THE AXIOM OF CHOICE AND ITS IMPLICATIONS

KEVIN BARNUM

Abstract. In this paper we will look at the Axiom of Choice and some of the various implications it has. These implications include a number of equivalent statements, and also some less accepted ideas. The proofs discussed will give us an idea of why the Axiom of Choice is so powerful, but also so controversial.

Contents

1. Introduction

The Axiom of Choice states that for any family of nonempty disjoint sets, there exists a set that consists of exactly one element from each element of the family. It seems strange at first that such an innocuous sounding idea can be so powerful and controversial, but it certainly is both. To understand why, we will start by looking at some statements that are equivalent to the axiom of choice. Many of these equivalences are very useful, and we devote much time to one, namely, that every vector space has a basis. We go on from there to see a few more applications of the Axiom of Choice and its equivalents, and finish by looking at some of the reasons why the Axiom of Choice is so controversial.

2. The Axiom of Choice and Its Equivalents

2.1. The Axiom of Choice and its Well-known Equivalents.

We will start by looking at some of the most famous equivalents of the Axiom of Choice. Many of these equivalences are used extensively throughout mathematics. Here is a statement of the axiom itself.

Date: DEADLINES: Draft AUGUST 18 and Final version AUGUST 29, 2013.

Axiom of Choice 2.1 (The Axiom of Choice)**.** *For any family A of nonempty, disjoint sets, there exists a set that consists of exactly one element from each element of A.*

Axiom of Choice 2.2 (The Choice-Function Principle)**.** *For any family A of nonempty sets, there exists a function $f : A \to \bigsqcup_{a \in A} a$ such that for every $a \in A, f(a) \in a$.*

For the next two equivalences, we need to define some types of ordering.

Definition 2.3. On a set S, we say that $<$ is a *partial ordering*, if for $s_1, s_2, s_3 \in S$, (1) $s_1 \leq s_2$, $s_2 \leq s_1 \Rightarrow s_1 = s_2$, (2) $s_1 \leq s_2 \leq s_3 \Rightarrow s_1 \leq s_3$, and (3) $s_1 \leq s_1$.

Definition 2.4. We say that $(S, <)$ is a *total ordering* or *linear ordering*, if $(S, <)$ is a partial ordering, and for every $s_1, s_2 \in S$, either $s_1 \leq s_2$, or $s_2 \leq s_1$.

Definition 2.5. We say that a total ordering $(S, <)$ is a *well-ordering* if for every $T \subseteq S$, there exists an $s_0 \in T$ such that $s_0 \leq s \, \forall s \in T$.

Axiom of Choice 2.6 (Well-Ordering Principle)**.** *For any set S, there is a relation $<$ such that $(S, <)$ is a well-ordering.*

The next equivalent statement, Zorn's Lemma, is very important, as we will later see, in proving several results in number theory.

Axiom of Choice 2.7 (Zorn's Lemma)**.** *If $(A, <)$ is a partial ordering such that $A \neq \emptyset$ and every subset of A simply ordered by $<$ has an upper bound, then A has a maximal element under $<$.*

Definition 2.8. A family of sets A has *finite character* if for every set a, a belongs to the family A if and only if every finite subset of a belongs to A.

We include this next equivalence because it will be necessary for when we prove Lemma 3.1.

Axiom of Choice 2.9 (Tukey's Lemma)**.** *If A is a nonempty family of sets with finite character, then A has a maximal element with respect to $\subseteq$.*

Before we prove the equivalence, we will have to introduce one more concept: transfinite recursion. Though a full exploration of the topic is outside the scope of this paper, we include it as a lemma as it is necessary for the theorem that follows. For more a more detailed examination of transfinite recursion, see [7].

Lemma 2.10 (Transfinite Recursion)**.** *Let V be the class of all sets, and let Ord be the class of all ordinals. Given a function $F : V \to V$, there exists a function $G : Ord \to V$ such that if we define $a_\alpha := G(\alpha)$, then for each α,*
$a_\alpha = F(\langle a_\xi : \xi < \alpha \rangle)$.

