].

The above example is very simple but it serves to illustrate the concept of a model; much more elaborate models can be constructed.¹⁵

We can now show that **AC** is the cause of **BTP**. (Please note that the following is taken from [Herr06], p. 133.) Since Lebesgue measure in \mathbb{R}^3 is additive, invariant, and normed, **BTP** implies that there exist bounded subsets of \mathbb{R}^3 which are not Lebesgue measurable. Thus, since there exist models of **ZF** in which all bounded subsets of \mathbb{R}^3 are Lebesgue measurable,¹⁶ **AC** must be the culprit.

4 Problems without AC

We will now suppose that **AC** is false and explore two undesirable results caused by this alternative supposition.

4.1 Problems with Cardinal Arithmetic

In **ZF** there exist sets whose sizes we cannot compare. To prove this—after recovering from the shock of such a terrible result—we shall start by considering two different definitions of finiteness:

Definition 4.1. A set X is *finite* iff $|X| = n$ for some $n \in \mathbb{N}$.¹⁷ A set X is *infinite* iff it is not finite.

Definition 4.2. A set X is *Dedekind-infinite*, or just *D-infinite*, iff there exists a proper subset Y of X such that $|X| = |Y|$. A set X is *Dedekind-finite*, or just *D-finite*, iff it is not *D-infinite*.

Before we get stuck in, we had better recall some notation:

Notation 4.3. Let A and B be non-empty sets. Then

- (i) $|A| = |B|$ means that there exists a bijection $f : A \rightarrow B$; and
- (ii) $|A| \leq |B|$ means that there exists an injection $f : A \rightarrow B$.

We now state and prove some results leading from the above definitions.¹⁸ We start with a lemma:

Lemma 4.4. A set X is *D-infinite* iff $\aleph_0 \leq |X|$.

¹⁵Model theory is in fact fundamental to symbolic logic and metamathematical results such as Gödel's 1931 Incompleteness Theorems. For a thorough introduction to the field see [Hedm04].

¹⁶An example of such a model is Shelah's Model II ($\mathcal{M}38$ in [HoRu98] (p. 168)).

¹⁷For the purposes of this definition we include 0 in \mathbb{N} .

¹⁸Lemma 4.4 & proof, Theorems 4.5, 4.7, and Definition 4.9 are taken from [Herr06], pp. 44, 47–48.

Proof. (\Rightarrow) Let $f : X \rightarrow X$ be an injection onto some proper subset of X . Take some $y \in X \setminus f(X)$. We can define recursively an injection $g : \mathbb{N} \rightarrow X$ by $g(0) = y, g(n+1) = f(g(n))$.

(\Leftarrow) Let $\nu : \mathbb{N} \rightarrow X$ be an injection. Then the map $v : X \rightarrow X \setminus \{\nu(0)\}$ defined by

$$v(x) = \begin{cases} \nu(n+1) & \text{if } x = \nu(n), \\ x & \text{otherwise} \end{cases}$$

is a bijection. Since $X \setminus \{\nu(0)\}$ is a proper subset of X , we have that X is D -finite. \square

For the next result we assume **AC**.

Theorem 4.5.* *A set X is finite iff it is D -finite.*

Proof. (\Leftarrow) Suppose that X is infinite. We shall construct a sequence (x_n) such that $m \neq n \Rightarrow x_m \neq x_n$. Take some x_1 in X . By **AC** we can construct our sequence recursively by taking some x_n from the set $X \setminus \{x_1, x_2, \dots, x_{n-1}\}$ for each $n \in \mathbb{N}$. The map $f : \mathbb{N} \rightarrow X$ defined by $f(n) = x_n$ is clearly injective. Thus, by Lemma 4.4, X is D -infinite. Therefore, by considering the contrapositive of our implication (infinite $\Rightarrow D$ -infinite) we see that if a set is D -finite, then it is finite.

(\Rightarrow)¹⁹ Consider a set X such that $|X| = n$ for some $n \in \mathbb{N}$. Then any proper subset of X contains at most $n-1$ elements. Thus X is D -finite. \square

Lemma 4.6. *If a set X is infinite and $|X| \leq \aleph_0$, then $|X| = \aleph_0$.*

Proof. Consider some injection $f : X \rightarrow \mathbb{N}$. Let $\mathbb{N}_f := \{n \in \mathbb{N} : n = f(x) \text{ for some } x \in X\} \subseteq \mathbb{N}$. Then $f : X \rightarrow \mathbb{N}_f$ is a bijection and so $|X| = |\mathbb{N}_f|$. Since \mathbb{N}_f is infinite, we can construct a bijection between \mathbb{N}_f and \mathbb{N} . Thus $|X| = \aleph_0$. \square

Theorem 4.7. *The following are equivalent:*

- (i) *a set X is finite iff it is D -finite;*
- (ii) *for every set X we have $|X| \leq \aleph_0$ or $\aleph_0 \leq |X|$.*

Proof. (i) \Rightarrow (ii) Suppose that X is finite. Then clearly $|X| \leq \aleph_0$. Now suppose that X is infinite. Then X is D -finite and so, by Lemma 4.4, $\aleph_0 \leq |X|$.

