UNIVERSITY OF LEEDS

MATH3001

THINKING ABOUT SETS

(The Axiom of Choice)

Author
J. G. LLOYD

Supervisor
Dr. Michael Rathjen

March 26, 2020

Contents

Section 0 - Introduction

Set theory is the section of mathematical logic that studies Sets. It is a philosophical approach to looking at mathematics. It breaks down the whole of maths, and aims to rebuild it, piece by piece, with only the neccesary assumptions. Those being the concepts of a *Set*, and certain ideas relating to a set. Namely, *Memberships* and *Equality*.

Description 0 - A *Set* is a collection of objects contained together. For example, $\{x, y, z\}$ is the set consisiting of *Elements* x, y, and z. Elements are *Members* of a set.

Note that this description is non-rigorous and therefore not a definition; this is representative of the fact that we are building all of mathematics on the notions of a set. Furthermore, when attempting to define a set, we cannot be as formative as one may like. The same goes for membership and equality.
Although we cannot define these 3 concepts, we can create axioms that state precisely how they exist and what we can do with them. We may refer to these 3 concepts as the *3 basic assumptions*. The first topic we will look at is Zermelo-Frankanel Set theory (ZFC for short), a branch of set theory that does this.

Section 1 - Zermalo-Frankanel Set theory

Section 1.1 Introduction to Zermelo-Frankanel Set theory (ZFC)

Zermelo-Frankanel Set theory is a specific branch of set theory created in the early 20th century. It was developed in order to avoid paradoxes within the way we think about mathematics. The most famous one was Russell's Paradox - a paradox that defines a collection of objects that cannot exist.

Notation

- $x \in y$ means "x is a member of y".
- $x \notin y$ means "x is not a member of y".
- $x : \mathrm{R}(x)$ means "all x such that x satisfies the conditions of some relation $\mathrm{R}(x)$".

Theorem 1.1.1 - *Let R be the set $\{x : x \notin x\}$. R follows the laws of mathematics as it defines some specific boundaries on elements of a set to create a*

collection of objects. However, it cannot exist.

Proof:

- Assume, for contradiction, that $x \in x$, then, by definition of R, $x \notin x$. But if $x \notin x$, then $x \in x$. This is a contradiction. Therefore, we have that the statement 'x is a member of x' is neither true nor false, yet this is impossible, so the set can't exist.

An example of the paradox being used in a real life scenario is the Barber's paradox; a nice way of portraying the paradox in a less abstract way.

Theorem 1.1.2 - *The Barber's Paradox* - There exits one barber. The barber shaves those, and only those, who do not shave themselves. This is not possible.

Proof:

If the barber shaves himself, then he is not in the collection of people who the barber shaves, but then he does not shave himself. If he does not shave himself, he shaves himself; by the conditions given about who the barber shaves. Again, a contradiction.

Although Russell's paradox is a paradox in mathematics, it is not a paradox in ZFC, as we will soon find out. ZFC aims to avoid problems just like this one, and to have a system of looking at mathematics without flaws. Like mentioned previously, ZFC has 9 basic axioms to base it's ideas on:

Remark - As we go through stating the axioms, we will define certain concepts and notions as and when relevant, to either remind or inform the reader of the necessary information needed to understand each axiom, [illegible]

Axiom 1.1.3 - *The Axiom of Separation* - Let A be a set, and P be a property. The collection $\{x : x \in A \text{ and } \mathrm{P}(x)\}$ is a set.

Notice that Axiom 1.1.3 does not succumb to Russell's paradox, because in Russell's paradox, no set A is defined for x to be in.

Axiom 1.1.4 - *The Empty Set Axiom* - There exists a set consisting of no members. ($\exists A\ \forall x(x \notin A\ )$). This set is denoted ø.

Axiom 1.1.5 - *The Axiom of Extensionality* - Two sets are equal if and only if they have the same members.

Definition 1.1.6 - The *ordered pair* (x, y) of elements x and y is the set $\{\{x\}, \{x, y\}\}$.

One may assume it to be unnecessarily tedious to define such a simple notion of a an order pair in such a way. Why not define it as $\{x, y\}$? The logic behind this is that defining the pair as a set allows us to express the fact that the first element of the ordered pair is the element existing in both $\{x\}$ and $\{x, y\}$ and the second element of the ordered pair is the element existing in only $\{x, y\}$. If we kept it as $\{x, y\}$, this would not define an order.

Axiom 1.1.7 - *The Pairing Axiom* - An unordered pair $\{x, y\}$ of 2 sets x and y is a set.

Section 1.2 - Encodings and Unions

Before we define the next axiom, let us remind ourselves of the notion of union:

Definition 1.2.1 - The *Union* of a set A denoted $\bigcup$A is the collection of all the members of the members of A

Examples:
- D := $\{x, y, z\}$, $\bigcup$D := $\emptyset$
- O := $\{x, y, y, y, z, \{y, z\}\}$, $\bigcup$O := $\{y, z\}$
- N := $\{\emptyset, \{\emptyset\}, \{x, \{\{x, y\}\}\}\}$, $\bigcup$N := $\{\emptyset, x, \{\{x, y\}\}\}$

Axiom 1.2.2 - *The Union Axiom* - For every set A, $\bigcup$A is a set.

We can conclude from this axiom, that sets D, O, N are all sets, and from the pairing axiom, can also conclude that {D, O, N} is a set.

An important feature to grasp with ZFC is that all members of sets are sets. All ideas we assemble are formed from ingredients that have a route of our 3 basic assumptions from *Description 0*. It is vital to understand that any element is a set, or at least the representation of it is a set, as what else could it be? For example, 1 is only what we choose to represent/define it as. sinh(510) is what what we choose to define it as. An important and often confusing aspect of ZFC is the extent to which we decide what the rules are, and how ZFC requires a more philosophical way of thinking about maths from what most people are used to. Although we cannot have the number 1 in the form of 1 unit of an arbitrary measurement, we can encode it using sets, representing it in the form some set we choose to assign to it. The same goes for any other object.

For instance, ZFC looks at and encodes the natural numbers in the following way:

Encryption 1.2.3 - Let us Encode the natural number 0 as the empty set ø. This intuitively makes sense, as the empty set has 0 elements, so there is a sensible connection between the two objects here.

- Encode the what we know as the natural number 1 as the set of the empty set {ø}.

- Encode the natural number 2 as the set of the set of (the empty set and the set of the empty set) {ø, {ø}}
...
- Encode the natural number n as the set of n-1.

We can look at this sequence of numbers (0, 1, 2, 3, ... , n, ...) as (ø, {ø}, {ø, {ø}}, ...). Namely, 0 := ø, 1 := {0} := {ø}, 2 := {0, 1} = {ø, {ø}}, 3 := {0, 1, 2} := { ø, {ø}, {ø, {ø}}}, etc.

The encryption above sums up quite nicely how we are able to encode numbers in terms of sets in an efficient and rigorous way. We are inducing a system such that for any natural n, m; n $<$ m if and only if n $\in$ m.