Theorem 2.11. *The following statements are equivalent (i) The Axiom of Choice, (ii) The Choice-Function Principle, (iii) The Well-Ordering Principle, (iv) Zorn's Lemma, (v) Tukey's Lemma.*

Proof. (i) $\Rightarrow$ (ii): Consider a family of nonempty sets A. For each set $a \in A$, let $X_a := \{(a, x) | x \in a\}$ and let $\alpha := \{X_a | a \in A\}$. Because each element of A is nonempty, each $X_a \in \alpha$ must be nonempty as well.

Now consider $a, b \in A$, with $Y, Z \in \alpha$, and $Y \neq Z$. Let $Y = \{(a, y) : y \in a\}$ and $Z = \{(b, z) : z \in b\}$. Now $Y \neq Z \Rightarrow a \neq b \Rightarrow Y \cap Z = \emptyset$. In other words, for every distinct $Y, Z \in \alpha$, Y and Z are disjoint. Thus, we can use the Axiom of Choice to choose one pair $(a, y) \in Y$ for every $Y \in \alpha$, and we can define $\beta := \{(a, y) | (a, y) \text{ chosen as described}\}$ and β will contain exactly one element in common with each element of α. Now, we can define $f(a)$ to be the function that when given the set a, returns the element of a that a is paired with in β. This f is a choice function.

(ii) ⇒ (iii): Consider a set S and let F be a choice function for all nonempty subsets of S. We must find an ordinal number α such that the sequence $a_0, \ldots, a_\xi, \ldots (\xi < \alpha)$ is a one-to-one enumeration of S. We create a sequence of a_ξ by transfinite recursion starting with $a_0 = F(S)$ and for all subsequent a_ξ, we have $a_\xi = F(S \backslash \{a_\eta : \eta < \xi\})$. This new set $\{a_0, \ldots, a_\xi, \ldots\}$ will eventually contain every element of S and the elements are well-ordered by the subscript.

(iii) ⇒ (iv): Consider $(S, <)$, a nonempty, partially ordered set and assume that every chain in S has an upper bound. By the Well-Ordering Principle, S can be well-ordered, so for some ordinal α, there exists an enumeration of $S = \{p_0, \ldots, p_\xi, \ldots\}$.

We now create a new set C by transfinite recursion, for which $c_0 = p_0$ and $\forall \xi > 0$, $c_\xi = p_\gamma$, where p_0 is the least element of S, and p_γ is an upper bound of $C_\xi = \{c_\eta : \eta < \xi\}, p_\gamma \notin C_\xi$. $\{c_\eta : \eta < \xi\}$ is always a chain and p_γ exists unless $c_{\gamma-1}$ is a maximal element of p. Thus, we will eventually obtain a maximal element of S.

(iv) ⇒ (v): Consider A a nonempty family of sets with finite character. A is partially ordered by $\subseteq$. For some chain C in A, let $B = \cup\{X : X \in C\}$, then every finite subset of B is in A, which means that B is in A and B is an upper bound of C. Therefore, by Zorn's lemma, there is a maximal element of A.

(v) ⇒ (i): Consider A, a family of disjoint nonempty sets. We define $B := \bigsqcup_{a \in A} a$ and $G := \{b \subseteq B : |b \cap a| \leq 1 \, \forall a \in A\}$.

We now prove G has finite character by contradiction. First, assume that G does not have finite character. There are two cases.

Case (1): There exists a $b \in G$ such that not every finite subset of b is in G. Then there exists b_0, a finite subset of b such that $b_0 \notin G$. But clearly since $b_0 \subset b$, $|b_0 \cap a| \leq 1$, which means $b_0 \in G$, and thus is a contradiction. Therefore, for all $b \in G$, every finite subset of b is in G.

Case (2): there exists a $d \notin G$ with every finite subset of d in G. But $d \notin G$ implies that there exists some $a \in A$ for which $|d \cap a| \geq 2$. But then there is a finite subset d_0 of d that consists of only two elements for which $|d_0 \cap a| = 2$ for some $a \in A$, and thus every finite subset of d cannot be in G. Thus, G must have finite character.

