2010

The Axiom of Choice

Cristian Allen
Virginia Commonwealth University

College of Humanities and Sciences
Virginia Commonwealth University

This is to certify that the thesis prepared by Cristian Gerardo Allen titled “The Axiom of Choice” has been approved by his or her committee as satisfactory completion of the thesis requirement for the degree of Master of Science.

Dr. Andrew Lewis, College of Humanities and Sciences

Dr. Kevin Beanland, College of Humanities and Sciences

Dr. David Ameen, College of Humanities and Sciences

Dr. John Berglund, Graduate Chair, Mathematics and Applied Mathematics

---

Dr. Robert Holsworth, Dean, College of Humanities and Sciences

---

Dr. F. Douglas Boudinot, Graduate Dean

Date

The Axiom of Choice

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

Cristian Gerardo Allen
Master of Science

Director: Dr. Andrew Lewis, Associate Professor, Department Chair
Department of Mathematics and Applied Mathematics

Virginia Commonwealth University
Richmond, Virginia
May 2010

Acknowledgment

I am grateful that Virginia Commonwealth University has given me such a strong education in mathematics. Many thanks are due to Dr. Andrew Lewis for accepting me as a student and being a role model. Many of the principles that I have learned from him go beyond mathematics and will help me have a fruitful career. I would like also to thank Dr. Kevin Beanland for supporting me throughout these two years of my masters education. As my analysis professor, his ideas of the subject and its positive aspects have rubbed off onto me as I begin to find myself more interested in it as time passes. I would like to thank Dr. David Ameen for his early support of my decision to pursue a career in teaching mathematics in the public schools, and then his support of my decision to come back to school and get a masters degree. His attitude towards teaching and life helped shape my own.

Contents

Abstract

THE AXIOM OF CHOICE

By Cristian Gerardo Allen, Master of Science.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2010.

Director: Dr. Andrew Lewis, Associate Professor, Department Chair, Department of Mathematics and Applied Mathematics.

We will discuss the 9th axiom of Zermelo-Fraenkel set theory with choice, which is often abbreviated ZFC, since it includes the axiom of choice (AC). AC is a controversial axiom that is mathematically equivalent to many well known theorems and has an interesting history in set theory. This thesis is a combination of discussion of the history of the axiom and the reasoning behind why the axiom is controversial. This entails several proofs of theorems that establish the fact that AC is equivalent to such theorems and notions as Tychonoff's Theorem, Zorn's Lemma, the Well-Ordering Theorem, and many more.

Introduction

In this thesis we will discuss the history of an axiom of mathematics that very controversial in the early 1900s. The axiom is referred to as the axiom of choice (AC) and finds its roots in the branch of mathematics called set theory. The principle purpose of set theory is to establish foundational axioms from which most mathematical statements can be proved. Of particular interest in this thesis is the set theory introduced by Ernst Zermelo in 1908 which was further refined by Abraham Fraenkel in 1922. We will refer to the axiomatic set theory introduced by these two mathematicians as Zermelo-Fraenkel set theory, or ZF for short. ZF consists of 8 axioms which give mathematics a solid and rigorous foundation.

It was eventually realized that some statements in mathematics could not be proved from the 8 axioms of ZF, even though some of the statements were believed to be true among many mathematicians. Mathematicians were using an axiom not included in ZF in many undisputed proofs of statements of mathematics. George Cantor frequently and without question used AC while he was developing the concept of cardinality. In addition to the use of AC to prove many acceptable statements, AC was being used to prove many unnatural statements as well. Cantor in 1897, but as late as 1905 by Zermelo, proved the well-ordering principle using AC. This created a problem for most mathematicians since they believed that the well-ordering principle would need to be assumed as an axiom, since there it would not be able to be proven using the 8 axioms. The immediate conclusion by many was that AC was not true and that the proofs that had used it to that point must be examined again. On the other hand many mathematicians began to study ZFC (the ZF set theory coupled with the axiom of choice).

As the controversy expanded, more statements were proved using AC, some of which were intuitive and some of which were not. These results included Zorn's Lemma, Tychonoff's Theorem for compact spaces, the well-ordering principle, and Tukey's Lemma. These turned out to be equivalent to AC. Other important results were shown to be provable from AC but not equivalent, such as the Hahn-Banach Theorem of functional analysis. In 1924, Stefan Banach and Alfred Tarski proved that a sphere of any radius could be separated into as few as 5 pieces and then reassembled into two spheres of the same radius as the original. As counter-intuitive as this seems, it is either false, or the axiom of choice is false.

The need to resolve this conflict became a central problem of mathematics in the 20th century. Many began to doubt the "consistency" of any formulation of axioms that was strong enough to prove statements about arithmetic. One of the 23 unsolved problems of mathematics posed by David Hilbert in 1900 was to determine whether the axioms of arithmetic were consistent. That is, with the axioms, could contradictory statements be proven? In 1931, Kurt Godel proved two groundbreaking theorems, referred to as the first and second incompleteness theorems [5]. The first states that any theory of axioms that is strong enough to prove arithmetic results cannot be consistent and also complete. Either the theory will imply contradictory statements (inconsistent) or the there will exist true statements that are not provable in the system (not complete). His second theorem stated that a system of axioms could not prove its own consistency unless said system was inconsistent. Godel built a model called the Constructible Universe, often referred to as Godel's L, in which AC was niether consistent nor inconsistent. In 1962, Paul Cohen used a method of his own design called forcing to prove that the axiom of choice (as well as the continuum hypothesis) could neither be proved nor disproved using the axioms of Zermelo-Fraenkel [2] [3].

This thesis is concerned with three aspects of the Axiom of Choice (AC). The first is a historical setting and an attempt to organize the thesis so that the reader can understand the timeline of the history of AC. Secondly, I will be showing theorems whose difficulty ranges from trivial to complex, whose proofs in many texts, papers and other sources use the Axiom of Choice in the proof (even though these texts could have avoided using the axiom). Thirdly, I will also give examples of theorems whose proofs require the axiom of choice, that is to say, are equivalent to the axiom of choice.

In section 2 I introduce many basic definitions that will be used throughout the other 5 sections. These include the definition of a relation, a function, a set, the intersection of sets, and the union among others. This section establishes the denotation used throughout the thesis.

In section 3 I present the statement of the Axiom of Choice and then give some basic theorems from different fields that use AC in the proofs. That is, these proofs do not require the axiom but are helpful as an introduction to the concept of what the axiom can allow us to achieve. At the end of the section I will give one theorem which shows a result that not only relies on AC in its proof, but is proved equivalent to AC. In particular, it is proved that every infinite set has a countably infinite subset, the cartesian product of two copies of the natural numbers is countable, the union of countable sets is countable, every irrational number can be "approximated" using rational numbers, and the theorem that AC is equivalent to the statement that the Cartesian product of nonempty sets is nonempty.

In section 4 is the most important section of the thesis. In it I develop the theory of orders so that I can prove the equivalence of the Axiom of Choice to the well-ordering theorem and Zorn's lemma. I then prove the equivalence of AC to Tukey's lemma, the Hausdorff Maximal principle and Kuratowski's lemma (all through Zorn's lemma). This section is the most important because the results in the last three sections of the thesis mostly follow from Zorn's lemma and the other maximal principle, but are thought of as equivalent to the Axiom of Choice.

A complete list of all equivalencies proved in this thesis is given in Appendix A. Results which are not equivalent or whose equivalence is only stated and not proved is contained in Appendix B.

Section 5 is the beginning of three sections on Algebra, General Topology and Real Analysis (in that order) in which the use of the Axiom of Choice is used in the proofs of important results in each of the fields. In particular, in section 5, we will prove that AC implies Krull's theorem, which states that every nonempty ring with unity has a maximal ideal.

In section 6 we show that AC is equivalent to Tychonoff's theorem. Tychonoff's theorem has been refered to as the most important result in General Topology [11]. This result requires many definitions and lemmas to prove and as such, section 6 contains the most material of the last 3 sections of the thesis.

In section 7 we prove, using AC, a result which is refered to as one of the three pillars of Functional Analysis, the Hahn-Banach theorem. The proof uses the Hausdorff maximal principle, which is shown in section 4 to be equivalent to AC, but the Hahn-Banach theorem is weaker than AC since the reverse direction cannot be proved. The Hahn-Banach theorem is equivalent to a weaker form of AC which we will not cover but is called the Boolean Prime Ideal theorem.

The goal of this thesis is to give the motivation for understanding why, as a begining student of mathematics, we should be critical of the axioms that we assume. Mathematicians have been critical of Euclid's 5th axiom of geometry and have thus given new opportunities in the subjects of geometry to study the results which do not rely on that axiom. In the same context, it is important to study the results that rely on the Axiom of Choice, and those that do not. Currently, mathematicians use AC freely and without second thought, but its truth should not be taken because of the beauty of the results which it proves [12]. AC solves many problems of mathematics across many different fields, but as Shoenfield states, "The

more problems a new axiom settles, the less reason we have for believing the axiom is true."
[15]

Preliminaries

The purpose of this section is to set the notation used throughout the thesis. I will assume that the reader has seen a development of formal logic and refer the interested reader to a text in the bibliography by Shoenfield [16].

DEFINITION 2.1. The word **set** will be used to describe a collection of objects and/or elements which may or may not be specifically given.

Most often, capital letters will be used to describe sets $A, B, C, ...$, lower case letters will be used to describe the objects and/or elements of the sets $a, b, x, y, ...$, and script letters will be reserved mostly for collections of sets $\mathcal{A}, \mathcal{B}, \mathcal{C}, ...$

DEFINITION 2.2. We will refer to the set that contains no elements as the **empty set** and will denote it as $\emptyset$.

DEFINITION 2.3. We will say that a set A is a **subset** of a set B if every element of A also belongs to the set B. We abbreviate this as $A \subset B$.

This leaves open the possibility of A containing the same exact elements as B and so we make a distinction by defining a proper subset below.

DEFINITION 2.4. We say that a set A is a **proper subset** of B if every element of A belongs to B but there are elements of B that do not belong to A.

DEFINITION 2.5. The **complement** of a set $A \subset B$ is designated as $A^c = \{x | x \notin A \text{ and } x \in B\}$.