Definition 1.2.4 - The *(pairwise) union* of a two sets A and B denoted A $\cup$ B is the set containing all elements that exist in A or B (or both).
(A $\cup$ B := $\{x : x \in$ A or $x \in$ B$\}$)

Section 1.3 - Infinity, Power set and Replacement

Axiom 1.3.1 - *The Axiom of Infinity* - There exists a set A such that ø $\in$ A and for all $x \in$ A, $x \cup \{x\} \in$ A. ($\exists I$(ø $\in I$ and $\forall$ x (x $\in I \Rightarrow$ x $\cup$ {x} $\in I$)). (This is the same as saying there exists a set that has a non-finite amount of elements).

Definition 1.3.2 - *Subset* - Let A, B be sets. A is a subset of B if every member of A is also in B. When this the case, we say A $\subseteq$ B.

Definition 1.3.3 - *Powerset* - Let A be a set. The Powerset P(A) is the collection of all Subsets of A. P(A) := {B : B $\subseteq$ A }

Axiom 1.3.4 - *Powerset Axiom* - For every set A, P(A) is A set.

Examples:

- R := $\{y, z, \{z\}\}$ is a set $\Rightarrow$ P(R) := $\{\{y, z, \{z\}\}, \{z, \{z\}\}, \{y, \{z\}\}, \{y, z\}, \{z\}, \{\{z\}\}, \{y\}, \emptyset\}$ is a set.

- A := $\{x, \emptyset, \{\emptyset\}\}$ is a set $\Rightarrow$ P(A) := $\{\{x, \emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}, \{x, \{\emptyset\}\}, \{x, \emptyset\}, \{\{\emptyset\}\}, \{\emptyset\}, \{x\}, \emptyset\}$ is a set.

Axiom 1.3.5 - *Axiom of Replacement* - Let P be property in two variables and A be a set. (For every $x \in$ A there exists y such that $P(x, y)$ holds) $\Rightarrow$ The collection $\{y : P(x, y)$ holds for some $x \in$ A$\}$ is a set.

Axiom 1.3.6 - *The Axiom of Foundation* - Let S be a non-empty set. There exists $x \in$ S such that for every $y \in$ S, y $\notin x$.

Example:

- S := 3 := $\{0, 1, 2\}$. $x := 0 := \emptyset$ is our x that fits the axiom here. This is because For every $y \in \{0, 1, 2\}$, $y \notin \emptyset$.

Definition 1.3.7 - The Cartesian product A $\times$ B of two sets A and B is the collection of ordered pairs who's first entry is a member of A and second entry is a member of B.

Remark 1.3.8 - The Cartesian product A $\times$ B of two sets A and B is a set.

There is one more axiom, the *Axiom of Choice*, a very important axiom in ZFC. It is considered an additional axiom to the 9 basic ones It gives rise to many philosophical ideas and questions. It is historically controversial, and we will delve into it in Sections 3, 4 and 5.

Section 2 - Category theory and ETCS

Section 2.1 Introduction to Category theory

Although Category theory is not a branch of set theory, it has a similar idea, the idea to rebuild mathematics from a few axioms and rules, in order to represent it in a less flawed, more elegant manner. Category theory is a mathematical structure that aims to formalise mathematical notions in terms of nodes and edges, labelled objects and arrows respectively. There are two main ideas associated with a category. The first being that a category can compose arrows associatively. The second, being that for all objects, there exists an identity arrow. Let us define a category more formally:

Definition 2.1.1 - A *category* C is a mathematical structure that has the following properties:

- A collection $ob(C)$ of *objects*

- For every $A, B \in ob(C)$ there exists a collection $C(A, B)$ of *arrows* from A to B

- For every $A, B, D \in ob(C)$, there exists a *Composition Function*;

$$C(A, B) \times C(B, D) \Rightarrow C(A, D)$$
$$(f, g) \Rightarrow g \circ f$$

- For every $A \in ob(C)$ there is an *identity* arrow $1_A \in C(A, A)$.

- We write $\mathrm{Hom}_C(A, B)$ and f : $A \Rightarrow B$ to refer to $C(A, B)$and f $\in C(A, B)$ respectively.

There are 2 main axioms that Category theory obeys:

Axiom 2.1.2 - *The Associativity axiom* - For all functions f : P $\Rightarrow$ Q, g : Q $\Rightarrow$ R, h: R $\Rightarrow$ S, we have h $\circ$ (g $\circ$ f) := (h $\circ$ g) $\circ$ f

Axiom 2.1.3 - *The Unit axiom* - For all functions f : A $\Rightarrow$ B, we have that 1_B $\circ$ f := f and f $\circ$ 1_A := f.

Let us now introduce some examples of categories:

- The Category of **Set** has it's objects representing '*sets*' and it's arrows representing '*functions*'.

- The Category of **Grp** has it's objects representing '*groups*' and it's arrows representing '*group homomorphisms*'.

- The Category of $\mathbf{Vect}_k$ has it's objects representing '*vector spaces over a field k*' and it's arrows representing '*linear maps*'.

- The Category of **CRing** has it's objects representing '*rings*' and it's arrows representing '*ring homomorphisms*'.

Definition 2.1.4 - An arrow f : $A \Rightarrow B$ is an *Isomorphism* iff there exists g : $B \Rightarrow A$ such that fg:=1_B and gf :=1_A. We denote this g as the inverse function f^{-1}. We denote an Isomorphism f : $A \Rightarrow B$ as $A \cong B$

Definition 2.1.5 - Let C be a category, I be an object. I is called an *Initial object* if for every $A \in C$ there is exactly one arrow $I \Rightarrow A$.

Definition 2.1.6 - Let I be an object, C be a category. I is called a *terminal object* if for every A $\in$ C there is exactly one arrow $A \Rightarrow I$.

Category Theory has applications within Computer Science, such as in *Programming Language theory*. However, we shall not explore this in this paper.

Section 2.2 Introduction to ETCS

The *Elementary theory of the Category of Sets (ETCS)* is a formulation of Set theory that uses categories instead of sets to express the building blocks of mathematical structures. It was created by Lawvere in 1964 as an alternative to ZFC because ZFC was considered to have a major flaw. In ZFC, the use of the word 'set' is, by some, considered unhelpful and contradictory. In ZFC, the elements of sets are always sets too. So, take the set of the real numbers. If the elements of the real numbers are always sets, then π is a set. Let us introduce the question "What are the elements of π?". This question has no intuitive answer. Perhaps one might say that this means π has no elements. But Axiom 1.1.5 (The Axiom of Extensionality) tells us that two sets are equal if and only if they have the same members. This means $\pi := \emptyset$. This doesn't make sense, furthermore being argued as a flaw in ZFC, the motivation for ETCS. To avoid this flaw, ETCS avoids looking at membership, and uses the two main notions of functions and sets.

ETCS, like ZFC, has a few axioms, let us define these as follows:

- **Axiom 2.2.1** - The *composition* of 2 functions is associative, and has an identity. *(See Axiom 2.1.2 and 2.1.3)*

- **Axiom 2.2.2** - There exists a set with exactly one element (A *Terminal set*).

- **Axiom 2.2.3** - There exists a set with exactly no elements. Namely the empty set $\emptyset$.

- **Axiom 2.2.4** - Let X and Y be sets and f, g: $X \Rightarrow Y$ be functions. Suppose that f(x) := g(x) for every $x \in X$. Then f := g.