Because G has finite character, we can now apply Tukey's lemma to say there exists a maximal set $\tilde{b} \in G$. We now want to show that $\tilde{b}$ is the set we want in the Axiom of Choice. Suppose, for contradiction, that $\tilde{b} \cap a = \emptyset$ for some $a \in F$. Then we can pick any $u \in X$ and then $\tilde{\tilde{b}} = \tilde{b} \cup \{u\}$ is a proper extension of $\tilde{b}$ with $\tilde{\tilde{b}} \in G$. But this contradicts the maximality of C. Thus, $|C \cap X| = 1 \forall X \in A$. So $\tilde{b}$ is the desired set that satisfies the Axiom of Choice. □

2.2. Some Other Less Well-known Equivalents of the Axiom of Choice.

The axiom of choice has many, many equivalences, many more than are within the scope of this paper. The equivalences proved above are among the most important. These next equivalences, though also important, are not used quite as regularly as those above, however, we include these for their use later in proving Theorem 3.7.

Axiom of Choice 2.12 (The Axiom of Multiple Choice)**.** *For every family A of sets, there exists a function f on A such that for every $a \in A$, $f(a)$ is a finite nonempty subset of a.*

Definition 2.13. An *antichain* is a chain in a partially ordered set that consists entirely of mutually incomparable elements.

Axiom of Choice 2.14 (Antichain Principle)**.** *Every partially ordered set has a maximal antichain.*

Axiom of Choice 2.15. *Every linearly ordered set can be well ordered.*

Definition 2.16. The *power set* of a set S is the set of all subsets of S (including S itself and the empty set).

Axiom of Choice 2.17. *The power set of every well-ordered set can be well-ordered.*

The equivalence we are about to prove holds in Zermelo-Frankel set theory, a set theory that consists of the axioms that along with the Axiom of Choice, most mathematicians accept as true today. It is possible to create other set theories, however, for which this equivalence chain does not hold. In particular, if we work in a set theory that includes objects other than sets, called atoms, the equivalence fails.

Theorem 2.18. *The following are equivalent in Zermelo-Frankel set theory (ZF): (i) The Axiom of Choice, (ii) The Axiom of Multiple Choice, (iii) The Antichain Principle, (iv) Axiom of Choice 2.15, (v) Axiom of Choice 2.17.*

Proof. (i) ⇒ (ii): The proof is trivial because we have already shown that The Axiom of Choice is equivalent to the Choice-Function Principle, which is clearly stronger than the Axiom of Multiple Choice.

(ii) ⇒ (iii): We consider $(P, <)$, a partially ordered set. We know from the Axiom of Multiple Choice that there exists a function f such that for all $X \neq \emptyset, X \subseteq P$, $f(X)$ is a nonempty finite subset of X.

We next define $g(X)$ for nonempty $X \subseteq P$ to be the set of all elements of $f(X)$ that are minimal by $<$. Then $g(X)$ is a nonempty finite antichain.

We now construct a maximal antichain by transfinite recursion. Let $A_0 = g(P)$ and for α in the set of ordinal numbers, $A_\alpha = g(X)$, where X is the set of $x \in P$ which are incomparable with every element of $\cup\{A_\beta : \beta < \alpha\}$.

We finally show that A, the union of every set A_α is a maximal antichain in P. Assume there exists $a \in P$ such that a is incomparable with every element of A. Then there also exists some element of P that is comparable to a, minimal by $<$, and not in A. But by the construction of A, a must be in some A_α, which is a contradiction. Thus, A is a maximal antichain in P.

(iii) ⇒ (iv): Consider the linearly ordered set $(Q, <)$. As we have already shown in

Theorem 2.11, Q can be well-ordered if we can find a choice function on the power set of Q.

Let P be the set of pairs (X, x) for all $X \subset Q$, $x \in X$. Define the partial ordering $\prec$ on P by $(X, x) \prec (Y, y)$ if and only if $X = Y$ and $x < y$. By the Antichain Principle, $(P, \prec)$ has a maximal antichain A and A is clearly a choice function that pairs a set with an element of that set.