(ii) \Rightarrow (i) Suppose that X is finite. Then by Theorem 4.5(\Rightarrow) (whose proof does not require **AC**), X is D -finite. Now suppose that X is D -finite. If $\aleph_0 \leq |X|$, then by Lemma 4.4, X is D -infinite which is a contradiction. Thus $|X| \leq \aleph_0$. If X is infinite, then by Lemma 4.6, $|X| = \aleph_0$. But this is a contradiction since \mathbb{N} is D -infinite. Therefore X is finite. \square

¹⁹Notice that the proof of this implication does not require **AC**.

We can in fact do even better than the above theorem:

Theorem 4.8. *The following are equivalent:*

- (i) *a set X is finite iff it is D -finite;*
- (ii) *for every pair of sets A and B we have $|A| \leq |B|$ or $|B| \leq |A|$.*²⁰

Proof. (i) \Rightarrow (ii) We assume A and B to be infinite (if A and B are finite then (ii) is obviously true.). By Theorem 4.7, we have that $|A| \leq \aleph_0$ or $\aleph_0 \leq |A|$; and $|B| \leq \aleph_0$ or $\aleph_0 \leq |B|$. Thus we only need consider the cases: (a) $|A| \leq \aleph_0$ and $|B| \leq \aleph_0$; (b) $|A| \leq \aleph_0$ and $\aleph_0 \leq |B|$; and (c) $\aleph_0 \leq |A|$ and $\aleph_0 \leq |B|$. (a) Since A and B are infinite, by Lemma 4.6, $|A| = |B| = \aleph_0$. Thus $|A| \leq |B|$ and $|B| \leq |A|$. (b) Since A is infinite, by Lemma 5.6, $|A| = \aleph_0$. Thus $|A| \leq |B|$. (c) Suppose that (ii) is false. Then both of the following are false: $\aleph_0 \leq |A| \leq |B|$ and $\aleph_0 \leq |B| \leq |A|$. This is a contradiction since $\aleph_0 \leq |A|$ and $\aleph_0 \leq |B|$.

(ii) \Rightarrow (i) This follows immediately from Theorem 4.7. \square

Definition 4.9. A set X has *Dedekind-cardinality*, or just *D -cardinality*, iff X is infinite and D -finite.

From Theorem 4.5, we see that in **ZFC** no set has D -cardinality. Thus, by Theorem 4.8, we can compare the cardinality of any two sets using the relation \leq . This is what we would expect: given any two sets surely we can say whether or not one is ‘bigger’ than the other? The problem comes from the fact that there are models of **ZF** in which there exist sets with D -cardinality,²¹ which means that if we do not accept **AC** then we have to accept that we can have sets whose cardinalities we cannot compare. In fact the existence of such sets leads to even more disaster:

4.2 Problems with Continuity

In **ZF**, we have the following:

Proposition 4.10. *Functions $f : \mathbb{R} \rightarrow \mathbb{R}$ may be sequentially continuous at some point $a \in \mathbb{R}$ but not continuous at a .*

Clearly this is no good. To prove this rather worrying result we will need a definition and a lemma.

Definition 4.11. Let X be a subset of \mathbb{R} . A point $a \in \mathbb{R}$ is an *accumulation point* of X iff for every $\varepsilon > 0$ the interval $(a - \varepsilon, a + \varepsilon)$ contains infinitely

²⁰Note that by the Schröder–Bernstein Theorem, which we do not prove here, if $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

²¹An example of such a model is Cohen’s original model ($\mathcal{M}1$ in [HoRu98] (pp. 146–147)).

many points of X .

Notice that an accumulation point a of X does not necessarily have to be in X . For example, 1 is an accumulation point of the open interval $(0, 1)$.

Lemma 4.12. *Every bounded and infinite subset of \mathbb{R} has an accumulation point in \mathbb{R} .*

*Proof.*²² Let X be a bounded and infinite subset of \mathbb{R} . Then there exist $m, M \in \mathbb{R}$ such that $m \leq x \leq M$ for every $x \in X$. Since X is infinite, we see that the interval (m, M) contains infinitely many points of X . Moreover, we see that at least one of the intervals $(m, \frac{m+M}{2})$, $(\frac{m+M}{2}, M)$ contains infinitely many points of X . We pick the interval which contains infinitely many points of X (if both do, we choose the interval whose infimum is smaller). Let (a_1, a_2) be this interval (so either $a_1 = \frac{m+M}{2}$ or $a_2 = \frac{m+M}{2}$). We now see that at least one of the intervals $(a_1, \frac{a_1+a_2}{2})$, $(\frac{a_1+a_2}{2}, a_2)$ contains infinitely many points of X . We pick the interval which contains infinitely many points of X (if both do, we choose the interval whose infimum is smaller). Let (a_3, a_4) be this interval. We repeat this process, each time splitting the interval in two and choosing the interval which contains infinitely many points of X . We see that the sequence (a_n) converges to an accumulation point of X . \square