- **Axiom 2.2.5** - For all sets X and Y, their cartesian product $X \times Y$ exists.

- **Axiom 2.2.6** - For all sets X and Y , there is a function set g: $X \Rightarrow Y$

- **Axiom 2.2.7** - For all functions f : $X \Rightarrow Y$ and elements $y \in Y$, there exists an inverse image of y under f; namely $f^{-1}(y)$.

- **Axiom 2.2.8** - The subsets of a set X correspond to the functions from X to {0,1}

- **Axiom 2.2.9** - The natural numbers form a set.

- **Axiom 2.2.10** - Every surjection has a right inverse. (A right inverse of a surjection S: $X \Rightarrow Y$ is a choice, for every $y \in Y$ of an element of the nonempty set $S^{-1}(y)$).

This concludes our axioms. We will not expand on this within the paper.

Section 3 - The Axiom of Choice and It's Equivalences

Section 3.1 The Axiom of Choice and The Well-ordering theorem

Axiom 3.1.1 - *The Axiom of Choice* - If A is pairwise disjoint set of non-empty sets, there exists a set B such that which consists of exactly one element from each set in A.

Russell, B. (1919) highlights, in "Introduction to Mathematical Philosophy" that to choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, whereas for shoes, the Axiom of Choice is not required. The idea here is that left and right socks are indistinguishable, while left and right shoes are distinguishable. This means that the axiom of Choice is needed to show that an element can be taken from each set of socks. The axiom of choice is not needed with shoes, because you can define the new set as 'each left shoe' or 'each right shoe'.

The Axiom of choice was formulated in 1904 by Ernst Zermelo in order to formalize his proof of the Well-ordering theorem. *ZFC* is precisely the axioms defined in section 1 with the Axiom of Choice, and *ZF* is these axioms, but without the Axiom of choice.

Definition 3.1.2 - Let A be a set. Let $<$ be a binary relation on A. A *(Totally) ordered set* is a pair (A,$<$) with the following conditions:

- Anti-reflexivity - ($x \not< x$) for every $x \in$ A.
- Transitivity - (($x < y$, $y < z$) $\Rightarrow x < y$) for all $x, y, z \in$ A
- Totality - For all $x, y \in$ A, either $x < y$ or $y < x$.

Note - A non-strict order, (A,$\leq$) will be defined as $x \leq y$. This is when $x < y$ or $x := y$.

Definition 3.1.3 Let B $\subseteq$ A be a subset of an ordered set (A,$\leq$). The *least element* of B is the unique $x \in$ B such that $x \leq y$ for every $y \in$ B. An ordered set (A,$<$) is *Well ordered* if every non-empty subset of A has a least element.

Theorem 3.1.4 - *The Well-ordering theorem* - Given any set A, there exists a well-order $<$ on A.

When observing certain sets, it is clear that they are well ordered. For example, the natural numbers under the ordering $<$ is well ordered. ($1 < 2 < 3 < 4$ etc, and 1 is the least element). This seems easy. However, if we took, for example, the integers without zero under that same order, we would find that there *is* no least element here. Let us take the following ordering $<^*$ on the non-zero integers:

$a <^* b$ if
either $|a| < |b|$ and $a \neq -b$
or $|a| = |b|$ and $a < 0$ and $b > 0$

The integers without zero are well ordered here, as (-1 $<^*$ 1 $<^*$ -2 $<^*$ 2 $<^*$ -3 etc).

Although we may believe that it would be possible to brew up a well-order for any set, we can of course not be completely convinced until we have a proof. Ernest Zermelo gave the first proof of the Well-ordering theorem, back in 1904, under the assumption of the axiom of choice. He considered the proof to be a "Fundamental law of thought of great consequence, particularly noteworthy for it's universal validity" (Heijenoort, V. (1967), p. 139). Zermelo proves, in fact, that the Axiom of Choice implies the Well-ordering theorem. The following is is an outline of Zermelo's first proof of the Well Ordering theorem:

Theorem 3.1.5 - *Axiom of Choice* $\Rightarrow$ *Well-ordering theorem (Zermelo):*

Proof: (Heijenoort, V. (1967), pp. 140-141)

Let M be some arbitrary set. Denote the cardinality (size) of M as $|M|$. Let m be an arbritrary element of M. Let M' be a non-empty subset of M with

cardinality $|M^{'}|$. Define $M/M^{'}$ as the subset of M "complimentary" to $M^{'}$. Here, we have partitioned M into 2 distinct subsets; $M^{'}$ and $M/M^{'}$.

- *(Recall that two subsets $M^{'}$, $M^{''}$ of M are distinct if there exists an element in either $M^{''}$ or $M^{'}$ but not both.)*
- *(Recall that the set of all subsets $M^{'}$ of M as P(M) : Definition 1.4.3 - Powerset)*

Suppose that for every $M^{'} \subseteq M$ there exists exactly one unique arbitrarily chosen associated element $m_1^{'} \subseteq M^{'}$. Define this chosen $m_1^{'}$ as the distinguished element of $M^{'}$. Here, we need the axiom of choice, because otherwise there would be no way to 'choose' an element arbitrarily from a set.

Let us define our function for this choice as the following 'Choice Function':

$\gamma : \{P(M)/\emptyset\} \to M$ such that for every $M^{'} \in P(M)$ we have that $\gamma(M^{'}) \in M^{'}$. Here all this function is doing is taking some subset $M^{'} \subseteq M$, and returning an element of that subset. Let us introduce an example of the above, to help with intuition:

Examples:

- $M := \{1,2,3,4,5\}$, $M^{'} := \{2,4\}$. $\gamma(M^{'}) := 4$
- $M := \{x \in \mathbb{R} : 3x \in \mathbb{Z}\}$, $M^{'} := \{\mathbb{Z}\}$, $\gamma(M^{'}) := 507$

Now, let us introduce the notion of a *γ-set*.

Definition 3.1.6 - Fix some γ, so every subset of M is mapped to a member of m, the 'distinguished element'. A *γ-set* is a set M^{γ} that has the following properties:

- (a) $M^{\gamma} \subseteq M$

- (b) M^{γ} is well ordered for some $<$

- (c) Whenever a is an element of M^{γ}, and A is the unique set which consists of the elements $x \in$ M such that x $< a$ (call this the *Associated Segment* of M), we have that a is precisely the distinguished element of $M/$A. In other words, $a = \gamma(M/\text{A})$

We may think of this as a recipe to formulate the choice of our distinguished element. Let us give an example of this to help visualise it:

Examples:

(Let M be a set.)

- Let m be a member of M, existing such that $\gamma(M)$=m. Then take $\{m\}$. We aim to show that this is a γ-set. ((a) is clearly satisfied by definition of m). This is well-ordered by $<^*$ such that for every $x \in M$, $x \not<^* m$. ((b) is then satisfied). So our associated segment 'A' here is $\emptyset$. Given that $\gamma(M/\text{A}) = \gamma(M/\emptyset) = \gamma(M) = m$, we have that (c) is satisfied, and therefore, $\{m\}$ is a γ-set.