(iv) $\Rightarrow$ (v): Consider X, a well-ordered set. It is not difficult to see that the power set of X can be linearly ordered using a lexicographic ordering. Therefore, by Axiom of Choice 2.15, the power set of X can be well-ordered.

(v) $\Rightarrow$ (i): We assume the power set of every well-ordered set is orderable by Axiom of Choice 2.17. It will suffice to show that for all limit ordinals α, the family V_α of all sets of rank less than α can be well-ordered.

Let α be some fixed limit ordinal, and let κ be the least ordinal such that there exists a one to one mapping from κ to V_α. By Axiom of Choice 2.17, the power set of κ can be well-ordered.

We define W_β on $V_\beta = \cup_{\gamma<\beta} V_\gamma$ to be a well-ordering of V_β such that $\beta_1 < \beta_2 \Rightarrow V_{\beta_1} < V_{\beta_2}$ for all $\beta < \alpha$ with $W_0 = 0$, and β a limit ordinal. Now, if $\beta = \gamma + 1$, V_β is the power set of V_γ. V_γ is well-ordered by W_γ and thus has a one to one correspondence with ordinal $\lambda < \kappa$. We therefore have a well-ordering W_β of the power set of V_γ. □

3. Applications of the Axiom of Choice

In this section we turn to some of the many ways that the Axiom of Choice is used in mathematics.

3.1. Equivalence Between The Axiom of Choice and the Claim that Every Vector Space has a Basis.

In this section we will prove another equivalence of the Axiom of Choice: the claim that every vector space has a basis. This next lemma, the first part of the equivalence, has been known for a long time.

Lemma 3.1. *The Axiom of Choice implies every vector space has a basis.*

Proof. Consider V, a vector space. A is the family of all linearly independent sets of vectors in V. A has finite character, therefore, by Tukey's lemma, there exists a maximal linearly independent set B. Clearly, then, B is a basis. □

Unlike the previous theorem, the next part of the equivalence wasn't proven until fairly recently, 1984 [4].

Lemma 3.2. *In ZF, every vector space has a basis implies the axiom of multiple choice.*

Proof. For some cardinal I, consider the family of nonempty sets $\{X_i : i \in I\}$ and assume that the sets X_i are pairwise disjoint. Define $X := \sqcup_{i \in I} X_i$ and adjoin all elements of X as indeterminates to some field k. We get the field $k(X)$.

We define the i-degree of a monomial to be the sum of exponents of elements of X_i in the monomial. We also say that the rational function $f \in k(X)$ is i-homogenous of degree d if it is the quotient of two polynomials such that all monomials in the denominator have i-degree n for some n, and those in the numerator have i-degree $n + d$.

The rational functions that are i-homogenous of degree $0\ \forall i \in I$ constitute a subfield K of $k(X)$ and $k(X)$ is a vector space over K. We define V to be the subspace spanned by X and B to be the basis of the K-vector space V.

Now, $\forall i \in I, x \in X_i$, we can express x as a finite K-linear combination of elements of B:

$$x = \sum_{b \in B} \alpha_b(x) \cdot b \tag{3.3}$$

where $\alpha_b(x) = 0$ for almost all $b \in B$.

Similarly, for $y \in X_i$, we have

$$y = \sum_{b \in B} \alpha_b(y) \cdot b \tag{3.4}$$

or, if we multiply equation (3.4) by y/x, which is in K because x and y are elements of the same X_i and clearly have the same degree, we get

$$\begin{aligned} y &= \sum_{b \in B} \left(\frac{y}{x}\right) \alpha_b(x) \cdot b \\ \Leftrightarrow \frac{\alpha_b(y)}{y} &= \frac{\alpha_b(x)}{x}. \end{aligned} \tag{3.5}$$

And since B is a basis, by equations (3.4) and (3.5), we get

$$\alpha_b(y) = \left(\frac{y}{x}\right) \cdot \alpha_b(x). \tag{3.6}$$

But now we see that $\alpha_b(x)/x$ depend only on i, not on the specific $x \in X_i$, so $B_{bi} := \frac{\alpha_b(x)}{x}$ for some $x \in X_i$ is well-defined.