*Proof of Proposition 4.10.*²³ Consider a model of **ZF** where there exists $X \subset \mathbb{R}$ that has D -cardinality.²⁴ Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the function defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in X, \\ 0 & \text{if } x \notin X. \end{cases}$$

Suppose that X is unbounded to the right. Then the set $Y := \{x \in X : 1 < x\} \subseteq X$ is infinite D -finite. The function $h : \{x \in \mathbb{R} : 1 < x\} \rightarrow \{x \in \mathbb{R} : 0 < x < 1\}$ defined by $h(x) = \frac{1}{x}$ is a bijection. We see that $h(Y)$ is bounded and infinite D -finite. We can similarly construct a bounded and infinite D -finite set if X is unbounded to the left, and so for the rest of the proof we can assume that X is bounded.

By Lemma 4.12, there exists an accumulation point $a \in \mathbb{R}$ of X . We have two cases: either $a \in X$ or $a \notin X$.

First suppose that $a \notin X$. Then $f(a) = 0$. We will first show that f is not continuous at a . Take $\varepsilon = \frac{1}{2}$. Consider some $\delta > 0$. Since a is an accumulation point of X , there exists $\varphi \in (a - \delta, a + \delta)$ such that $\varphi \in X$. So $|f(\varphi) - f(a)| = |1 - 0| = 1 > \varepsilon = \frac{1}{2}$. Therefore f is not continuous at a .

We will now show that f is sequentially continuous at a . Let (x_n) be a sequence in \mathbb{R} such that $(x_n) \rightarrow a$. Consider the set $Q := \{x_n : x_n \in X\} \subset X$. Since X is D -finite, by Lemma 4.4, Q must be finite. Thus there exists $N \in \mathbb{N}$

²²Notice that this proof does not use **AC** since we state how to construct the sequence (a_n) .

²³Part of this proof is taken from [Herr06], p. 73.

²⁴An example of such a model is Cohen's original model ($\mathcal{M}1$ in [HoRu98] (pp. 146–147)).

such that $n \geq N \Rightarrow x_n \notin X \Rightarrow f(x_n) = 0$. So $f(x_n) \rightarrow 0 = f(a)$ as $n \rightarrow \infty$. Therefore f is sequentially continuous at a .

Now suppose that $a \in X$. The set $Z := X \setminus \{a\}$ is infinite D -finite and a is an accumulation point of Z . We replace X by Z and apply the same reasoning as above. \square

5 To Choose or Not to Choose

What are we to do with this most troublesome axiom? If we accept **AC** then **BTP** puts into question our whole idea of volume in Euclidean three-space. But if we reject **AC** then we lay ourselves open to sets with incomparable cardinalities and functions in \mathbb{R} which are discontinuous yet sequentially continuous. Well, the truth is this: no one is entirely sure. Set theoretical research is still current. But this is why **AC** is so interesting—it is an open question at the very heart of mathematics.

In this essay we have only scraped the surface of the topic of **AC**. For further study into the area this author suggests [Herr06].

References

- [Arms88] Armstrong, M.A., *Groups and Symmetry* (New York: Springer, 1988).
- [Hedm04] Hedman, S., *A First Course in Logic* (New York: Oxford University Press, 2004).
- [Herr06] Herrlich, H., *Axiom of Choice* (Berlin: Springer, 2006).
- [HoRu98] Howard, P., and Rubin, J.E., *Consequences of the Axiom of Choice* (Rhode Island: American Mathematical Society, 1998).
- [Rob47] Robinson, R.M., ‘On the decomposition of spheres’, *Fundamenta Mathematicae*, 34 (1947), pp. 246–260.
- [Stol79] Stoll, R.R., *Set Theory and Logic* (New York: Dover Publications, 1979).
- [Su90] Su, F.E., ‘The Banach-Tarski Paradox’ (December 1990) <<http://www.math.hmc.edu/~su/papers.dir/banachtarski.pdf>> accessed 13 March 2007.
- [Supp72] Suppes, P., *Axiomatic Set Theory* (Toronto: Dover Publications, 1972).
- [Wag86] Wagon, S., *The Banach-Tarski Paradox* (Cambridge: Cambridge University Press, 1986).

Acknowledgement

I would like to thank Professor Dr Horst Herrlich of the University of Bremen for his help with §4.

This paper was written for the undergraduate module *MA213 Second-Year Essay*. It was typeset using MiKTeX and WinShell. Please email any comments or corrections to Daniel.Wood@warwick.ac.uk.