DEFINITION 2.6. The **intersection** of two sets A, B is the set $A \cap B = \{x | x \in A \text{ and } x \in B\}$.Given a collection of sets $\mathcal{A} = \{A_i : i \in I\}$ we define $\cap \mathcal{A} = \cap \{A_i : i \in I\}$

DEFINITION 2.7. The **union** of two sets A, B is the set $A \cup B = \{x | x \in A \text{ or } x \in B\}$. Given a collection of sets $\mathcal{A} = \{A_i : i \in I\}$ we define $\cup \mathcal{A} = \cup \{A_i : i \in I\}$

DEFINITION 2.8. We define an **ordered pair** (x, y) as $(x, y) = \{\{x\}, \{x, y\}\}$.

DEFINITION 2.9. For any A and B, we define the **cartesian product** of the two sets as, $A \times B = \{(x, y) | x \in A \text{ and } y \in B\}$.

We also define the arbitrary cartesian product of a collection of sets X_i such that $i \in I$ and I is any index set. We define,

$$\prod_{i \in I} X_i = \{f : I \to \bigcup_{i \in I} X_i : f(i) \in X_i \text{ for all } i \in I\}$$

If each of the X_i are equal we then denote the product as,

$$\prod_{i\in I} X_i = X^I$$

DEFINITION 2.10.

(i) A **relation** is a set R whose elements are ordered pairs.

(ii) We denote the **domain** of R as $\mathrm{dom}(R) = \{x : \exists y((x,y) \in R)\}$.

(iii) We denote the **range** of R as $\mathrm{ran}(R) = \{y : \exists x((x,y) \in R)\}$.

(iv) We say f is a **function** iff f is a relation and $\forall x \in \mathrm{dom}(f)\exists! y \in \mathrm{ran}(f)((x,y) \in f)$.

*We use the notation $f : A \to B$ for the case when f is a function and $\mathrm{dom}(f) = A$ and $\mathrm{ran}(f) \subset B$. We also say that $f(x)$ is the unique y such that $(x,y) \in f$.

(v) Given R, we define $R^{-1} = \{(x,y) : (y,x) \in R\}$. We say that R^{-1} is the **inverse** of R.

(vi) Given $f : A \to B$, we say that f is **1-1**, iff f^{-1} is a function. We also say that f is **onto** iff $\mathrm{ran}(f) = B$. Lastly, f is a **bijection** iff f is 1-1 and onto.

The Axiom of Choice

In this section we will give the statement of the Axiom of Choice. Some authors give an alternate statement of the axiom, see Kunen [9]. Our version is that stated by Munkres [11]. After the statement, I will give some simple theorems and their proofs which use the Axiom of Choice. I establish the necessary definitions for each of the theorems and conclude the section with the statement that AC is equivalent to the Cartesian Product theorem.

Axiom 9- The Choice Axiom. Suppose that $\mathcal{B}$ is a collection of nonempty sets. Then there exists a function $\mathcal{C} : \mathcal{B} \to \bigcup_{X \in \mathcal{B}} X$ such that $\mathcal{C}(X) \in X$ for each $X \in \mathcal{B}$.

In later proofs, the assumption of AC will be used as shorthand for assuming that there does exist such a function as above, for every collection of nonempty sets.

We make the assumption in this thesis that there exists a set $\mathbb{R}$, called the real numbers, and we assume that the reader is familiar with the algebraic properties of $\mathbb{R}$.

DEFINITION 3.1. A subset A of $\mathbb{R}$ is **inductive** if it contains the number 1, which is a real number, and for every $x \in A$, $x+1 \in A$.

DEFINITION 3.2. Let $\mathcal{A}$ be the collection of all inductive sets of $\mathbb{R}$. We define the **natural numbers**, or $\mathbb{N}$, to be the set,

$$\mathbb{N} = \bigcap_{A \in \mathcal{A}} A$$

DEFINITION 3.3. We denote a **section**, S_n, of the natural numbers to be all of the natural numbers less than n.

DEFINITION 3.4. Let A be a set.

(1) A is **finite** if there exists a 1-1 and onto function between A and a section of $\mathbb{N}$.

(2) A is **countable** if it is finite or there exists a 1-1, onto function between A and $\mathbb{N}$.

(3) A is **uncountable** if it is not countable.

Now that we have defined the notion of countable we can state our first theorem of the thesis. In the proof of this theorem we begin a selection process of elements from a set A in order to create a sequence indexed by the natural numbers. Unfortunately, we give no exact rule or method of selecting these specific elements from the infinite set from which they are selected. The axiom of choice justifies this type of selection.

THEOREM 3.5. Every infinite set has a countably infinite subset.

Proof. Let A be an infinite set. A must then be nonempty so that we can select an element $a_0 \in A$. Since A is infinite we must have that $\{a \in A : a \neq a_0\} \neq \emptyset$, so select an a_1 from this set. Continue in this manner, creating a sequence of distinct elements $a_i \in A$, one for each $i \in \mathbb{N}$. We now have that the set $H = \{a_0, a_1, ...\}$ is countable and $H \subset A$. $\square$

I will not prove the next proposition but it is a standard result in Set Theory and the reader is refered to Munkres [11] for the proof.

PROPOSITION 3.6. Let A be a nonempty set. The following are equivalent,

(1) A is countable.

(2) There exists an onto function $f : \{1,2,...\} \to A$.

(3) There exists a 1-1 function $g : A \to \{1,2,...\}$.

DEFINITION 3.7. A set A has **cardinality** 0, denoted $|A| = 0$, if A is empty. A set A has **cardinality** n, denoted $|A| = n$, if there is a bijection $f : A \to S_{n+1}$. We have that $|A| \leq |B|$ if there is an 1-1 function $g : A \to B$.

LEMMA 3.8. $|\mathbb{N} \times \mathbb{N}| \leq |\mathbb{N}|$.

Proof. We construct a 1-1 function, $g : \mathbb{N} \times \mathbb{N} \to \{1,2,...\}$. Define $g(n,m) = 2^n 3^m$. Assume for contradiction that $g(n,m) = g(p,q)$ for some $(p,q) \neq (n,m)$. If $n < p$ we get that $3^m = 2^{p-n} 3^q$ with $p - n \in \mathbb{N}$. But this claims that 3^m is even, which cannot be the case. If $n > p$, similarly, we get a contradiction and thus it must be true that $n = p$. Now since $3^m = 3^q$, we have $m = q$. Therefore g is 1-1 and $|\mathbb{N} \times \mathbb{N}| \leq |\mathbb{N}|$. □

THEOREM 3.9. Let $|A_i| \leq \omega \ \forall i \in \mathbb{N}$. Then $|\bigcup_{i \in \mathbb{N}} A_i| \leq \omega$.

Proof. Since each $|A_n| \leq \omega$, there exists a function $f_i : A_i \to \mathbb{N}$ for every $i \in \mathbb{N}$ such that each f_i is 1-1. For each $b \in \bigcup_{i \in \mathbb{N}} A_i$, we know that there exists a least i such that $b \in A_i$ (b could be a member of multiple A_i's). We now define $F : \bigcup_{i \in \mathbb{N}} A_i \to \mathbb{N} \times \mathbb{N}$ by $F(b) = (i, f_i(b))$, where i is the least such that $b \in A_i$.

Since $|\mathbb{N} \times \mathbb{N}| \leq \omega$, we use the above lemma to find a 1-1 function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Now $g \circ F : \bigcup_{i \in \mathbb{N}} A_i \to \mathbb{N}$ is 1-1 and thus $\bigcup_{i \in \mathbb{N}} A_i$ is countable. □

The above proof used the Axiom of Choice, although subtly. By Proposition 6, if a set A is countable there exists an onto function from the natural numbers to A. The proposition does not say how many said functions exists; there could be uncountably many. For example, the number of onto functions $f : \mathbb{N} \to \mathbb{N}$ is uncountable. In the proof of the theorem, for each $n \in \mathbb{N}$ we selected one such function, countably many times. This type of arbitrary selection, which occurs arbitrarily many times, is the common theme of proofs that use the Axiom of Choice.

I will now present some elementary analysis definitions so that I can follow with a simple proof from a first course in analysis of why there always exists a sequence of rational numbers converging to any irrational number.

DEFINITION 3.10. Let $a \in \mathbb{R}$ and $\varepsilon > 0$.

(1) We define an **ε-neighborhood** of the point a as, $N_\varepsilon(a) = \{x \in \mathbb{R} : |x - a| < \varepsilon\}$.

(2) A **sequence** in $\mathbb{R}$ is a function $f : \mathbb{N} \to \mathbb{R}$. For each $n \in \mathbb{N}$ we denote the nth **term** of the sequence f as a_n where $a_n = f(n)$, and for short, denote f by $\{a_n\}_{n \in \mathbb{N}}$.

(3) A sequence $\{a_n\}_{n \in \mathbb{N}}$ in $\mathbb{R}$ is said to **converge** to a point $a \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $a_n \in N_\varepsilon(a)$ for all $n \geq N$.

THEOREM 3.11. Let p be an irrational number. There exists a sequence $\{q_n\}_{n \in \mathbb{N}}$, such that each q_n is a rational number, that converges to p.

Proof. Choose $q_1 \in N_1(p), q_2 \in N_{\frac{1}{2}}(p), \ldots$ and in general choose $q_n \in N_{\frac{1}{n}}(p)$ for all $n \in \mathbb{N}$. The claim is that $\{q_n\}_{n=1}^{\infty}$ is a sequence of rational numbers that converge to p.

Let $\varepsilon > 0$. Choose $N > \frac{1}{\varepsilon}$. Then for all $n \geq N$ we have that $q_n \in N_\varepsilon(p)$. This is due to the fact that $\frac{1}{N} < \varepsilon$ and $a_n \in N_{\frac{1}{n}}(p) \subset N_{\frac{1}{N}}(p) \subset N_\varepsilon(p)$. By definition, $\{q_n\}_{n=1}^{\infty}$ converges to p. $\square$

AC was used in the above proof with the selection of a rational number within each ε-neighborhood. There are countably many said rational numbers to choose from in any given neighborhood on the real line and we made infinitely many of these choices, one for each $n \in \mathbb{N}$. Without AC we would need a more constructive choice. If we denote the set of the rational numbers in the first neighborhood D, we see that the set D is **denumerable** (i.e. there exists a 1-1 function $f : \mathbb{N} \to D$). Then we can select from the first neighborhood, $f(1)$, and from the second neighborhood, $f(n)$ such that $n \neq f(1)$, etc. There is always an $f(n)$ with this property since f is onto. This is the idea of a more constructive proof. It is an interesting question to ask whether we should always search for a proof that does not use AC because of the uncertainty of the truth value of AC. Some mathematicians take this approach.