We can see that B_{bi} is i-homogenous of degree -1. Therefore, when B_{bi} is written as a quotient of polynomials in reduced form, some variables from X_i are in the denominator. We now let F_i be the set of those elements of x_i that occur in the denominator of B_{bi} for some $b \in B$. F_i is a nonempty finite subset of X_i as required. □

It is now clear that the claim every vector space has a basis and the axiom of choice are equivalent, because the equivalence of the Axiom of Multiple Choice and the Axiom of Choice has already been proven. We thus get the theorem we desire.

Theorem 3.7. *Every vector space has a basis is equivalent to the Axiom of Choice.*

3.2. **Some More Applications of the Axiom of Choice.**

We now look at some more evidence of the power of the Axiom of Choice and its use throughout mathematics. As we will see, there are many types of problems where the Axiom of Choice and its equivalents can help us.

We start with a puzzle of sorts, whose solution, if you can find it, requires the Axiom of Choice.

Question 3.8. *An evil wizard has threatened a village where an infinite number of gnomes reside. The wizard will cast a spell that will cause a hat to appear on the head of every gnome. Each hat will either be red or blue, but each gnome will be unable to see that hat on his or her head. The wizard will leave the gnomes alone only if only a finite number of gnomes guess the color of the hat on their heads incorrectly. The gnomes can strategize before the wizard puts the hats on* *their heads, but they cannot talk or communicate with each other once the hats are on their heads. The gnomes have very good eyesight and can see the hat of every other gnome. The wizard can listen to the gnomes strategize and choose the most evil possible placement of hats. What should the gnomes do?*

The next problem is the one that first interested me in the Axiom of Choice as it is not possible to find a solution without the axiom. We use the definitions of weakly convex and convex from Emil Artin [1, p.1,4-5], but others who study the concept may use the term midpoint convex instead of weakly convex.

Definition 3.9. A function f is *weakly convex* if it satisfies the following inequality on the interval for which it is defined

$$f\left(\frac{x_1+x_2}{2}\right) \le \frac{1}{2}(f(x_1)+f(x_2)) \tag{3.10}$$

Definition 3.11. Define the difference quotient:

$$\varphi(x_1,x_2) = \frac{f(x_1)-f(x_2)}{x_1-x_2},$$

where f is a real-valued function defined on an open interval. f is *convex* if for every x_3 of the interval, and for any pair of numbers $x_1 > x_2$ distinct from x_3, the following inequality holds:

$$\varphi(x_2,x_3) \le \varphi(x_1,x_3)$$

Lemma 3.12. *Every real-valued, convex function defined on an open interval is continuous.*

Proof. Consider a function f that is convex in (a,b). For $a < s < x < u < b$, and $t \to x$, for t close enough to x, we have

$$\varphi(x,s) \le \varphi(x,t) \le \varphi(x,u) \Leftrightarrow \frac{f(x)-f(s)}{x-s} \le \frac{f(x)-f(t)}{x-t} \le \frac{f(x)-f(u)}{x-u}$$

And now we take $f(x)-f(t) = \left(\frac{f(x)-f(t)}{x-t}\right)\cdot(x-t)$ And combining these two equations we have:

$$\begin{aligned}
&\left(\frac{f(x)-f(s)}{x-s}\right)\cdot(x-t) \le \left(\frac{f(x)-f(t)}{x-t}\right)\cdot(x-t) \le \left(\frac{f(x)-f(u)}{x-u}\right)\cdot(x-t)\\
&\Rightarrow 0 \le f(x)-f(t) \le 0\\
&\Rightarrow f(x)-f(t) = 0
\end{aligned}$$