- Take the unique $a, b \in M$ such that $\gamma(M/\{a\}) = b$ and $\gamma(M) = b$. Take the set $\{a, b\}$. Let us aim to show that this is a γ-set. (criteria (a) is clearly satisfied by the definition of a and b). This set is well-ordered by the ordering $<^k$ where $x <^k y$ if and only if $x = a$ and $y = b$ (criteria (b) is therefore satisfied). When we take b, we see that the associated segment A_b is {a} (precisely all the elements x in M that satisfy x $<^k$ b). Here $\gamma(M/A_b) = \gamma(M/\{a\}) = b$. When we take a, we see that the associated segment A_a is $\emptyset$. So here $\gamma(\text{M}/A_a)$ $= \gamma(M/\emptyset) = \gamma(M) = a$. Thus, we have that both elements of $\{a, b\}$ fit part (c) of the criteria for a gamma set, so $\{a, b\}$ is a γ-set.

Following the notion of γ-sets, Zermelo was able to show the following 3 remarks:

Remark 1: Take 2 distinct γ-sets; M'^{γ} under the well-ordering $<'$ and M''^{γ} under the well ordering $<''$. Then one of the two sets is isomorphic to an initial segment of the other.

Remark 2 - If 2 γ-sets have an element in common, call it b, then the associated segment for the two γ-sets are the same.

Remark 3 - If 2 γ-sets have 2 elements a and b, then we either have a $<$ b in both sets or b $<$ a in both sets.

Note - We say that x is a γ-element if $x \in M'^{\gamma}$ for some M'^{γ}

Let us define L $= \bigcup M^{\gamma}$. Namely, L is the union of all γ-sets. We claim, that L is a γ-set, and more specifically, L $= M$. Assume, for contradiction, that L does not equal M. This means we can let $a = \gamma(M/\text{L})$, a mere output of the choice function 'choosing' an element that is in M but not L. But then L $\cup \{a\}$ is a γ-set because it ticks all 3 boxes that define a γ set:

(a) L $\subseteq M$ and $\{a\} \subseteq M$ so clearly L $\cup \{a\}$ is a subset of M.

(b) As L is a γ set, it is well-ordered on some $<$, so define $<^k$ to be the ordering that lets $l <^k a$ for all $l \in$ L, and $l, b \in \text{L} \rightarrow l <^k b$ if $l < b$. Therefore L $\cup \{a\}$ is well ordered (under $<^k$).

(c) Take an arbitrary element of L $\cup$ $\{a\}$. If the element taken is a, then we know a determines the set (the associated segment) A_a = {x $\in M$: x$<^k$a} such that $a = \gamma$(M/A). Notice A is merely just L here. If the element is not a, namely it is b for some $b <^k a$ then we have that b is a member of L, which means that b is an element of some γ-set. So b determines A_b (the associated segment) = $\{x \in$ M : $x < b\}$ and $b = \gamma(M/$A). We therefore have that any element of L $\cup$ $\{a\}$ suffices criteria (c) of being a γ-set, and so is itself a γ-set.

So we have that L $\cup$ $\{a\}$ is a γ-set, but that means that L is not the largest such γ-set and so clearly cannot be the union of all γ-sets. This is a contradiction so we can confirm that L = M. We know that L is a γ-set, and is therefore well-ordered, so if L = M, then we have that M is well-ordered. Therefore, every set can be well ordered. □

Zermelo's proof was widely respected but also created much discussion as whether or not it presupposes too much. It is for this reason that in 1908, Zermelo introduced a second proof of the Well-ordering theorem, one he believed to rely less on the assumptions of set theory. It was considered to have more mathematical validity, and his result was generally accepted by the mathematical world. (Heijenoort, V. (1967))

Section 3.2 A Cycle of Equivalences of the Axiom of Choice

There are a variety of statements considered equivalent to the axiom of choice. Some different versions can be more useful when proving certain theorems, as we will see in chapter 6. First, we will look at how the Well-ordering theorem implies the axiom of choice. This, combined with Zermelo's proof of the converse in Section 4.1, suffices to show how the 2 statements are logically equivalent.

Theorem 3.2.1 - *Well-ordering theorem* $\Rightarrow$ *The Axiom of Choice*

Proof: (Mantova, V.L. (2019))

Let A be a set of non-empty sets. Let B = $\bigcup$A (A set by The union Axiom). As B is a set, there exists some ordering $<$ such that (B,$<$) is well-ordered (by the well ordering theorem). Pick some $x \in$ A. All members of x are members of members of A so x is a non-empty subset of A. Now define f(x) as the function sending x to the least element of x under the ordering $<$. We can define f as f = A$\times$Y= $\{z : z = (x, y)$ for some $x \in$ A, $y \in$ Y$\}$ where Y = $\{y \in$ B : Y is the least element of B under $<\}$. Y is a set by Separation, and A$\times$Y is a set by Remark 1.4.8. This function is a choice function and therefore represents the existence of the Axiom of Choice. □

We have shown how the Axiom of choice and the Well-ordering theorem imply each other, but now we shall introduce some more concepts that also imply the Well ordering theorem. The following chain reaction occurs when looking at the statements that imply The Axiom of choice:

Hausdorff's maximal principle $\Rightarrow$ *Zorn's Lemma* $\Rightarrow$ *Well ordering theorem* $\Rightarrow$ *Axiom of Choice* $\Rightarrow$ *Hausdorff's maximal principle.*

Before we define these statements, let us introduce some useful definitions:

Definition 3.2.2 - A *Partially Ordered Set* (Or *Poset*) is a pair (P,$\leq$) where A is a set, $\leq$ is a binary relation on A, and the following hold:

(a) *Reflexivity*: $a \leq a$ (For every $a \in$ P)

(b) *Antisymmetry*: ($a \leq b$ and $b \leq a$) $\rightarrow a = b$ (For every $a, b \in$ P)

(c) *Transitivity*: ($a \leq b$ and $b \leq c$) $\rightarrow a \leq c$ (For every $a, b, c \in$ P)

Definition 3.2.3 - Let (P, $\leq$) be a Poset. Let A $\subseteq$ P. An element $p \in$ P is an *Upper Bound* of A if for every $a \in$ A, we have that $a \leq p$.

Definition 3.2.4 - A *Chain* in a Poset (A, $\leq$) is a subset C$\subseteq$A that is Ordered on $\leq$. A chain C is *Maximal* if no proper superset D $\supset$ C is a chain in (A, $\leq$).

We will now introduce two equivalences of the axiom of choice:

Theorem 3.2.5 - *Hausdorff's maximal principle (HMP)* - Any chain in a Poset can be extended to a maximal Chain.

Lemma 3.2.6 - *Zorn's Lemma (ZL)* - Let (A, $\leq$) be a Poset such that each chain in A has an upper bound. Then (A, $\leq$) has a maximal element.