We now show an equivalent form of the Axiom of Choice which will be used later to prove the equivalency of Tychonoff's theorem and AC.

THEOREM 3.12. AC holds if and only if the cartesian product of nonempty sets is nonempty.

Proof. In the forward direction, let $\mathcal{B} = \{A_j | j \in J\}$ such that every $A_j \neq \emptyset$. By the Choice Axiom, $\exists \mathcal{C}(C : \mathcal{B} \to \bigcup_{j \in J} A_j$ such that $\mathcal{C}(A_j) \in A_j$. Let $f : J \to \mathcal{B}$ by $f(j) = A_j, \forall j \in J$.

Now we have that $\mathcal{C} \circ f : J \to \bigcup_{j \in J} A_j$ such that $(\mathcal{C} \circ f)(j) \in A_j$, $\forall j \in J$. This function is an element of the cartesian product of the A_j.

Assuming the latter, let $\mathcal{B}$ be a collection of nonempty sets indexed by J. We can say that $\mathcal{B} = \{A_j | j \in J\}$. Define a function $g : \mathcal{B} \to J$ by $g(A_j) = j$. By our assumption, $\prod_{j \in J} A_j$ is nonempty and so there exists an $f : J \to \bigcup_{j \in J} A_j$ such that $f(j) \in A_j, \forall j \in J$. Consider $(f \circ g) : \mathcal{B} \to \bigcup_{j \in J} A_j$. This is a choice function on $\mathcal{B}$. $\square$

Order Theory

Order theory in mathematics is concerned with the ordering of sets. We will define several different types of orders, some of which put more restriction than other on the set we are ordering. These will include the partial order, the total order, and the well-order. A natural question that one might have is, or any set, is there an order on the set, and if so, what type of order? We will prove in this section that the axiom of choice implies that every set can be well-ordered, which is a type of order that we will define, and in fact the converse holds; If every set can be well-ordered, then AC is true. We will give the proof for this statement.

We then state Zorn's lemma, which introduces maximal elements. What type of sets have maximal elements and which do not? Maximality is a middle ground in between finiteness and infiniteness. All finite sets have a maximal element but not all infinite sets have one. Having a maximal element gives some idea to how the set is structured.

We will define what it means for a family of sets to have "finite character", which is different form being finite and was proposed by the mathematician Tukey. We will then use Zorn's lemma to prove Tukey's lemma and then Tukey's lemma to prove the Hausdorff maximal principle. We then present the proof that the Hausdorff maximal principle implies Zorn's lemma, establishing the equivalence of all three, and thus their equivalence to AC and the well-ordering theorem.

Lastly, we will prove the equivalence of Zorn's lemma with the Kuratowski lemma.

DEFINITION 4.1. Let A be a set. A relation R is called a **partial order** if;

(1) for any $a \in A$ it is true that aRa

(2) for any $a, b, c \in A$ with aRb and bRc, then aRc.

A set A together with a partial order R is often referred to as a partially ordered set (A, R), or poset for short.

DEFINITION 4.2. Let A be a set. A relation R is called a **total order** if;

(1) R is a partial order on A.

(2) if for any $a, b \in A$, either aRb, bRa, or $a = b$.

DEFINITION 4.3. Let A be a set. A total order R is called a **well-order** if the order has the property that for every nonempty subset $B \subset A$, there is a R-least element of B. Formally, $\exists b \in B$ such that $\forall c \in B$, bRc.

In addition, the set A is then said to be **well-ordered** by R.

The well-ordering theorem. For any set X, there is a relation R such that R is a well-order for X.

First we now show that the well-ordering theorem implies AC.

THEOREM 4.4. If for any set X, there is a relation R such that R is a well-order for X, then AC.

Proof. Let $\mathcal{B}$ be a collection of nonempty sets. Let $X \in \mathcal{B}$. By the Well-ordering Theorem, X can be well-ordered by some relation R. Define a function $\mathcal{C} : \mathcal{B} \to \bigcup_{X \in \mathcal{B}} X$ by,

$$\mathcal{C}(X) = \text{The } R\text{-least element of } X.$$

We can see that $\mathcal{C}(X) \in X$ since the R-least element of X is in X, and thus $\mathcal{C}$ is the desired choice function for $\mathcal{B}$. □

I now define upper (lower) bound, maximal element, and chain, so that I can state Zorn's lemma.

DEFINITION 4.5. An **upper bound** for an ordered set A is an element u, not necessarily belonging to A, such that for any $a \in A$, aRu. Likewise, a **lower bound** for a set A is an element b such that for any $a \in A$, aRb.

DEFINITION 4.6. A **maximal** element $a \in A$ is one such that there is no element $b \in A$ such that $a < b$, where the symbol $<$ is used to describe the case if aRb but $b \neq a$.

DEFINITION 4.7. A **chain** of a partially ordered set A is a subset of A that is totally ordered.

LEMMA 4.8. (Zorn's Lemma) Let X be a partially ordered set in which every chain has an upper bound. Then X has at least one maximal element.

I now move one step closer to showing that AC, the well-ordering theorem, and Zorn's lemma are equivalent. I have shown above that the well-ordering theorem implies the Axiom of Choice. Now I show that Zorn's Lemma implies the well-ordering theorem.

THEOREM 4.9. Suppose that for every X which is a partially ordered set in which every chain has an upper bound, then X has at least one maximal element. Then we have that for any set X, there is a relation R that well-orders X.

Proof. Let X be a set. We will show that there is an order relation that well orders X. Let W be the collection of all well-ordered subsets of X, i.e. $W = \{(A,R) : A \subseteq X, R \text{is a well-order on } A\}$. Let's put a partial order on W. We say that $(A,R) \leq_r (B,S)$ iff

(1) $A \subseteq B$

(2) For every $a_1, a_2 \in A$, $a_1 R a_2$ implies $a_1 S a_2$.

(3) aSb, for all $a \in A$ and for all $b \in B - A$.

Let $\mathcal{C}$ be a chain in W. We will show that $\mathcal{C}$ has an upper bound and then apply Zorn's Lemma to W. Define $C = (A^*, R^*)$ where $A^* = \bigcup\{A : (A,R) \in \mathcal{C}\}$ and R^* to be the relation such that if $a, b \in A^*$ and $a \in A_0, b \in A_1$ with $A_0 \subseteq A^*, A_1 \subseteq A^*$, and WLOG $A_0 \subseteq A_1$, then $a\, R^* b$ if and only if $a\, R_{A_1}\, b$, where R_{A_1} is the relation that well-orders A_1.

Then we can see that for any $(A,R) \in \mathcal{C}$ we have $(A,R) \leq_r (A^*, R^*)$ since,

(1) $A \subseteq A^*$

(2) R^* restricted to A, is R

(3) aRa_0 for all $a \in A$ and for all $a_0 \in A^* - A$.

So every chain in W has an upper bound and Zorn's Lemma gives us a maximal element (X_0, R_0) of W. But $X_0 = X$, since if we assume not, we can select an $h \in X - X_0$ and then $(X_0 \cup \{h\}, R_0')$ where R_0' is just R_0 restricted to X_0 with the addition of $xR_0'h$ for all $x \in X_0$. Then we would have $(X_0, R_0) \leq_r (X_0 \cup \{h\}, R_0')$ by definition which contradicts that (X_0, R_0)=(X, R_0) and so R_0 is a well-order on X. □

We now define the property of finite character and state Tukey's Lemma. We prove Tukey's Lemma using Zorn's Lemma, Tukey's is then used to prove the Hausdorff maximal principle, which is in turn used to prove Zorn's Lemma. This establishes the equivalence of all 3 of these maximal principles.

DEFINITION 4.10. A family of sets $\mathcal{F} \subseteq \mathcal{P}(X)$ for some set X, is said to be of **finite character** if;

(1) For each $A \in \mathcal{F}$, every finite subset of A belongs to $\mathcal{F}$.

(2) $B \subseteq \mathcal{P}(X)$ and every finite subset of B belongs to $\mathcal{F}$ implies $B \in \mathcal{F}$.

Tukey's Lemma Let $\mathcal{F}$ be a family of finite character. Then for any set B an element of $\mathcal{F}$, there exists a set $C \in \mathcal{F}$ such that C contains B and C is maximal with respect to the subset relation.

We now prove Tukey's Lemma using Zorn's lemma.

THEOREM 4.11. If for every partially ordered set in which every chain has an upper bound there is a maximal element, then for all $\mathcal{F}$, a family of finite character, and for any set X an element of $\mathcal{F}$, there exists a set $Y \in \mathcal{F}$ such that Y contains X and Y is maximal with respect to the subset relation.

Proof. Let $\mathcal{F} = (F_i)_{i \in I}$ be a family of subsets of $\mathcal{P}(X)$ for some set X such that $\mathcal{F}$ is of finite character. By Zorn's Lemma it is enough to show that every chain in $(F_i)_{i \in I}$ has an upper bound. We partially order the family by inclusion ($\subseteq$), i.e. $F_i \leq F_j$ iff $F_i \subseteq F_j$ for all $i, j \in I$. With this partial order we can see that for any element of the family, say $B \in \mathcal{F}$, the set $\{B\}$, is a chain by the reflexive property, and thus it will suffice to show that there is an upper bound to this chain, which is an element of $\mathcal{F}$, and hence C.

What we prove is more general. Let $(F_i)_{i\in J\subseteq I}$ be any chain in our family. We claim that this chain has an upper bound, $\bigcup(F_i)_{i\in J\subset I}$, and more importantly the upper bound is a member of the family.

Let D be a finite subset of $\bigcup(F_i)_{i\in J\subset I}$. Then every element $d \in D$ satisfies $d \in F_i$ for some $i \in J$. Since D is finite and since $(F_i)_{i\in J\subset I}$ is a chain, there is a largest F_k with $k \in J$ such that every element of D is in F_k. That is to say, $D \subset F_k$. But F_k is in the family of finite character and thus all of its subsets are elements of the family, i.e. $D \in (F_i)_{i\in I}$.