So then $f(x)$ must be continuous. □

Theorem 3.13. *There exists a function that is weakly convex, but not convex.*

Proof. From Lemma 3.1, we have a basis B for the vector space $\mathbb{R}$ over $\mathbb{Q}$. One of the elements of that basis, call it r_1 must be a rational number, and WLOG, let $r_1 = 1$. We now take a function f that switches two basis vectors r_{i_1}, r_{i_2} , for some fixed $i_1 < i_2$, other than r_1, so that we have $f\left(\frac{p}{q}\right) = \frac{p}{q}$ for $p, q \in \mathbb{Z}, q \neq 0$. Thus, we have for $x_1, x_2 \in \mathbb{R}$, $r_i \in B$, and $\alpha_i, \beta_i \in \mathbb{Q}$

$$\begin{aligned}
f(x_1) &= \sum_{\substack{i=1\\ i\neq i_1,i_2}} (\alpha_i r_i) + \alpha_{i_1} r_{i_2} + \alpha_{i_2} r_{i_1}\\
f(x_2) &= \sum_{\substack{i=1\\ i\neq i_1,i_2}} (\beta_i r_i) + \beta_{i_1} r_{i_2} + \beta_{i_2} r_{i_1}\\
f\left(\frac{x_1+x_2}{2}\right) &= \sum_{\substack{i=1\\ i\neq i_1,i_2}} \left(\left(\frac{\alpha_i+\beta_i}{2}\right) r_i\right) + \left(\left(\frac{\alpha_{i_1}+\beta_{i_1}}{2}\right) r_{i_2}\right) + \left(\left(\frac{\alpha_{i_2}+\beta_{i_2}}{2}\right) r_{i_1}\right)\\
&= \frac{1}{2}\left(\sum_{\substack{i=1\\ i\neq i_1,i_2}} (\alpha_i r_i) + (\alpha_{i_1} r_{i_2}) + (\alpha_{i_2} r_{i_1}) + \sum_{\substack{i=1\\ i\neq i_1,i_2}} (\beta_i r_i) + (\beta_{i_1} r_{i_2}) + (\beta_{i_2} r_{i_1})\right)\\
&= \frac{1}{2}(f(x_1)+f(x_2))
\end{aligned}$$

so clearly equation (3.10) holds. This function is discontinuous at infinitely many points, and thus it cannot be convex because, as we proved in Lemma 3.12, convex implies continuous. ☐

The next theorem is very important to functional analysis.

Definition 3.14. For a function p in some vector space V over the real numbers, we say that p is *sublinear* if $p(x+y) \leq p(x) + p(y)$ for $x, y \in V$, and $p(rx) = rp(x)$ for $x \in V$, $r \in \mathbb{R} \geq 0$.

Theorem 3.15 (Hahn-Banach). *For p, a sublinear function on the space V, we take f, a linear function on a subspace K of V, such that $f(x) \leq p(x)$ on K. There exists a linear function F on V, which extends f and $F(x) \leq p(x)$ holds for all $x \in V$.*

Proof. Consider p and f as defined above. We let S be the set of all pairs $< X, g >$, where X is a subspace of V that contains K, and g is a linear function on X that extends f, and for which $g(x) \leq p(x)$ holds for all $x \in X$. $S \neq \emptyset$ since clearly $< K, f > \in S$. S is partially ordered by extension if we say that
$< X_1, g_1 > \prec < X_2, g_2 >$ if X_1 is a proper subset of X_2. Every subset of S simply ordered by $\prec$ has an upper bound because each $X \leq V$. We can now use Zorn's Lemma, to yield a maximal member of S, $< M, F >$.

Now suppose $M \neq V$, then there exists $u \in V \backslash M$. Let $U := M \oplus < u >$. $\forall x \in U$, there is a unique representation of the form $x = z + \alpha u$ where $x \in M$, $\alpha \in \mathbb{R}$. For some constant γ, we can define G on u by $G(z + \alpha u) := F(z) + \alpha\gamma$. G is a linear function that properly extends F. We now want to show:

$$G(x) \leq p(x)\ \forall x \in U \tag{3.16}$$

For $x, y \in U$, we have

$$\begin{aligned}
&F(y) - F(x) = F(y-x) \leq p(y-x) \leq p(y+u) + p(-u-x)\\
&\quad \Rightarrow -p(-u-x) - F(x) \leq p(y+u) - F(y)
\end{aligned}$$

And since the left hand side of the resulting equation is independent of y and bounded above, while the right hand side is independent of x and bounded below, there exists a constant c such that we get the following equations:

$$c \leq p(y+u) - F(y) \tag{3.17}$$

$$-p(-u-x) - F(x) \leq c \tag{3.18}$$

for all $x, y \in M$.