Theorem 3.2.7 - *Hausdorff's maximal principle (HMP)* $\Rightarrow$ *Zorn's Lemma*

Proof: (Winfried, J. and Weese, M. (1996))

Assume the existence of HMP. Let (A, $\leq$) be a poset such that every chain in A has an upper bound. We know $\emptyset$ is a chain in A because $\emptyset \subseteq$ X for any set X, and $\emptyset$ is ordered under $\leq$. By HMP, there exists a chain within A that extends $\emptyset$ to a maximal chain. Fix this chain as C, and let x be an upper bound of C. Suppose, for contradiction, that (A, $\leq$) has no maximal element. Then there exists $y \in$ A such that $x < y$. Then C $\cup \{y\}$ is a chain of A that is a proper superset C $\cup \{y\} \supset$ C. Then C is not maximal, a contradiction. Hence

(A, $\leq$) must have a maximal element. $\square$

Now let us consider the next step in our sequence of implications:

Theorem 3.2.8 - *Zorn's Lemma* $\Rightarrow$ *Well-ordering theorem.*

Proof: (Foland, G.B. (1984))

Let X be a set. If X = $\emptyset$ then the Well ordering theorem clearly holds, so let X be non-empty. Let W $\subseteq$ P(X) be the set of well ordered subsets sets of X. Then W is partially ordered by the ordering $\subseteq$. More specifically, we can say that for all w_1,$w_2 \in$ W, $w_1 \subseteq w_2$ iff w_1 is an initial segment of w_2. There is therefore an upper bound for all chains within W. Namely, the union of all chains in W. Thus the premise within Zorn's Lemma holds, so it can be deduced that W has a maximum element. Suppose, for contradiction, that this is not a well-ordering on X itself. Let (B,$\subseteq$) be the maximal well-ordering on X. Assume $x_0 \in$ X\B. We can now define an ordering $\leq^{'}$ on X $\cup \{x_0\}$ as follows:

$\leq^{'}$:=($x \leq y$ for $x, y \in$ B) and ($x_0 \leq x$ for $x \in$ B and $x_0 \in$ X\B)

But then $\leq^{'}$ extends $\leq$ which contradicts the maximality of $\leq$. Therefore, the Well ordering of W is a well-ordering of X, and so X is well ordered. $\square$

Now, to connect the chain of implications, let us prove the following:

Theorem 3.2.9 - *Well-ordering theorem* $\Rightarrow$ *Hausdorff's maximal principle*

Proof: (Dudley, R.M. (2018))

Let (X, $\leq$) be a partially ordered set. We want to find a maximal chain, given some chain. Let W be a well-ordering on X (We may assume this due to the Well Ordering Theorem). We may label X as $(x_1, x_2, x_3, ...)$ where $x_1 W x_2 W x_2 W ...$ We may recursively define a function f such that $f(x_1) = (x_1)$ and

$$f(x_i) = \begin{cases} x_i & \text{if } \{x_i\} \cup \{f(x_j) : x_j W x_i \text{ and } j < i\} \text{ is a chain for } \leq \\ x_1 & \text{Otherwise} \end{cases} \tag{1}$$

The range of f is precisely a maximal chain of the chain containing x_1. Why is this? Well, what the function is essentially doing is looking at each element of X, one by one, in the order that is precisely W, and asks, "is this element in the same chain as x_1?" If it is, then the element is added to the range, and if not, then it isn't. The maximal chain is precisely the elements in the range. Hence, any chain in a Poset can be extended to a maximal chain. $\square$

We have now completed the following cycle of proofs: *Hausdorff's maximal principle* $\Rightarrow$ *Zorn's Lemma* $\Rightarrow$ *Well ordering theorem* $\Rightarrow$ *Axiom of Choice* $\Rightarrow$ *Hausdorff's maximal principle.*

Section 3.3 - Tukey's Lemma

Finally, let us introduce another Lemma in our sequence of implications, with need of a definition:

Definition 3.3.1 - Let F be a family of sets. F has *Finite Character* if for every set A, we have that $A \in F$ if and only if every finite subset of A is also in F.

Lemma 3.3.2 - *Tukey's Lemma (TL)* - If F is a non-empty family of sets with finite character, then F has a maximal element with respect to $\subseteq$.

Theorem 3.3.3 - *Zorn's Lemma* $\Rightarrow$ *Tukey's Lemma.* (Barnum, K. (2013), p. 3)

Proof:

(Let A be a non-empty family of sets with finite character. Due to its finite characterisation, A is partially ordered by $\subseteq$. For some chain C in A, let $B = \bigcup\{X : X \in C\}$, i.e. B is precisely the set of elements that lie within sets that are in the chain. But then due to the nature of the chain, specifically the partial ordering under $\subseteq$, we have that every finite subset of B is in the chain. This implies B is in A (again due to the finite characterisation). B is then clearly an upper bound of C. Therefore, by Zorn's lemma, A has a maximal element under $\subseteq$. $\square$

Theorem 3.3.4 - *Tukey's Lemma* $\Rightarrow$ *Zorn's Lemma*

Proof:

Assume Tukey's Lemma holds. Let $(P, \leq)$ be a poset in which every chain has an upper bound. We aim to show that $(P, \leq)$ has a maximum element. Let F be the family of chains in $(P, \leq)$. Every finite subset of a chain is a chain, so F has finite character. So, by Tukey's lemma, F has a maximal element C with respect to the ordering $\subseteq$. C is a chain in $(P, \leq)$, so has an upper bound p (by the conditions defined on $(P, \leq)$). As C is maximal, we have that $p \in C$. This is because if $p \notin C$ then we would have that $C \subseteq C \cup \{p\}$, contradicting the maximality of C in $(F, \subseteq)$. Therefore, p is the maximal element of the maximal chain in $(P, \leq)$, so $(P, \leq)$ has a maximal element, namely p. $\square$

You may be asking about the relevance of Tukey's lemma? We already had a nice circle of equivalences of the axiom of choice. As we will see in the next section, Tukey's Lemma can be a useful tool when looking at Vector Spaces.

Section 4 - Applications of the axiom of choice

Section 4.1 - Vector Spaces

We will now go in to some of the applications of The axiom of choice, the first being the use of it within Linear Algebra, specifically, *Vector Spaces*.

Definition 4.1.1 - A *Vector space* is a set V over a field F with binary operations for multiplication and addition such that for every $u, v \in$ V, and every scalar $a, b \in$ F:

- $u + v = v + u \in$ V
- $(u + v) + w = u + (v + w)$
- There is a vector $0 \in$ V such that $u + 0 = u$.
- There is a vector $(-u) \in$ V such that $u + (-u) = 0$
- $a(u) \in$ V, $a(u + v) = au + av$, and $(a + b)(u) = au + bu$
- $a(bu) = (ab)u$
- $1u = u$

Definition 4.1.2 - A sequence of Vectors $(v_1, v_2, ..., v_k)$ from a Vector Space V is linearly *Linearly Independent* if it is not the case that there exist scalars not all zero $a_1, a_2, ...ak$ such that $a_1v_1 + a_2v_2 + ... + a_kv_k = 0$. Namely, $a_1v_1 + a_2v_2 + ... + a_kv_k = 0$ implies $a_i = 0$ for every i $\in$ {1,...,k}

Definition 4.1.3 - If every vector in a vector space V can be written as a linear combination of vectors in a given set S, then S is called a *spanning set* of V, (S *spans* V or *Span*$(S) = V$).

Definition 4.1.4 - A set B of vectors in a vector space is called a *Basis* if B spans V and B is linearly independent.