We have now that all of the finite subsets of $\bigcup(F_i)_{i\in J\subset I}$ are in the family and so $\bigcup(F_i)_{i\in J\subset I} \in (F_i)_{i\in I}$ by the property of finite character. Since $\bigcup(F_i)_{i\in J\subset I}$ is an upper bound for our arbitrary chain, we conclude that all chains in the family have an upper bound and therefore by Zorn's Lemma our family has a maximal element. □

Now I will show that Tukey's Lemma implies the Hausdorff maximal principle.

The **Hausdorff maximal principle** states that if A is a partially ordered set, then there exists a maximal chain contained in A.

THEOREM 4.12. Let $\mathcal{F}$ be a family of finite character. If for any set X an element of $\mathcal{F}$, there exists a set $Y \in \mathcal{F}$ such that Y contains X and Y is maximal with respect to the subset relation, then for every partially ordered set A, A has a maximal chain.

Proof. Let A be a poset. We will show that A has a maximal chain. Let $\mathcal{F}$ be the family of all subsets of A that are totally ordered (chains).

(1) Let $B \in \mathcal{F}$. Then B is a chain. Every finite subset of B is also a chain and thus belongs to $\mathcal{F}$.

(2) Let all of the finite subsets of a set B belong to $\mathcal{F}$. Then these subsets are chains. Let $b_1, b_2 \in B$. $\{b_1, b_2\}$ is a finite subset of B and thus in $\mathcal{F}$ (and totally ordered). So either $b_1 \leq b_2$ or $b_2 \leq b_1$. So B is a chain and thus in $\mathcal{F}$.

By (1) and (2), $\mathcal{F}$ is of finite character and has a maximal element by Tukey's Lemma. Thus A has a maximal chain. □

What follows finishes off the equivalence of Zorn's Lemma, Tukey's Lemma, and the Hausdorff Maximal Principle (i.e. the Hausdorff maximal principle implies Zorn's lemma). Later, in the section on functional analysis, the Hahn-Banach Theorem will be proven using the Hausdorff Maximal Principle.

THEOREM 4.13. If every partially ordered set A has a maximal chain, then every partially ordered set A with the property that every chain has an upper bound, has a maximal element.

Proof. Let A be a poset in which every chain has an upper bound. By the HMP there exists a maximal chain in A, call it B. B has an upper bound, c, by assumption.

Claim: c is a maximal element of A.

Assume for contradiction that there exists $d > c$ with $d \in A$. Then $B \cup \{d\}$ is a chain in A since $d > b$ for all $b \in B$. $B \cup \{d\}$ is a chain in A that strictly contains B since $d \in B$. This contradicts the maximality of B and so c is a maximal element of A. □

We have proven the equivalence of Zorn's lemma, Tukey's lemma and the Hausdorff maximal principle. We have also shown that Zorn's lemma implies the well-ordering theorem, and that the well-ordering theorem implies AC. What remains is to show that AC implies Zorn's lemma and then everything presented in this section so far will be proved equivalent. We prove two lemmas first, then show using the two lemmas that AC implies Zorn's modified lemma, then show that Zorn's modified lemma implies the Hausdorff maximal principle, which above was shown to imply Zorn's lemma. We begin with a definition.

DEFINITION 4.14. Given a set L ordered by R, we use $\leq$, and $l_1, l_2 \in L$, we say that l_2 is a **direct successor** of l_1 if,

(1) $l_1 < l_2$ and,

(2) there does not exist an $l_3 \in L$ with $l_1 < l_3 < l_2$.

LEMMA 4.15. Let L be a partially ordered set in which every chain has a least upper bound. Then there does not exist a function $g : L \to L$ with the property that for every $x \in L$, $g(x)$ is a direct successor of x.

Proof. Assume that such a g exists. Fix an element $z \in L$ and let $L_0 = \{x \in L \ : z \leq x\}$. Then L_0 is a partially ordered set in which every chain has an upper bound and also for any $x \in L_0$, $g(x)$ is a direct successor of x.

Define a **tower** as any subset A of L_0 such that the following hold,

(1) $z \in A$,

(2) $g(A) \subseteq A$,

(3) For each chain in A, A contains the least upper bound of the chain.

Claim: The set of all towers is nonempty. This is true since L_0 is a tower. Obviously $z \in L_0$ since $z \geq z$ and thus meets the membership requirements of L_0. We have $g(L_0) \subseteq L_0$ since if $x \in L_0$, then $g(x) \geq g(z)$. So if $y \in g(L_0)$, then $y \geq g(z) > z$, and so by definition of L_0, $y \in L_0$. Lastly, for every chain in L_0, L_0 contains the least upper bound of the chain by assumption.

Let T be the intersection of all towers of L_0.

Claim: T is a tower. Since $z \in A$ for every tower A (by (1)), $z \in T$. Since $g(A) \subseteq A$ for every tower A and $T \subseteq A$ for every A, we have $g(T) \subseteq g(A) \subseteq A$. So if we have $g(T) \subseteq A$ for every tower of L_0, A, then $g(T) \subseteq \cap A = T$ by definition. Lastly, let C be a chain in T. Then C is in every tower of L_0 and so every tower of L_0 contains the least upper bound of C. Formally, if we let l.u.b.(C) denote the least upper bound of C then we have that l.u.b.$(C) \in \cap A = T$. So $C \in T$.

Claim: T is a chain.

Assuming the claim above that T is a chain, since we have shown that T is a tower, property (3) from above holds, and so T contains its least upper bound, call it t. By property (2), $g(t) \in T$. By assumption though, $g(t)$ is a direct successor of t, and so $g(t) > t$, the least upper bound of T, and thus $g(t)$ cannot be in T, which is a contradiction.

To prove the claim let $V = \{x \in T : \text{ for all } y \in T, \text{ either } x \leq y, \text{ or } y \leq x\}$. We have then that V is a chain and we would like to show that $V = T$. Having that T contains V, let's show that V contains T. If V is a tower then we have $T \subseteq V$ by definition.

The 3 properties of a tower for L_0 must hold for V for it to be a tower.

The first property, $z \in V$, is proved since z is in every tower and thus in T, and T is a subset of L_0 which was defined as all of the elements which are greater than or equal to z. So z is comparable with all of the elements T, and is thus in V.

The second property. Let $b \in V$. We must show that $g(b) \in V$. Define the set $W = \{w \in T : w \leq b \text{ or } g(b) \leq w\}$. Remember though that $b \in V \subseteq T \subseteq A \subseteq L_0$, where A is any tower of L_0. So $g(b)$ is a direct successor of b by assumption and so there are no elements of L_0 in between $g(b)$ and b. So $W = T$ since all of the elements of T that are less than or equal to b, or greater than or equal to $g(b)$ is all of T.

We conclude that for all $u \in T$, we have either $u \leq b$ or $g(b) \leq u$. That is to say that $u \leq g(b)$ and $g(b) \leq u$ for any $u \in T$. Thus $g(b)$ fits the definition of a member of V and that is what we wanted to show.

The third property. Let $\{C_\alpha\}$ be a chain in V and show that the least upper bound of $\{C_\alpha\}$ is also in V. Let b be the least upper bound of $\{C_\alpha\}$ and let u be arbitrary in T. We show that b is comparable with u.

We know that each C_α is in V and thus comparable with every element of T.

Case 1: If at least one of the $C_\alpha \geq u$, then $b \geq u$.

Case 2: If all of the $C_\alpha \leq u$, then u is an upper bound for our chain. Since b is the least upper bound, $b \leq u$.

Either way, b is comparable with u and so $b \in V$.

We have now that V is a tower, and so $T \subseteq V$, and hence $V = T$. Remember though that V was defined as a chain and thus T is a chain. So the least upper bound of T, call it t, is in T by property (3) of a tower; it follows then by property (2) that $g(t)$ is in T. But if t is the least upper bound of T, and $g(t)$ is a direct successor of t, $g(t)$ cannot be in T, and thus our contradiction.

So there does not exist a function $g : L \to L$ such that for every $x \in L$, $g(x)$ is a direct successor of x. This proves the first lemma.

□

Before proceeding to the second lemma, I present an example of a poset in which an element has more than one direct successor.

EXAMPLE 4.16. Let M be a partially ordered set. Let L be the collection of all chains which are contained in M. Then L is a partially ordered set by inclusion.

That is, we assume that $a_i \leq a_{i+1}$ for any $i \in \mathbb{N}$; then the following are all chains in L,

(1) $A = \{a_1, a_2, a_3\}$,

(2) $B = \{a_1, a_2, a_3, a_4\}$,

(3) $C = \{a_1, a_3, a_4\}$,

(4) $D = \{a_1, a_2, a_3, a_7\}$,

we have that $A \leq B, A \leq D$ but $A \nleq C$. B is a direct successor of A since there is no other chain greater than A, not equal to A, but less than B, not equal to B. Note also though, D is a direct successor to A for the same reasons.

LEMMA 4.17. . Let L be a partially ordered set in which every chain has a least upper bound. Then there does not exist a function $g : L \longrightarrow L$ with the property that for every $x \in L$, $g(x) > x$.

Proof. : Let L be a partially ordered set in which every chain has a least upper bound and assume that there does exist such a function g. Consider the set of all chains contained in L, call it $\mathfrak{L}$. Then $\mathfrak{L}$ is partially ordered by inclusion. We define $h : \mathfrak{L} \rightarrow \mathfrak{L}$;

(1) If $C \in \mathfrak{L}$ and has a greatest element x, then let $h(C) = C \cup \{g(x)\}$. We have then that $h(C)$ is a direct successor of C. Note that $g(x)$ need not be a direct successor of x for this to be true, rather just $g(x) > x$. Also, $h(C)$ is a chain since $g(x) > x \geq c$ for all $c \in C$.

(2) If $C \in \mathfrak{L}$ has no greatest element then we let $h(C) = C \cup \{\alpha\}$, where α is the least upper bound of C. C has a least upper bound by assumption since C is a chain in L. Now we note that $h(C)$ is a direct successor of C and that $h(C)$ is a chain.