Now for some $x = z + \alpha u \in U$, we want to show:

$$G(x) = F(z) + \alpha\gamma \leq p(z + \alpha u) = p(x)$$

This clearly holds for $\alpha = 0$. Now for $\alpha > 0$, we replace y in equation (3.17) by $\alpha^{-1}z$ to give us:

$$\begin{aligned} & c \leq p(\alpha^{-1}z + u) - F(\alpha^{-1}z) \\ \Leftrightarrow & F(\alpha^{-1}z) + c \leq p(\alpha^{-1}z + u) \\ \Leftrightarrow & \alpha^{-1}F(z) + c \leq \alpha^{-1}p(z + u\alpha) \\ \Leftrightarrow & F(z) + \alpha c \leq p(z + \alpha u) \end{aligned}$$

and we can similarly replace x in equation (3.18) by $\alpha^{-1}z$ to prove it for $\alpha < 0$. So $< U, G > \in S$ and clearly $< M, F > \prec < U, G >$. This contradicts the maximality of M. Therefore, $M = V$, and F is the function we want. □

We finish this section with two theorems that are fundamental to number theory and show the importance of Zorn's Lemma.

Theorem 3.19. *Every nonzero commutative ring has a maximal ideal.*

Proof. Consider S, the set of proper ideals in the commutative ring $R \neq 0$. Because the zero ideal is a proper ideal of R, $S \neq \emptyset$. S can be partially ordered by inclusion. We now define $\{I_\alpha\}$ to be a totally ordered set of proper ideals in R, and $I := \cup I_\alpha$.

We now show that I is an ideal. For $x, y \in I$, $x \in I_1$ and $y \in I_2$, for some I_1, I_2 in $\{I_\alpha\}$. WLOG, we can say that $I_1 \subset I_2$, so that because I_2 is an ideal, $x + y \in I_2 \subset I$. Similarly, for all $r \in R$, $rx \in I_\alpha \subset I$. Therefore, I satisfies the properties of an ideal, and thus is an ideal in R.

We next show that I is a proper ideal. Assume I is not a proper ideal. Then 1 must be an element of I. Therefore, 1 must be an element of I_0 for some $I_0 \in \{I_\alpha\}$. But then $I_0 = (1)$ is not a proper ideal, which contradicts that $\{I_\alpha\}$ is a set of proper ideals. Therefore, 1 cannot be an element of I, so I must be a proper ideal, $I \in S$.

We have shown that every totally ordered subset of S has an upper bound in S. Therefore, by Zorn's Lemma, S has a maximal element that is maximal for inclusion among all proper ideals. Thus, we have a maximal ideal in R. □

Theorem 3.20. *Every proper ideal in a nonzero commutative ring is contained in a maximal ideal.*

Proof. Consider a nonzero commutative ring R, and I a proper ideal in R. The quotient ring R/I is nonzero, so by Theorem 3.19, it contains a maximal ideal which we will call M. The inverse image of M under the quotient map $R \to R/I$ is a maximal ideal that contains I. □

4. Controversial Results

We have seen part of the reason the Axiom of Choice is so fundamental to math by looking at its equivalences and by looking at how it has been used to solve a number of important problems in various branches of mathematics. Now that the usefulness of the Axiom is clear, we now turn to some examples of why the axiom has been so widely debated throughout history. This theorem is one commonly cited in explanations of why the Axiom of Choice is controversial.

Theorem 4.1. *There exist non-measurable bounded sets of real numbers.*

Proof. Define $\mu(X)$ to be the Lebesgue measure of a set X of real numbers. For the real numbers in $[0,1]$, we define $x \sim y$ if $x - y \in \mathbb{Q}$. $\sim$ is an equivalence relation and we will denote the equivalence class for each $x \in [0,1]$ by $[x]$.