Theorem 4.1.5 - *Every Vector space has a basis*

We will introduce 2 proofs here, highlighting the importance of having numerous equivalents of the Axiom of Choice. The first proof incorporates Tukey's Lemma:

Proof 1: (Barnum, K. (2013), p. 5)

Let V be a vector space and let F be the family of all linearly independent sets of vectors in V. A set A is in F iff any finite subset of A is also in F, so F has finite character. This means, by Tukey's Lemma, that there exists some maximal linearly independent set B of vectors. As B is maximal, it is clearly a basis. □

This proof was nice and short, a quick application of Tukey's Lemma. But let's assume we didn't have Tukey's lemma. Then the proof would be much more tedious:

Proof 2: (Khatchatourian, I (2018), p. 7)

Let V be some vector space. Let P be the set of vectors A such that A $\subseteq$ V and A is linearly independent. P is partially ordered by Inclusion. We can notice that P $\neq \emptyset$ because any singleton that is a subset of the vector space is linearly independent. For this proof we introduce 2 lemmas.

Lemma 4.1.6 - *Let B be a maximal element of P. Then B is a basis for V.*

Proof: (Khatchatourian, I (2018), p. 7)

- Suppose that B is a maximal element of P under inclusion. B is linearly independent, so all that is left to show is that B spans V. Suppose, for a contradiction, that B does not span V. Namely there exists some $v \in$ V such that $v \notin$ Span(B). But then take C = B $\cup \{v\}$. C is clearly linearly independent as it is the union of 2 linearly independent sets (any singleton is linearly independent). This means C $\in$ P. It is easy to see that B $\subseteq$ C, but then this contradicts the fact that B is maximal. Therefore, B is a basis for V. □

Now, if we can show that every chain in P has an upper bound, we can combine Lemma 1 with the premise of Zorn's lemma to get to our final conclusion (That Every Vector space has a basis)

Lemma 4.1.7 - *Every chain in P has an upper bound.*

Proof: (Khatchatourian, I (2018), p. 8)

- Let C be a chain in P. Recall that a chain in P is a totally ordered subset of (P, $\subseteq$). So for all C_1,$C_2 \in$ C, $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$. C is a collection of sets of linearly independent vectors v $\in$ V. It is obvious that any empty chain has an upper bound, so let us assume that C is non-empty. Let X = $\bigcup$C. Let us claim that X is the upper bound for C that we are looking for. We need to show that for every $C_i \in$ C, $C_i \subseteq$ X, and also that X $\in$ P. As X is by definition

the set of the members of the members of C, it is clear that any member of C will be a subset of X. Now let us claim, for contraction that X $\notin$ P. So X is not linearly independent, and therefore X is a linearly dependent set of Vectors $\{v_1, v_2, ...v_n\}$ for some n.

Remark - One may assume that $n \in \mathbb{N}$ because If a set of vectors is linearly dependent, it is witnessed by finitely many elements. We shall not delve in to the proof of this here.

By definition of X, any member of X is a member of some member C_i of C. Given that C is a finitely long chain, there must exist some $C_k \in C$ such that $C_i \subseteq C_k$ for all C_i in C. But then as each v_i is a subset of some C_i, where each C_i is a subset of C_k, this means that $\{v_1, v_2, ...v_n\} \subseteq C_k$. But then C_k is linearly dependent which contradicts the fact that $C_k \subseteq$ P. Therefore, X $\in$ P.
So we have shown that X $\in$ P and that for every $C_i \in$ C, $C_i \subseteq$ X. Therefore X is an upper bound on C, and so every chain in P has an upper bound. $\square$

We now know from Lemma 2 that Every chain in P has an upper bound. Given that P is partially ordered an non-empty, we can deduce using Zorn's Lemma that P has a maximal element M. Then by Lemma 1, we can conclude that M is a basis for V. Therefore, Every vector space has a basis. $\square$

The difference in lengths of these two proofs really do highlight the handiness of having different equivalences of the axiom of choice. Clearly it is quicker to use Tukey's Lemma in the above case. However, in cases like the following, Zorn's Lemma is the best tool possible:

Theorem 4.1.8 - *Every spanning set of a non-zero vector space V contains a basis of V.*

Proof: (Conrad, K. (No date) p. 16)

Let V be a non-zero vector space. Let S be a spanning set of V. Let L be the set of linearly independent subsets of S. We can partially order L by inclusion. Let K be a subset of L that is totally ordered by inclusion. Take $\bigcup$K. $\bigcup$K is an upper bound on every $k_i \in$ K. This means that any totally ordered set of the partially ordered set L has an upper bound. Therefore, by Zorn's Lemma, L has a maximum element B. B is precisely a maximal linearly independent subset of S with respect to inclusion.

We know B is linearly independent, as it is a member of L, so if we can show that B spans V, then we can conclude that B is a basis for V (by the definition of Basis).

We know by definition of S that S spans V. If we can show that every element in S is in the span of B, then we can conclude that B spans V.

Assume, for contradiction, that there exists $v \in$ S that is not an element of the span of B. Then B $\cup \{v\}$ is a linearly independent set (as it's the union of 2 linearly independent sets). It is therefore a linearly independent subset of S (as all elements in B and in $\{v\}$ are in S). But then B is not the maximal element of L, which is a contradiction. Therefore, for every $x \in$ S, we have that x is in the span of B, so B spans V and therefore B is a basis for V.

So we have that for any V there exists B such that B spans V and B is linearly independent. Therefore, Every spanning set of a non-zero vector space V contains a basis of V. □

Section 4.2 - Ring Theory

We will now briefly touch on the application of The axiom of choice within Ring Theory.

Definition 4.2.1 - The reader may remind themselves of the notion of a *Ring* and and an *Ideal* in (Cohn, P.M. (2000)).

Theorem 4.2.2 - *Every non-zero commutative ring contains a Maximal Ideal.*

Proof: (Conrad, K. (No date) p. 4)

- Let S be the set of Proper Ideals in some commutative ring R $\neq 0$. As the 0 Ideal is an Ideal of any ring, and R $\neq 0$, then it is clear that 0 is a Proper Ideal of R. Therefore, S $\neq \emptyset$. S can be partially ordered by inclusion. Let us define I_p to be a totally ordered set of Proper Ideals in R, under the ordering $\subset$. Namely, a totally ordered subset of S. Let I = $\bigcup I_p$. The first thing to be done is to show that I is an Ideal:

- Do we have that $0_R \in$ I? Well, for every Ideal $I_k \in I_p$, we have that $0_R \in I_k$. Therefore, $0_R \in$ I

- Do we have that For all $x \in$ I, $(-x) \in$ I? Well for any $x \in$ I, there exists $I_k \in I_p$ such that $(-x) \in I_k$, so (-x) $\in$ I

- Do we have that For all $x, y \in$ I, $(x + y) \in$ I? Well, Take $x, y \in$ I. Then there exists I_k, $I_j \in I_p$ such that $x \in I_k$, $y \in I_j$. Since I_p is totally ordered by inclusion, either $I_j \subset I_k$ or $I_k \subset I_j$. Assume, without loss of generality, that $I_k \subset I_j$. Then $x, y \in I_j$ so $(x + y) \in I_j$ so $(x + y) \in$ I.