Claim: $\mathfrak{L}$ is a partially ordered set in which every chain has a least upper bound. A chain D in $\mathfrak{L}$ is a collection of chains from L that are all comparable and thus a least upper bound must be a chain as well. Define $V = \bigcup_{C \in D} C$, such that C is a chain. Then V is a chain for if $\alpha, \beta \in V$ the $\alpha \in C'$ for some $C' \in D$ and $\beta \in C''$ for some $C'' \in D$. WLOG, $C' \leq C''$ and so $\alpha, \beta \in C''$ and thus comparable since C'' is a chain. By definition, V is the least chain greater than or equal to every element of D, for assume that there is a chain E greater than or equal to all elements of D but less than V. Then for any chain $C \in D$ we have that $C \subseteq E \subseteq V$, but also any element of V is in some chain $C \in D$ and thus in E, so $V \subseteq E$, and $E = V$.

To reiterate, $\mathcal{L}$ is a partially ordered set in which every chain has a least upper bound. Also, there exists $h : \mathcal{L} \to \mathcal{L}$ such that for every $C \in \mathcal{L}$, $h(C)$ is the direct successor of C. This contradicts Lemma 1 and thus there must not exist a function $g : L \to L$ with the property that for every $x \in L$, $g(x) > x$.

□

Now that we have proved the two lemmas needed, I state a modified version of Zorn's lemma and show that the Axiom of Choice proves it.

Zorn's modified lemma.

Let L be a partially ordered set in which every chain has a least upper bound. Then L has a maximal element.

THEOREM 4.18. . The Axiom of Choice implies Zorn's modified lemma.

Proof. : Let L be a partially ordered set in which every chain has a least upper bound. For each $x \in L$, define $M_x = \{y \in L : y > x\}$.

Case 1: One of the M_x is empty. Then x is a maximal element since there does not exist $y \in L$ with $y > x$.

Case 2: None of the M_x are empty. By the Axiom of Choice there exists a function defined on the collection containing every M_x such that $f(M_x) \in M_x$ for all $x \in L$. Let $g : L \to L$ by $g(x) = f(M_x)$. Then we have that $g(x) = f(M_x)$ is an element of M_x. That is, $g(x) > x$. Our function g is defined for all $x \in L$ and so we contradict Lemma 2. Thus Case 2 does not occur.

Thus only Case 1 occurs, and we have a maximal element of L. □

We have shown using two lemmas that AC implies Zorn's modified lemma. The next proof shows that Zorn's modified lemma implies the Hausdorff maximal principle.

Hausdorff maximal principle. Let L be a partially ordered set. Then there is a maximal chain in L.

THEOREM 4.19. Zorn's modified lemma implies the Hausdorff maximal principle.

Proof. Let L be a partially ordered set. We aim to show that there is a maximal chain in L. Let H be the set of all chains of L. Then H is partially ordered by inclusion and each chain $\{C_\alpha\} \subseteq H$ has a least upper bound, namely $\bigcup_{C \in C_\alpha} C$ as shown in the proof of Lemma 2. By Zorn's modified lemma, H has a maximal element (a chain of L). This is a maximal chain of L, and our theorem is proved. $\square$

We have from Theorem 4.13 that the Hausdorff maximal principle implies Zorn's lemma. Now everything in the section has been shown equivalent, let us mention one last lemma due to Kuratowski, and tie it into the cycle of equivalences. Consider the Hausdorff maximal principle, which states that in every poset, there is a maximal chain. If we have an arbitrary chain in the poset, does it necessarily have to be contained in that maximal chain (or any of the maximal chains)? This question brings us to Kuratowski's lemma.

LEMMA 4.20. (Kuratowski's Lemma) Any chain in an ordered set is contained in a maximal chain.

THEOREM 4.21. Kuratowski's lemma is true if and only if Zorn's Lemma is true.

Proof. $\Rightarrow$ Suppose that P is a poset such that every chain has an upper bound and let C be a chain in P. By Kuratowski's lemma, C can be extended to a maximal chain C', which, by assumption has an upper bound $a \in P$ since it is a chain. Assume (for contradiction) that there is a $b \in P$ such that $a < b$. This then means that $b \notin C'$.

Claim: $C' \cup \{b\}$ is a chain. Every $c \in C'$ satisfies $c \leq a$. Also $a \leq b$. By transitivity of P, $c \leq b$ for all $c \in C'$. Thus $C' \cup \{b\}$ is a chain (extending C'), which contradicts that C' was maximal. Thus there is no $b \in P$ such that $a < b$ and thus, a is a maximal element in P.

$\Leftarrow$ Suppose P is a poset and C is a chain in P. If $P = \emptyset$, then our claim is true. We assume that $P \neq \emptyset$. Let $\mathcal{P}$ be the set of all chains in P that extend C. We want to show that $\mathcal{P}$ has a maximal element, i.e. a maximal chain that extends C. $\mathcal{P}$ is a poset in P since it is partially ordered by inclusion. Let $\mathcal{C}$ be a chain in $\mathcal{P}$ (i.e. a chain of chains) and let $C' = \bigcup \mathcal{C}$. We claim that C' is an upper bound of $\mathcal{C}$ such that $C' \in \mathcal{P}$.

Claim 1: C' is a chain in P: Let $x, y \in C'$ and we show that either $x \leq y$, or $y \leq x$. Since C' is a union of chains we know that $x \in A$ and $y \in B$ for some $A, B \in \mathcal{C}$. But $\mathcal{C}$ is a chain in $\mathcal{P}$ and thus $A \subseteq B$ (and thus $x, y \in B$) or $B \subseteq A$ (and thus $x, y \in A$). But if x and y belong to the same chain then we have that either $x \leq y$ or $y \leq x$.

Claim 2: C' is in $\mathcal{P}$: For C' to be in $\mathcal{P}$ we must show that C' is a chain that contains (extends) C. But C' is a chain from above. Let $a \in C$; then $a \in A$ for every $A \in \mathcal{P}$ and thus $a \in A$ for every $A \in \mathcal{C}$. So $a \in C'$. Thus $C \subseteq C'$.

Claim 3: C' is an upper bound of $\mathcal{C}$: Let $a \in A$ for any $A \in \mathcal{C}$. Then $a \in \bigcup \mathcal{C} = C'$ and thus $A \subseteq C'$.

We have proved that for any chain $\mathcal{C} \subseteq \mathcal{P}$, $\bigcup \mathcal{C}$ is an upper bound. Applying Zorn's Lemma we see that $\mathcal{P}$ has a maximal element, that is to say, is a maximal chain containing C. $\square$

Algebra

The Axiom of Choice has important consequences in Abstract Algebra when we consider its equivalent, Zorn's Lemma. Before Zorn's Lemma was published in 1935, algebraists used AC in the form of the well-ordering theorem. In 1905, Hamel gave a basis for $\mathbb{R}$ over the rational numbers [Moore]. Another result of importance in algebra is Krull's Theorem which states that every nonempty ring with unity has a maximal ideal. In this section we prove that Zorn's lemma implies Krull's theorem and give the references for the reverse direction which was proved first in 1979 by Hodges and then simplified in 1994 by Banaschewski.

DEFINITION 5.1. A **binary operation** $*$ on a set G is a function from $G \times G$ to G and for all $a, b \in G$ we write $a * b$ instead of $*(a, b)$.

DEFINITION 5.2. A **group** is an ordered pair $(G, *)$ where G is a set and $*$ is a binary operation such that,

(1) For all $a, b, c \in G$ we have $(a * b) * c = a * (b * c)$.

(2) For all $a \in G$ there exists an element $e \in G$ refered to as the **identity** element of G that satisfies $a * e = e * a = a$.

(3) For each $a \in G$ there is an element of G denoted a^{-1}, refered to as the **inverse** of a, that satisfies $a * a^{-1} = a^{-1} * a = e$ where e is the identity element of G.

We say that a group G is **abelian** if for any $a, b \in G$, $a * b = b * a$.

DEFINITION 5.3. A **ring** is a set together with two binary operations $+$ and $\times$,which are called addition and multiplication respectively, satisfying,

(1) $(R, +)$ is an abelian group.

(2) Multiplication is associative, i.e. $(a \times b) \times c = a \times (b \times c)$ for all $a, b, c \in R$.

(3) The right and left distributive laws hold in R: i.e. for all $a, b, c \in R$ we have,

$$(a+b) \times c = (a \times c) + (b \times c)$$

and also,

$$a \times (b+c) = (a \times b) + (a \times c).$$

DEFINITION 5.4. A ring R is said to be a ring with **unity** if there exists an element denoted 1, such that $1 \in R$ and for all $a \in R$,

$$1 \times a = a \times 1 = a.$$

DEFINITION 5.5. Let R be a ring and $I \subset R$ and $r \in R$. Then I is an **ideal** in R if

(1) I is a subring of R

(2) For any $i \in I$, $ri \in I$ and $ir \in I$.

THEOREM 5.6. Zorn's lemma implies that if R is a ring with unity and I is a proper ideal of R, then there exists a maximal ideal M such that $I \subseteq M \subsetneq R$.

Proof. Let $B = \{J | J \text{is an ideal and } I \subseteq J \subsetneq R\}$. Then B is a partially ordered set (ordered by inclusion), I is trivially in B and so $B \neq \emptyset$.

Claim: Every nonempty chain $\{A_\alpha\}_{\alpha \in J}$ in B has an upper bound, namely $A' = \bigcup_{\alpha \in J} A_\alpha$.

(1) A' is an ideal. Let $a, b \in A'$. Then $a \in A_\alpha, b \in A_\beta$ for some $\alpha, \beta \in J$. Without loss of generality let $A_\alpha \subseteq A_\beta$. Then $a, b \in A_\beta$. Thus by the ideal properties of A_β, $a + b \in A_\beta \subseteq A'$.

Let $r \in R, a \in A'$. Then $a \in A_\alpha$ for some $\alpha \in J$. Thus $ra \in A_\alpha \subseteq A'$ and $ar \in A_\alpha \subseteq A'$.

(2) $A' \subsetneq R$. If $A' = R$, then $1 \in A'$ and this would imply that $1 \in A_\alpha$ for some $\alpha \in J$. But $1 \in A_\alpha$ implies $A_\alpha = R$ and this contradicts the definition of B for any $\alpha \in J$.

(3) $I \subseteq A'$. This is true since $\{A_\alpha\}_{\alpha \in J} \neq \emptyset$ by assumption and $I \subseteq A_\alpha$ for all $\alpha \in J$ since each A_α is an "extension" of I be definition of $A_\alpha \subseteq B$. Thus $I \subseteq A'$.