Using the axiom of choice, we can choose one element out of each equivalence class. Therefore, there exists a set M of real numbers, with $M \subset [0,1]$, such that for each $x \in [0,1]$, there exist unique $y \in M$, and $r \in \mathbb{Q}$ such that $x = y + r$. We define $M_r := \{m + r | m \in M\}$, so that $[0,1] = \sqcup\{M_r | r \in \mathbb{Q}\}$ is a partition of $[0,1]$ into countably many disjoint sets.

Now assume that M is measurable. There are two cases: Case (i): $\mu(M) = 0$ which implies $\mu([0,1]) = 0$, which is obviously impossible. Case (ii): $\mu(M) > 0$. Then

$$\begin{aligned}\mu([0,1]) &\geq \mu(\cup\{M_r : r \in \mathbb{Q} \text{ and } 0 \leq r \leq 1\}) \\ &= \sum_{\substack{0\leq r\leq 1 \\ r\in\mathbb{Q}}} \mu(M_r) \\ &= \sum_{\substack{0\leq r\leq 1 \\ r\in\mathbb{Q}}} \mu(M) \\ &= \infty\end{aligned}$$

Therefore, M is not measurable. □

The next theorem, the famous Banach-Tarski Paradox, is arguably the most well-known example of why the Axiom of Choice is not universally accepted. The Banach-Tarski Paradox essentially states that a sphere can be disassembled into two spheres that are identical to the first one. The proof of this theorem requires several lemmas and some math that is outside of the scope of this paper. For detailed proofs, see [6], or [9].

Theorem 4.2 (Banach-Tarski)**.** *We define $X \approx Y$ if there exists both a finite decomposition of X into disjoint sets such that $X = X_1 \cup \ldots \cup X_m$ and a finite decomposition of Y into the same number of disjoint sets such that $Y = Y_1 \cup \ldots \cup Y_m$, and X_i is congruent to Y_i for every $i = 1, \ldots, m$. A closed ball U can be decomposed into two disjoint sets, $U = X \cup Y$ such that $U \approx X$ and $U \approx Y$.*

We have seen a number of equivalences of the Axiom of Choice and a number of problems that it can solve. In addition to its usefulness, however, we have looked at some proofs that it can be used in that seem much less intuitive. For these reasons, the Axiom of Choice remains one of the most fundamental, but nonetheless contested, concepts in all of mathematics.

Acknowledgments. I would first like to thank Professor Peter May for giving me the opportunity to be part of such a great program this summer that helped me learn so much more about math. I would also like to thank all the Professors who taught classes this summer. Special thanks to my graduate student mentors, Yun Cheng and Preston Wake, for pushing me to keep learning more and for always giving me new problems to try. Lastly I would like to thank the other students in the program for making it so enjoyable and for helping me pursue mathematics further, in particular Paul Duncan and Zac Wang.

References

[1] Emil Artin. The Gamma Function. Trans. Michael Butler. Holt, Rinehart, and Winston. 1964.
[2] Keith Conrad. Zorn's Lemma and Some Applications http://www.math.uconn.edu/∼kconrad/blurbs/zorn1.pdf.
[3] John Bell. The Axiom of Choice. College Publications. 2009.
[4] Andreas Blass. "Existence of Bases Implies the Axiom of Choice." Comtemporary Mathematics, Volume 31. 1984.
[5] Kenneth Harris. http://kaharris.org/teaching/582/Lectures/lec28.pdf.
[6] Thomas J. Jech. The Axiom of Choice. North-Holland Publishing Company 31-33. 1973.
[7] Thomas J. Jech. Set Theory. Springer. 2006.
[8] Don Monk. The Axiom of Choice. http://euclid.colorado.edu/∼monkd/m6730/gradsets05.pdf.
[9] Terrence Tao. The Banach Tarski Paradox. https://www.math.ucla.edu/∼tao/preprints/Expository/banach-tarski.pdf.