- Is it the case that for all $r \in$ R, $x \in$ I, we have $rx \in$ I and $xr \in$ I? Well, take any $r \in$ R, $x \in$ I. Then there exists $I_k \in I_p$ such that xr $\in I_k$ and rx $\in I_k$, so $xr \in$ I and $rx \in$ I.

Therefore all the conditions are satisfied, so we can confirm that I is an Ideal in R.

The next thing to show is that I is a Proper Ideal of R. Namely I $\neq$ R. Assume I is not a Proper Ideal of R. Namely, I = R. But then, as R is a ring, we have that $1 \in$ R, and so $1 \in$ I. But then, there exists $I_k \in I_p$ such that $1 \in I_k$. But then, as I_k is an Ideal, we have that for every r $\in$ R, 1r $\in I_k$ and r1 $\in I_k$. So, I_k is not a Proper Ideal of R, a contradiction. Therefore $1 \notin$ I and so I $\neq$ R, so I is a Proper Ideal of R, so I $\in$ S. As I is the union of a set of Proper Ideals in P, and we've shown that it is itself a Proper Ideal in P, it is clear that I is an upper bound of Proper Ideals in P. We have therefore shown that given any totally ordered subset of S (every $I_k \in I_p$) has an upper bound in S. By Zorn's Lemma, we can deduce that S contains a maximal element. As S consists of all the Proper Ideals in R, then of course the maximal element is a Proper Ideal, and whence a maximal Ideal. □

Section 5 - The Banach-Tarski Paradox

Although it is apparent that the axiom of Choice can be put to good use, proving significant results within mathematics, there are questions as to the extent to which is give rises to counter-intuitive implications, one of these being the Banach-Tarski Paradox. The Banchi-Tarski paradox is a theorem that was proven in the 1924, by Tefan Banach and Alfred Tarski, that relies on the Axiom of Choice.

Theorem 5.1 - *The Banach-Tarski Paradox* - If B is a closed ball in $\mathbb{R}^3$, then there exists a partition of B into pairwise disjoint pieces such that these pieces can be rearranged to yield two identical copies of the original ball.
This is clearly untrue in the real world, whereas in the mathematical world, it can be shown using the axiom of choice:

Proof: (Wapner, L.M. (2005) pp. 143-156)

Take the surface of the ball. This is precisely the set of uncountably infinite points $S = \{x, y, z : x^2 + y^2 + z^2 \leq 1\}$. If we can show that we can partition S into sets and then rearrange them so we have two sets S_1, S_2 such that S_1, S_2, and S all have the same size, then we have proved the theorem.

Step 1 - Defining Rotations

Let us begin by defining some rotations. Define θ as a clockwise rotation of 120° on the z-axis, and define γ as a clockwise rotation of 180° on the line z=x. We may define a countably infinite number of sequence of rotations using θ and γ.
Let us right any sequence of these rotations from right to left. For example:

- $\theta\gamma\gamma$ would represent a rotation of γ then γ then θ
- $\gamma\theta$ would represent a rotation of θ then γ.

Note - θ^3 and γ^2 represent a 360° turn, so any sequence consisting of either γ^2 or θ^3 can be reduced to nothing.

Examples:

- $\theta^3\gamma^2\theta^2\gamma\theta = \theta^2\gamma\theta$
- $\theta^9\gamma^7 = (\theta^3)^3(\gamma^2)^3\gamma = \gamma$

We define e to be the identity rotation, a rotation of 0°.

There is a countably infinite number of rotations defined in the above form. Let G be the countably infinite collection of rotations of B. Let the length of g∈G be precisely the number of times a symbol is used to define a rotation, e.g. the length of $\theta\gamma^3\theta$ is 3. We also note that every rotation in G has a unique physical representation. 2 different members of G will represent 2 different physical rotations. This is known as the "Uniqueness theorem".

Step 2 - Partitioning G

We aim to partition G into 3 subsets G_1, G_2 and G_3. We are to implement an algorithm to sort each g ∈ G into exactly one of these 3 subsets. We first look at the rotations of length 0 or 1. Then, of length 2, then 3, etc. We begin by assigning e to G_1, γ and θ to G_2, and θ^2 to G_3.
The sorting process is precisely as follows:

	If g∈G1	**If g∈G2**	**If g∈G3**
If the leftmost character of g is θ or θ^2	Let γg∈G2	Let γg∈G1	Let γg∈G1
If the leftmost character of g is γ	Let θg∈G2 and Let θ^2g∈G3	Let θg∈G3 and Let θ^2g∈G1	Let θg∈G3 and Let θ^2g∈G2

The process continues forever, and due to the uniqueness theorem, the intersection of $G_1 \cap G_2 = G_2 \cap G_3 = G_3 \cap G_1 = \emptyset$. This process works in such a way that G_1, G_2, G_3 are all neatly related. Take any g $\in G_1$, write a θ to the left of it, and you are left with a rotation in G_2. We may write $\theta G_1 = G_2$. Similarly, adding a θ^2 to the left of any rotation in G_1, and you are left with a rotation in G_3. Again, taking any rotation in G_1, add a γ to the left of it, and you will naturally be left with a rotation that is a member of either G_2 or G_3.

We may write this as follows:

- $G_2 = \theta G_1$
- $G_3 = \theta^2 G_1$
- $G_2 \cup G_3 = \gamma G_1$

Step 3 - Poles and Orbits

Definition 5.2 - A *Pole* of a rotation $g \in$ G is a point $(x, y, z) \in$ S such that the point remains unchanged after a rotation. Any rotation has 2 poles, namely, the 2 points on the surface of the sphere that that rotation intersects. As there are a countably infinite number of rotations in G, and each rotation has 2 poles, we can be sure that there are a countably infinite number of poles.

Example - θ has poles (0,0,1) and (0,0,-1) as a rotation of 120° would not changed the location of these points.

Let $P = \{g \in$ G : G is a pole$\}$ (A set with countably infinitely many members). Recall that $S = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ (A set with uncountably infinitely many members), so S/P $= \{(x, y, z) : x^2 + y^2 + z^2 \leq 1$ and (x, y, z) is not a pole$\}$. Note that there are a countably infinite number of points in P but an uncountably infinite number of points in S, so there are an uncountably infinite number of points in S/P.

Take any point $a \in S/P$. There exist other points in S/P such that a can reach those points via some rotation $g \in$ G. In fact, *every* $a \in S/P$ has a countably infinite number of these, given that there are a countably infinite number of unique rotations in G. We can say that a and b are in the same *Orbit* if there exists $g \in$ G such that a and b are reachable from one another by g. There are uncountably many orbits within S/P.

Step 4 - The Axiom of Choice

Now, we are at a stage where we require the axiom of choice. There are an uncountably infinite amount of points s $\in$ S that are not poles (points in S/P). Suppose we define a new set C. Let C consist of exactly one member of every orbit in S/P. This is where we need the axiom of choice, as the axiom of choice is the precise tool that allows us to create this set C.