Thus $A' \in B$ and is an upper bound for our nonempty chain $\{A_\alpha\}_{\alpha \in J}$. By Zorn's Lemma, B has a maximal element M. In other words, M is an ideal such that $I \subseteq M \subsetneq R$ and is maximal with this respect. $\square$

The most commonly seen use of the above result is Krull's Theorem which was proven in 1929 by Wolfgang Krull. Krull was a German mathematician whose research interest was in commutative algebra and whose teaching career saw him advising 40 students in his 41 years between the University of Erlangen-Nï£¡rnberg and the Universitï£¡t Bonn.

Krull's Theorem is actually equivalent to the above theorem, which is equivalent to AC/Zorn's Lemma. The backwards direction, Krull's Theorem $\Rightarrow$ Zorn's Lemma, can be seen in a paper by Wilfrid Hodges (1979) or in a shorter and more modern paper by Banaschewski in 1994 (see bibliography).

COROLLARY 5.7. Krull's Theorem.

Every nonempty ring R with unity has a maximal ideal.

Proof. Let $I = \{0\}$, the zero ideal. By the theorem above, there exists M such that $I \subseteq M \subseteq R$ and M is a maximal ideal. Since every nonempty ring R with unity has the zero ideal, every such R has a maximal ideal. □

Two things of interest are (1) that the associativity of multiplication was not used and could be dropped from the assumptions to give a theorem about non-associative rings, and (2) it is necessary for the ring to have unity since in general, rings need not have maximal ideals. Take any abelian group which has no maximal subgroups and define $ab = 0$ for all $a, b \in R$; in such a ring the ideals are exactly the subgroups and thus there are no maximal ideals [Dummit and Foote].

General Topology

In this section we define the basic definitions of General Topology that are needed to prove that Tychonoff's theorem is equivalent to the Axiom of Choice. We begin with the definition of a topology, a basis, and a subbasis. Then we define a topology on the product of sets, called the product topology.

DEFINITION 6.1. A **topology** on a set X is a collection of subsets of X, usually denoted τ, that have the following properties:

(1) $\emptyset, X \in \tau$.

(2) Let $\mathbb{U}$ be a collection of elements of τ. Then $\bigcup \mathbb{U} \in \tau$.

(3) Consider $U_1, U_2, \ldots, U_n$ where each $U_i \in \tau$. Then $\bigcap_{i=1}^{n} U_i \in \tau$.

For a set X and a topology τ, we denote $O \subset X$ as an **open set** if $O \in \tau$. We also refer to a set X as a **topological space** if there is a topology τ on X.

DEFINITION 6.2. Let X be a set. We denote a collection $\mathbb{B}$ of subsets of X as a **basis** for a topology on X if the following properties hold:

(1) For each $x \in X$, $\exists B \in \mathbb{B}$ such that $x \in B$.

(2) Let $B_1, B_2 \in \mathbb{B}$. Then if $x \in B_1 \cap B_2$, there exists $B_3 \in \mathbb{B}$ with $x \in B_3$ and $B_3 \subseteq B_1 \cap B_2$.

We can now state that every basis generates a topology defined by the open sets as follow: A subset O of X is open in X if for every $x \in O$ there is a basis element B such that

$x \in B \subseteq O$. The fact that this collection is a topology is a standard result and the reader is referred to Munkres [11] for the proof.

DEFINITION 6.3. A **subbasis** S for a topology on X is a collection of subsets of X whose union equals X.

As with the basis, a subbasis also generates a topology in which the open sets are defined to be the collection of all unions of finite intersections of elements of S. The proof that this collection is a topology is a standard result and the reader is again referred to Munkres [11].

DEFINITION 6.4. Let X, Y be topological spaces. We define a topology on $X \times Y$ by taking as a basis the collection $\mathcal{B}$ of all sets of the form $U \times V$ where $U \in X, V \in Y$ and U, V are both open sets. This topology is called the **product topology** on $X \times Y$.

To define the product topology for an arbitrary number of sets we define the subbasis,

$$\mathcal{S}_\beta = \{\pi_\beta^{-1}(U_\beta) | U_\beta \in X_\beta, U_\beta \text{ is open}\}$$

and,

$$\mathcal{S} = \bigcup_{\beta \in J} S_\beta$$

The topology generated by the collection $\mathcal{S}_\beta$ is the **product topology** on $\prod_{\alpha \in J} X_\alpha$

We now define an open cover and compactness. We follow these with a proof of the tube lemma, which is helpful in showing that products of compact spaces are compact. After the tube lemma, we prove that the product of two compact spaces is always compact.

DEFINITION 6.5. We define an **open cover** of a set X to be a collection of open sets $\mathcal{O}$ such that $\bigcup \mathcal{O} = X$.

DEFINITION 6.6. A set X is **compact** if for every open cover $\mathcal{O}$ of X there exists a finite subcollection of sets of $\mathcal{O}$, say $O_1, O_2, ..., O_n$, that cover X.

LEMMA 6.7. (The tube lemma) Consider the product space $X \times Y$, where Y is compact. Let $x_0 \in X$. If N is an open set of $X \times Y$ that contains $x_0 \times Y$, then there exists an open set $W \subset X$ such that $x \in W$ and $W \times Y \subset N$.

Proof. For every point in $x_0 \times Y$, there exists a basis element of the form $U \times V \subset N$ containing x_0 (by the 2nd property of a basis). Selecting one such basis element for every point in $x_0 \times Y$ we have a collection that we shall call $\mathcal{O}$. Since Y is compact, $x_0 \times Y$ is also compact. This allows us to select from $\mathcal{O}$ a finite subcollection of basis elements, $U_1 \times V_1, U_2 \times V_2, ..., U_n \times V_n$ which covers $x_0 \times Y$. Now let $W = U_1 \cap ... \cap U_n$. Since each U_i is open, and a finite intersection of open sets is open, W is open. Since $x_0 \in U_i$ for every i, $x_0 \in W$. Now why is it true that $W \times Y \subset N$?

Let $x' \times y' \in W \times Y$. Consider the point $x_0 \times y' \in x_0 \times Y$. We know that $y' \in V_i$ for some i (by construction). Fix this i and V_i. We also know that $x' \in W$ which implies that $x' \in U_i$ for all i. Thus $x' \times y' \in U_i \times V_i \subseteq N$. So $W \times Y \subset N$. □

THEOREM 6.8. The product of two compact spaces is compact.

Proof. Let X,Y be compact spaces. Let $\mathcal{O}$ be an open covering of $X \times Y$. Let $x_0 \in X$. We can cover $x_0 \times Y$ using a finite number of sets from $\mathcal{O}$, call them $O_1,\ldots,O_n$. Using the tube lemma, we can find a W_{x_0} such that $x_0 \in W_{x_0}$, $W_{x_0} \times Y$ covers $x_0 \times Y$, and $W_{x_0} \times Y$ is covered by finitely many elements $(O_1,\ldots,O_n)$. For each $x \in X$ we can find a W_x in this same manner, so that the collection of all the $W_x \times Y$ is an open cover for $X \times Y$ and the collection of the W_x is an open cover for X. Since X is compact, there exists a finite collection $W_1,\ldots,W_k$ that covers X. We can conclude that $(W_1 \times Y) \cup (W_2 \times Y) \cup \ldots \cup (W_k \times Y)$ is a finite cover for $X \times Y$ each of which can be covered with a finite number of open sets from $\mathcal{O}$. Thus $X \times Y$ can be covered with a finite number of open sets from $\mathcal{O}$. □

Although the proof above did not require AC, when we generalize to an arbitrary product of compact spaces we must use a version of AC, Zorn's lemma. First, we prove a necessary characterization of compactness based on the definition of the finite intersection property.

DEFINITION 6.9. Let $\mathcal{C}$ be a collection of subsets of X such that every finite subcollection of $\mathcal{C}$ has a nonempty intersection. Then $\mathcal{C}$ is said to have the **finite intersection property**.

THEOREM 6.10. Let X be a topological space. Then X is compact if and only if for every collection $\mathcal{C}$ of closed sets in X that has the finite intersection property, the intersection of all the elements of $\mathcal{C}$ is nonempty.

Proof. Assume that X is compact and that $\mathcal{C}$ is a collection of closed sets in X with the finite intersection property.

Then $\mathcal{A} = \{X - C : C \in \mathcal{C}\}$ is a collection of open sets. We have that the finite subcollection of $\mathcal{A}$, $\{A_1,\ldots,A_n\}$ covers X if and only if for the corresponding $C_i = X - A_i$,

$$\bigcap_{i=1}^{n} C_i = \emptyset.$$

But $\mathcal{C}$ has the finite intersection property and so the latter is never true. Thus no finite subcollection of $\mathcal{A}$ covers X and since X is compact, $\mathcal{A}$ does not cover X.

So there must exist an element $y \in X - \bigcup \mathcal{A}$. That is to say, $y \in C$ for every $C \in \mathcal{C}$ and thus $\bigcap_{C \in \mathcal{C}} \neq \emptyset$.

The reverse direction is contained in the above proof. □

We are now ready to prove 3 lemmas that will be used in the proof that the Tychonoff theorem is equivalent to the Axiom of Choice.

LEMMA 6.11. If we have Zorn's lemma, then if X is a set and $\mathcal{A}$ is a collection of subsets of X having the finite intersection property, then there exists $\mathcal{D} \supset \mathcal{A}$ such that $\mathcal{D}$ has the finite intersection property and is maximal with respect to this property.

Proof. Let $\mathbb{A}$ denote the set consisting of all collections $\mathcal{B}$ that contain $\mathcal{A}$ and have the finite intersection property. Then $\mathbb{A}$ is partially ordered (strict) by proper inclusion and it remains to show that every chain in $\mathbb{A}$ has an upper bound in $\mathbb{A}$ (to use Zorn's lemma).

Let $\mathbb{B} \subset \mathbb{A}$ (i.e. $\mathbb{B}$ is a set consisting of collections of subsets of X) such that $\mathbb{B}$ is totally ordered. We have that $\mathcal{C} = \bigcup_{\mathcal{B} \in \mathbb{B}} \mathcal{B}$ is an upper bound for $\mathbb{B}$. It remains to show that $\mathcal{C} \in \mathbb{A}$. Since each $\mathcal{B} \in \mathbb{B}$ contains $\mathcal{A}$ we have that $\mathcal{C} \supset \mathcal{A}$. Now we show that $\mathcal{C}$ has the finite intersection property. Let $C_1, \ldots, C_n$ be elements of $\mathcal{C}$. Then each of the C_i must be contained in $\mathcal{B}_i$ for some i. So for each C_i find a $\mathcal{B}_i$ such that $C_i \in \mathcal{B}_i$ and form the finite collection $\{\mathcal{B}_1, ..., \mathcal{B}_n\} \subset \mathbb{B}$. Since this collection is in $\mathbb{B}$, it is totally ordered, and also being finite implies that it has a largest element, say $\mathcal{B}_k$. Being the largest element implies that $\mathcal{B}_i \subset \mathcal{B}_k$ for $i = 1, 2, ..., n$. Now we see that the sets $C_1, ..., C_n$ are all elements of $\mathcal{B}_k$. Since $\mathcal{B}_k \in \mathbb{B} \subset \mathbb{A}$, $\mathcal{B}_k$ has the finite intersection property. Thus the intersection of the sets $C_1, ..., C_n$ is nonempty. So we have showed that an arbitrary totally ordered subset of $\mathbb{A}$ has an upper bound and now apply Zorn's lemma to conclude that $\mathbb{A}$ has a maximal element, call it $\mathcal{D}$. That is to say,

(1) $\mathcal{D}$ is a collection of subsets of X such that $\mathcal{D} \supset \mathcal{A}$ and has the finite intersection property.

(2) $\mathcal{D}$ is maximal in $\mathbb{A}$ with respect to the finite intersection property. □

LEMMA 6.12. Let X be a set. If $\mathcal{D}$ is a collection of subsets of X that is maximal with respect to the finite intersection property, then any finite intersection of elements of $\mathcal{D}$ is not only nonempty, but also an element of $\mathcal{D}$.

Proof. Let B be the intersection of finitely many elements of $\mathcal{D}$ and assume that $B \notin \mathcal{D}$. B is nonempty because $\mathcal{D}$ has the finite intersection property. Define a larger collection of subsets of X denoted $\mathcal{E} = \mathcal{D} \cup B$. We want to prove that $\mathcal{E}$ has the finite intersection property. Let F be finitely many elements of $\mathcal{E}$. Case 1: If none of the elements of F is B then they must all belong to $\mathcal{D}$ and thus have a nonempty intersection. Case 2: If one of the elements of F is B we conclude that if there are any other elements in $\mathcal{E}$ then they must belong to $\mathcal{D}$ and we denote them $D_1, ..., D_m$ so that the intersection of F is of the form,

$$D_1 \cap D_2 \cap ... \cap D_m \cap B$$

which is nonempty because B is a finite intersection of elements of $\mathcal{D}$ and $D_1 \cap \ldots \cap D_n$ is also a finite intersection of elements of $\mathcal{D}$. So this intersection is nonempty which concludes the proof that $\mathcal{E}$ has the finite intersection property. This could not be the case though since $\mathcal{D}$ was maximal with respect to this property and thus assuming that $B \notin \mathcal{D}$ was our mistake. Thus it can only be true that $B \in \mathcal{D}$. □

LEMMA 6.13. Let $A \subset X$ such that A intersects every element of $\mathcal{D}$, where $\mathcal{D}$ has the properties discussed in Lemma 6.11. Then A is an element of $\mathcal{D}$.

Proof. Let $\mathcal{E} = \mathcal{D} \cup A$. Choose finitely many elements from $\mathcal{E}$. Case 1: None of these finitely many elements is A. Then we have finitely many elements of $\mathcal{D}$ and thus their intersection is nonempty since $\mathcal{D}$ has the finite intersection property. Case 2: One of the finitely many elements is A. So we have an intersection of the form $D_1 \cap \ldots \cap D_n \cap A$ where each $D_i \in \mathcal{D}$. We know that the intersection of the D_i's is an element of $\mathcal{D}$ from the lemma above, so A must intersect this intersection and it is therefore nonempty. □

The following is known as Tychonoff's theorem. In 1930 Tychonoff proved only that the product of arbitrary copies of the closed (and bounded, thus compact) interval $[0,1]$ was compact. The method of proof was later generalized though in 1937 by Eduard Cech [Moore], whom you may know from the Stone-Cech Compactification Theorem.

THEOREM 6.14. (Tychonoff) An arbitrary product of compact spaces is compact in the product topology. That is to say, if $X = \prod_{\alpha \in I} X_\alpha$ where each X_α is compact, then X is compact.

Proof. Let $\mathcal{A}$ be a collection of subsets of X that have the finite intersection property. We will show that $\bigcap_{A\in\mathcal{A}} \bar{A}$ is nonempty (indirectly) which will prove that X is compact by the alternate definition of compact. Using a previous lemma we can choose $\mathcal{D}$, a collection of subsets of X that contains $\mathcal{A}$ and is maximal with respect to the finite intersection property. Showing that $\bigcap_{D\in\mathcal{D}} \bar{D}$ is nonempty will prove that $\bigcap_{A\in\mathcal{A}} \bar{A}$ is nonempty.

For every $\alpha \in I$ let $\pi_\alpha : X \to X_\alpha$ be the projection mapping onto the α^{th} coordinate of an element of X. Now let $H = \{\pi_\alpha(D)|D \in \mathcal{D}\}$. Each $\pi_\alpha(D)$ is a subset of X_α. Since $\mathcal{D}$ has the finite intersection property, so does H, since intersecting elements of H is equivalent to intersecting elements of X that have zeroes on all coordinates except the α^{th} coordinate. Since each X_α is compact, we can choose a point x_α such that $x_\alpha \in \bigcap_{D\in\mathcal{D}} \pi_\alpha\bar{(D)}$. Define $\bar{x}$ to be the point $(x_\alpha)_{\alpha\in I}$ of X. If we show that $\bar{x} \in \bar{D}$ for every $D \in \mathcal{D}$ then we will have shown that $\bigcap_{D\in\mathcal{D}} \bar{D}$ is nonempty.

Let $\pi_\beta^{-1}(U_\beta)$ be a subbasis element such that $\bar{x} \in \pi_\beta^{-1}(U_\beta)$. This means that U_β is an open set in X_β that contains $\bar{x}$. We now have that (U_β) intersects the projection of D in $\pi_\beta(y)$ for some $y \in D$. So $y \in \pi_\beta^{-1}(U_\beta) \cap D$.

It follows from the lemma above that every subbasis element containing $\bar{x}$ is contained in $\mathcal{D}$, and thus every basis element containing $\bar{x}$ is in $\mathcal{D}$. Since $\mathcal{D}$ had the finite intersection property, it follows that every basis element containing $\bar{x}$ intersects every element of $\mathcal{D}$ and so $\bar{x}$ is a point of closure of D, i.e. $\bar{x} \in \bar{D}$ for all $D \in \mathcal{D}$. And so $\bigcap_{D\in\mathcal{D}} \bar{D}$ is nonempty, and X is compact. □

We will now show that Tychonoff's theorem implies the Cartesian Product theorem which is, as shown in section 3, equivalent to the axiom of choice.

THEOREM 6.15. The Tychonoff Theorem implies the Cartesian Product Theorem, and hence AC.

Proof. Let $\mathcal{B} = \{B_i | i \in I\}$ be a collection of nonempty subsets of X. Choose an element of X, call it p, such that $p \notin \bigcup_{i \in I} B_i$. Set $D_i = B_i \cup \{p\}$ for all $i \in I$. Give each D_i the topology $\tau_{D_i} = \{\emptyset, D_i, B_i, \{p\}\}$. Each D_i is compact since any open cover consists of 3 open sets or fewer and hence there exists a finite subcover. By the Tychonoff Theorem, $\prod_{i \in I} D_i = D$ is compact. Define for each $i \in I$, the set $U_i = \{f \in D : f(i) = p\}$. Each of these is open since $U_i = \pi_i^{-1}(\{p\})$ and projection mappings are continuous in the product topology (i.e. the inverse image of an open set is always open).

$\{U_i | i \in I\}$ is not an open cover of D for it were it would have a finite subcover since our product is compact, but we show that a finite subcover is not possible. Assume that $U_{i_1}, U_{i_2}, ..., U_{i_n}$ is a finite subcover of D. We define a function $(x_i)_{i \in I} \in D$ by,

$$x(k) = \begin{cases} b_{i_j} & : k = i_j \text{ for some } j \in \{1, 2, ..., n\} \\ p & : otherwise \end{cases}$$

This is a finite choice of elements which is allowed in ZF. We see that $x \in D$ but that $x \notin \bigcup_{j \in \{1,2,...,n\}} U_{i_j}$. This implies that $\{U_i | i \in I\}$ is not an open cover for D and so there exists a tuple $y \notin \bigcup_{i \in I} U_i$. But then the tuple y has no coordinate with value $\{p\}$ for if it did, then it would have been in the union of the U_i. Therefore $y \in \prod_{i \in I} B_i$. □

Functional Analysis

In this section we are interested in extending a linear functional that is defined only on a subspace of X to all of X. Instead of the trivial extension, we want an extension that preserves some of the properties of our functional. The result, which is an important result in functional analysis, is known as the Hahn-Banach theorem.

We define now a vector space, a linear functional and an extension of a linear functional. Then we prove the Hahn-Banach theorem using the Hausdorff maximal principle.

A **vector space** is an ordered tuple $(X, F, +, *)$ such that X is a set, F is a field, and $+$ and $*$ are binary operations and the following conditions are true for any $x, y, z \in X$ and $\alpha, \beta \in F$,

(1) $x + y = y + x$.

(2) $x + (y + z) = (x + y) + z$.

(3) There exists an element denoted 0 such that $x + 0 = x$.

DEFINITION 7.1. A **linear functional** is a real-valued function on X (a vector space) such that for all $\alpha, \beta \in \mathbb{R}$ and for all $x, y \in X$,

$$f(\alpha x+\beta y)=\alpha f(x)+\beta f(y).$$

DEFINITION 7.2. Given two linear functionals g_1, g_2, we say that g_2 is an **extension** of g_1 if

(1) $\text{dom}(g_1) \subseteq \text{dom}(g_2)$.

(2) For any $x \in \text{dom}(g_1), g_1(x) = g_2(x)$.

In the following theorem we give the sufficient conditions so that we can extend a linear functional from a subspace of X to all of X.