Remarks:

- C is a set consisting of elements from an uncountably infinite collection of sets, which means that C is itself, countably infinite.
- We know that C and P have no points in common, as every element of C is an element of an orbit, and every element of an orbit is a member of S/P, (So not a member of P).
- Given that C has *exactly* one element from each orbit, no elements $a, b \in$ C can be rotated to reach on another, as otherwise they would be in the same orbit.
- Given that C has a representative from every orbit in S/P, we have that if *all* the elements in C were rotated via *all* the rotations $g \in$ G, all points in S/P would be reached.

Let us now apply every set of rotations $g \in G_1$ to every element $c_i \in$ C. Let us denote the collection of points reached here as K_1. We can also do the same for all g $\in G_2$ and g $\in G_3$, forming sets K_2 and K_3 similarly. Given that C consisted of an element from every orbit in S/P, we have that the process of rotating each element of C with every element of G_1, G_2 and G_3 has essentially pinpointed every element in every orbit and partitioned each of these elements into K_1, K_1, and K_3. Precisely, the whole of S/P has been partitioned here.

Step 5 - Hausdorff's paradox

We now have that the sphere S is partitioned into 4 disjoint sets; K_1, K_2, K_3, and P. Let us recall from the way we partitioned G in step 2, that $G_2 = \theta G_1$, $G_3 = \theta^2 G_1$, and $G_2 \cup G_3 = \gamma G_1$. We can therefore deduce that K_2=θK_1, K_3=$\theta^2 K_1$, and $K_2 \cup K_3 = \gamma K_1$. But what does this mean?
It means there is a one-to-one mapping from elements in K_1 to elements in K_2, from elements in K_1 to elements in $K_2 \cup K_3$, and from elements in K_2 to elements in K_3. We say they are *congruent* and can write this as follows:

$$K_1 \cong K_2 \cong K_3 \cong K_2 \cup K_3$$

(Hausdorff's Paradox)

This a paradox because of the following:

Given that K_1, K_2, K_3 are congruent and partition S/P, as P is infinitely small compared with S, we can deduce that K_1, K_2, K_3 each contain (roughly) 1/3 of the the sphere's surface. As $K_2 \cap K_3 = \emptyset$, this means $K_2 \cup K_3$ must contain 2/3 of the sphere's surface. However, as each K_i is congruent to $K_2 \cup K_3$, this means each K_i contains approximately 2/3 of the sphere's surface, contradictory to what was said above. This is Hausdorff's Paradox - How can K_i be both 1/3 and 2/3 of the sphere's surface?

Each K_i being congruent to $K_2 \cup K_3$ means thats each K_i can be partitioned into 2 sets K_{i1} and K_{i2}, one being congruent to K_2 and the other being congruent to K_3. But then now we have that S/P = $K_1 \cup K_2 \cup K_3 = K_{11} \cup K_{12} \cup K_{21} \cup K_{22} \cup K_{31} \cup K_{32}$. But each of these are congruent to K_2 or K_3. So we have 6 sets each congruent to a set roughly 1/3 the size of the sphere's surface. Let us separate the 6 sets into two sets of 3.

$$K_{11} \cup K_{12} \cup K_{21}$$

is a partition of the sphere's surface and

$$K_{22} \cup K_{31} \cup K_{32}$$

is also a partition of the sphere's surface.

So, $K_{11} \cup K_{12} \cup K_{21} \cup K_{22} \cup K_{31} \cup K_{32}$ is two copies of the S/P, resulting from just one copy of the S/P!

We have now almost shown that we can duplicate the surface of our ball, but not quite. We have created two copies of S/P but we need to consider the poles. We may add the set P to one of the S/P copies, but what about the other copy?

Remark - The surface of a sphere missing a countable number of points can be decomposed into a number of pieces, and reassembled into the entire sphere. Specifically, S/P can be decomposed and reassembled into our sphere S.
This can be done through a process called *Shifting from Infinity*. It is the idea that there is a countably infinite number of holes where the poles of rotation would have been, and we can rotate the sphere is such a way that each missing pole point shifts onto another point, so the missing points can essentially be "plugged in" This process can be read into further in (Wapner, L.M. (2005) pp. 100-103)

Step 6 - From the Sphere to the ball

We have that S can be partitioned and reassembled into 2 duplicate spheres, S_a and S_b. Although we have shown that we can duplicate the sphere S, we need to show that we can duplicate the Ball B. For every subset of $P, K_1, K_2, K_3 \subseteq S$ (and S), extend each member of the set towards the centre (0,0,0) (not including the centre) For each set, define our extended sets as $P^e, K_1^e, K_2^e, K_3^e, S^e$ for P, K_1, K_2, K_3, S respectively. Given that $K_1 \cong K_1 \cong K_3 \cong K_2 \cup K_3$, we can inherently infer that $K_1^e \cong K_2^e \cong K_3^e \cong K_2^e \cup K_3^e$, and that S^e can be partitioned into 2 identical balls S_a^e and S_b^e. However, these balls do not have a centre point (0,0,0).

For S_a^e, we can implement the point (0,0,0) from the original ball, and for S_b^e, we can use the previously mentioned process of shifting from infinity for the missing point (0,0,0). We now have 2 balls S_a^e and S_b^e which are identical to our original ball B. $\square$

We have now seen how the Axiom of choice has many applications such as in Linear Algebra, yet also gives rise to questionable results, as above. It is for these reasons combined that the Axiom of choice is not only considered one of the most fundamental ideas in all of mathematics, but also one of the most controversial.

BIBLIOGRAPHY:

Barnum, K (2013). The Axiom of Choice and It's Implications. [ONLINE] Available at: http://math.uchicago.edu/ may/REU2014/REUPapers/Barnum.pdf. [Last Accessed 01/03/2020]. p. 3; 5

Cohn, P.M., (2000). An Introduction to Ring Theory. 1st ed. London: Springer-Verlag London Limited.

Conrad, K (No date). Zorn's Lemma and Some Applications. [ONLINE] Available at: https://kconrad.math.uconn.edu/blurbs/zorn1.pdf. [Last Accessed 20/01/2020]. p. 4; 16

Dudley, R.M. (2018). Real Analysis and Probability. 2nd ed. New York: Taylor and Francis group. p. 16

Folland, G.B., 1984, Real Analysis - Modern Techniques and their applications, Canada, John Wiley and Sons, Inc. pp. 5-6

Heijenoort, V. (1967), From Frege to Godel : A Source Book in Mathematical Logic, 1879-1931. Cambridge, Massachusetts, Harvard University Press, pp. 139; 140-141.

Khatchatourian, I (2018). 11. The Axiom of Choice. [ONLINE] Available at: http://www.math.toronto.edu/ivan/mat327/docs/notes/11-choice.pdf . [Last Accessed 25/02/2020]. p. 7; 8

Mantova, V.L. (2019), Advanced Models and Sets, Lecture Notes, MATH5120M, University of Leeds, delivered December 2019, p. 90

Russell, B. (1919), Introduction to Mathematical Philosophy, London, George Allen and Unwin. pp. 156-157.

Wapner, L.M. (2005). The Pea and the Sun - A mathematical Paradox. 1st ed. Wellesley, Massachusetts: AK Peters, Ltd pp. 100-103; 143-156;

Winfried, J. and Weese, M., (1996). Discovering Modern Set Theory. I - The basics. 1st ed. USA, The American Mathematical Society p. 134