THEOREM 7.3. (Hahn-Banach)

Let p be a real-valued function defined on a vector space X such that,

(1)$p(x+y) \leq p(x)+p(y)$ for all $x, y \in X$ (subadditivity).

(2)$p(\alpha x) = \alpha p(x)$ for all $\alpha \in \mathbb{R}, \alpha \geq 0$.

Let f be a linear functional defined on a subspace $S \subseteq X$ such that,

(1) $f(s) \leq p(s)$ for all $s \in S$.

Then there exists an extension of f, a linear functional $F : X \to \mathbb{R}$ such that,

(1) $F(x) \leq p(x)$ for all $x \in X$.

(2) $F(s) = f(s)$ for all $s \in S$.

Proof. The proof will be in two parts.

Part I will demonstrate the existence of the extension F that satisfies the properties above.

Part II will show that F is defined for all $x \in X$.

Define H to be the set of all g such that g is a linear functional on a subspace of X with $g(x) \leq p(x)$ whenever $g(x)$ is defined.

We now partially order H by defining for any $g_1, g_2 \in H, g_1 \leq g_2$ iff g_2 is an extension of g_1.

Seeing that $\{f\} \subseteq H$, take f as a chain, then the Hausdorff Maximal Principle (AC) claims that there is a maximal chain $\{g_\alpha\}_{\alpha \in I}$ so that every $g_\alpha \in H$ and $f \subseteq \{g_\alpha\}_{\alpha \in I}$.

Now define F so that,

(1) $\mathrm{dom}(F) = \bigcup(\mathrm{dom}(g_\alpha))$ for all $\alpha \in I$

(2) $F(x) = g_\alpha(x)$ whenever $x \in \mathrm{dom}(g_\alpha)$.

The function F is well-defined since if $x \in \mathrm{dom}(g_\alpha)$ and $x \in \mathrm{dom}(g_\beta)$ then we must have $g_\alpha(x) = g_\beta(x)$ since both functions are in the chain and one of them is the greater with respect to the extension.

Claim: F is a functional.

(1) Let x, y be in the domain of F. Then $x \in \mathrm{dom}(g_\alpha)$ for some α and $y \in \mathrm{dom}(g_\beta)$ for some β, where $\alpha, \beta \in I$.

WLOG, say that $g_\alpha \leq g_\beta$. Then $x, y \in \text{dom}(g_\beta)$. Thus for any $\lambda, \mu \in \mathbb{R}$ we have that,

$$F(\lambda x + \mu y) = g_\beta(\lambda x + \mu y) = \lambda F(x) + \mu F(y).$$

So F is a functional.

(2) Claim: F is a maximal extension of f (i.e. there is no proper extension of f).

Let G be an extension of F. But then for any α, $g_\alpha \leq F \leq G$ and thus G is an extension of g_α and hence an element of our chain. But then we have $G \leq F$, and so $G = F$.

To show that F is defined on all of X we will prove the following claim.

Claim: For each g which is defined on a proper subspace $T \subsetneq X$ and satisfying $g(t) \leq p(t)$ for all $t \in T$, there is a proper extension of g, denoted h.

*So if F were not defined on all of X then by the claim it would have a proper extension h, but we just showed that this cannot be.

Let T be a proper subspace of X and let y be an element of $X - T$. Let g be as described above. We extend g to the subspace $U = \{\lambda y + t : t \in T\}$.

Then for any $t_1, t_2 \in T$ we see that,

$$g(t_1) + g(t_2) = g(t_1 + t_2) \leq p(t_1 + t_2) \leq p(t_1 - y) + p(t_2 + y),$$

and so we have,

$-p(t_1 - y) + g(t_1) \leq p(t_2 + y) - g(t_2),$

which finally gives,

$$\sup_{t \in T}[-p(t-y) + g(t)] \leq \inf_{t \in T}[p(t+y) - g(t)].$$

We then define $h(y) = \alpha$ where α is any real number such that,

$$\sup_{t \in T}[-p(t-y) + g(t)] \leq \alpha \leq \inf_{t \in T}[p(t+y) - g(t)].$$

And so we have that for any $\lambda \in \mathbb{R}$ and $t \in T$, define,

$h(\lambda y + t) = \lambda h(y) + g(t).$

Hence h is a linear functional by definition and an extension of g. It remains to show that $h(a) \leq p(a)$ for all a in the domain of h.

Case 1: $a = \lambda y + t$ with $\lambda > 0$.

Then we have that,

$h(\lambda y + t)$

$= \lambda h(y) + h(t)$

$= \lambda \alpha + g(t)$

$= \lambda(\alpha + g(t/\lambda))$

$\leq \lambda((p(t/\lambda + y) - g(t/\lambda) + g(t/\lambda)))$

$= \lambda p(t/\lambda + y)$

$= p(t + \lambda y)$

$= p(\lambda y + t))$.

Case 2: $a = \lambda y + t$ with $\lambda < 0$.

Then let $\lambda = -\mu$ for some $\mu > 0$. We have that,

$h(\lambda y + t)$

$= \lambda h(y) + h(t)$

$= \lambda \alpha + g(t)$

$= -\mu \alpha + g(t)$

$= \mu(-\alpha + g(t/\mu))$

$\leq \mu((p(t/\mu - y) - g(t/\mu) + g(t/\mu)))$

$= \mu p(t/\mu - y)$

$= p(t - \mu y)$

$= p(t + \lambda y)$

$= p(\lambda y + t))$.

And thus h is a proper extension of g satisfying $h(a) \leq p(a)$ for all $a \in \text{dom}(h)$. Our claim is proved.

Hence F must be defined for every $x \in X$, otherwise it would have a proper extension which contradicts its maximality. □

Bibliography

Bibliography

[1] B. Banaschewski, *A new proof that "'Krull implies Zorn"',*Math. Logic Quart. **40** (1994), 478-480.

[2] Cohen, Paul. "The Independence of the Continuum Hypothesis". Proceedings of the National Academy of Sciences of the United States of America, Vol. 50, No. 6. (Dec. 15, 1963), pp. 1143-1148.

[3] Cohen, Paul. "The Independence of the Continuum Hypothesis, II" Proceedings of the National Academy of Sciences of the United States of America, Vol. 51, No. 1. (Jan. 15, 1964), pp. 105-110.

[4] Dummit, D and Foote, R. Abstract Algebra, third edition. John Wiley and Sons, Inc. 2004.

[5] Godel, Kurt. "ÃIJber formal unentscheidbare SÃd'tze der Principia Mathematica und verwandter Systeme, I." Monatshefte fÃijr Mathematik und Physik 38: 173âĂŞ98. 1931.

[6] W. Hodges, *Krull implies Zorn,* J. London Math. Soc. **19** (1979), 285-287.

[7] Jech, Thomas J. The Axiom of Choice. North-Holland Publishing Company, Amsterdam. London. 1973.

[8] Kaplansky, Irving. American Mathematical Society; 2nd edition 2001.

[9] Kenneth Kunen *Set Theory: An Introduction to Independence Proofs.*Elsevier, 1980. Random House, N.Y.

[10] Moore, Gregory H. "'Zermelo's axiom of choice." Studies in the History of Mathematics and Physical Sciences 8. Springer-Verlag. New York. 1982.

[11] James Munkres *Topology.*Second Edition. Prentice Hall, 1975.

[12] Potter, Michael. Set Theory and Its Philosophy. Oxford University Press. New York. 2004.

[13] H. L. Royden *Real Analysis* Second Edition. The Macmillan Company, New York. 1968.

[14] Rubin, H. and Rubin J.E. Equivalents of the Axiom of Choice, II. North-Holland Publishing Company, Amsterdam. 1985.

[15] Shoenfield, Joseph. "Axioms of set theory", in Barwise 1977, pg 321-44.

[16] Shoenfield, Joseph. Mathematical Logic. AK Peters, Ltd. 2001.

Appendix A

Cyclic Lists of Equivalencies of AC

These lists provide a reference to the pages in which equivalences of AC are proved.

List 1: The equivalence of AC, the well-ordering theorem, and Zorn's Lemma.

Page 26, AC implies Zorn's modified lemma.

Page 27, Zorn's modified lemma implies the Hausdorff maximal principle.

Page 20, Hausdorff maximal principle implies Zorn's lemma.

Page 17, Zorn's lemma implies the well-ordering theorem.

Page 15, The well-ordering theorem implies AC.

List 2: The equivalence of AC with the Kuratowski lemma, Tukey's lemma, and the Hausdorff maximal principle.

Page 26, AC implies Zorn's modified lemma.

Page 27, Zorn's modified lemma implies the Hausdorff maximal principle.

Page 20, Hausdorff maximal principle implies Zorn's lemma.

Page 28, Zorn's lemma implies Kuratowski's lemma.

Page 28, Kuratowski's lemma implies Zorn's lemma.

Page 18, Zorn's lemma implies Tukey's lemma.

Page 19, Tukey's lemma implies Hausdorff maximal principle.

Page 20, Hausdorff maximal principle implies Zorn's lemma.

Page 17, Zorn's lemma implies well-ordering theorem.

Page 15, Well-ordering theorem implies AC.

List 3: The equivalence of AC with Tychonoff's theorem.

Page 26, AC implies Zorn's modified lemma.

Page 27, Zorn's modified lemma implies the Hausdorff maximal principle.

Page 20, Hausdorff maximal principle implies Zorn's lemma.

Page 38, Zorn's lemma implies Lemma 6.11.

Page 40, Lemma 6.11 implies Tychonoff's theorem.

Page 42, Tychonoff's theorem implies the Cartesian product theorem.

Page 12, The Cartesian product theorem implies AC.

Appendix B

Important Results of AC

In this appendix we give the page numbers of results of AC which are not equivalent (inferior) or which are equivalent but whose second direction is only referenced and not proved in this thesis.

Page 32, AC implies Krull's theorem.

Page 44, Hausdorff maximal principle (AC) implies the Hahn-Banach theorem.

Vita

Cristian Gerardo Allen was born on February 14, 1983 in Alajuela, Costa Rica. After moving to Richmond, Virginia at an early age with his mother he attended public schools until his graduation from James River High School in 2001. He received his Bachelor of Arts in Mathematics from Virginia Commonwealth University in 2008 and subsequently taught for 1 year at Varina High School in Henrico County, Virginia. He then gained acceptance to the Master of Science of Mathematics program at Virginia Commonwealth University and will graduate in 2010.