# The Continuum Hypothesis and Forcing

Connor Lockhart

December 2018

**Abstract**

In this paper we introduce the problem of the continuum hypothesis and its solution via Cohen forcing. First, we introduce the basics of first order logic and standard ZFC set theory before elaborating on ordinals, cardinals and the forcing concept. The motivation of this paper is expository and should be approachable for anyone familiar with first order logic and set theory.

## Contents

# 1 Introduction and the Continuum Hypothesis

The continuum hypothesis (also referred to as CH) was first formulated in 1878 by Georg Cantor following his work on the foundations of set theory. Its formulation is often stated as **There is no set whose cardinality is strictly between that of the integers and the real numbers.** This can also be reformulated to state that the successor cardinal to $\aleph_0$ is the cardinality of the reals. Such was suspected, but not proven, by Cantor and his contemporaries. The first major advance on the problem was presented by Gödel in 1940 showing its consistency with ZFC axioms, and independence was finally shown in 1963 by Cohen. The technique Cohen created for showing independence, known as "forcing," was highly successful and elaborated by many contemporaries. The method that is used in this paper is guided by Cohen's original proof.

# 2 First Order Logic

Before proving any results of the continuum hypothesis, the reader must be comfortable with ideas and theorems from first order logic. This section can serve as a refresher or a hasty introduction to the ideas.

## 2.1 Formal Languages

The mathematical theorems we use in our day-to-day mathematics are usually not written in logical quantifiers because of their complexity but it is important to understand the principles behind formal languages.

**Definition 2.1** (The Symbols of Formal Language)**.**

$$\text{The 5 propositional connectives } \begin{pmatrix} \sim & \& \text{ or } \wedge & \vee & \rightarrow & \leftrightarrow \\ \text{not} & \text{and} & \text{or} & \text{implies} & \text{If and only If} \end{pmatrix}$$

$$\text{The universal and existential quantifiers } \begin{pmatrix} \forall & \exists \\ \text{For all} & \text{There exists} \end{pmatrix}$$

Variables $x_1, x_2, x_3...$ are usually denoted $x, x', x'', x''', ...$ to preserve symbol count but the distinction is not very important.

Propositional connectives serve as connectives between statements and let us build coherent logical phrases. These symbols are not the most efficient, since we can express some symbols using combinations of others. $\forall$, for example, is equivalent to $\sim \exists \sim$. Constants will usually be denoted in the same manner as variables, but with a c instead of an x.

**Example 2.1** (Uniqueness of Addition)**.** If we wish to represent addition in a formal expression we can think of addition as some relation between triples of numbers, namely $r + s = t$. If we have a formal language where $R_1$ is a ternary relation symbol then letting $R_1$ be addition we can represent the uniqueness of addition by

$$\forall x, \forall y, \forall z \forall u \ ((R_1(x, y, z) \& R(x, y, u)) \rightarrow z = u$$

**Example 2.2** (Existence of additive identity)**.** Here is the much easier formal representation of the existence of the additive identity by

$$\forall x \exists y \ (R_1(x, y, y))$$

## 2.2 Model Theory

Here I will introduce some terms and ideas that will allow us to expand our formal language to be able to describe fully-fledged mathematical objects, and to determine whether some idea or theorem is true not just in general but within a certain system or model. The system begins with an alphabet and it defined inductively from there.

**Definition 2.2. Alphabet**
The alphabet of a first order logic is a set containing:

- An infinite list of constant symbols: $Such\ as\ a, a', a'', a'' \cdots$
- An infinite list of variable symbols: $Such\ as\ x, x', x'', x''' \cdots$
- An infinite list of function symbols: $Such\ as\ f, g, h \cdots$
- An infinite list of relation symbols: $Such\ as\ P, Q, R \cdots$
- Logical connectors and quantifiers
- An equality sign and parentheses

The meaning and notation of the logical connectors and quantifiers is the same as in the previous section, as well as the equal sign and parentheses taking on their usual meanings. Now that we have an underlying set, we can define additional structures on this set.

**Note 1.** A function or relation of arity n takes n inputs. Therefore, addition in the usual sense is a 2 arity function since it take 2 separate numbers and returns a third, the sum. Addition as a relation has 3 arity, since it takes two numbers and a sum and is a true relation if those the numbers create a valid sum (and false if they do not).

**Definition 2.3. Term**
A term is a string of symbols from the alphabet defined recursively from a set of rules to help define which strings have meaning and can be used later on.

- Every constant and variable symbol is a term.
- If a function $f$ has arity n and $t_1, t_2, t_3, ..., t_n$ are terms then $f(t_1, t_2, t_3, ..., t_n)$ is a term
- A string of symbols is a term if it can be constructed by applying the above steps finitely many times.

**Definition 2.4. Formula**
Similar but more complex than a term, a formula is a string of symbols defined recursively as follows.

- If $t_1$ and $t_2$ are terms then $(t_1 = t_2)$ is a formula
- If R is a relation with arity $n$ and $t_1, t_2, ..., t_n$ are terms then $R(t_1, t_2, ..., t_n)$ is a term
- If $\phi$ is a formula then so is $\neg\phi$
- If $\phi$ and $\psi$ are formulas then so is $(\phi \wedge \psi)$
- If $\phi$ is a formula and x is variable then so is $(\exists x)\phi$
- A string of symbols is a formula if it can be constructed by application of the above rules in finitely many steps.

If terms are the basic building blocks of logical statements, then formulas are the things that can be made from those building blocks in a well-constructed and standardized way.

**Note 2.** Formulas of the form given by the first two rules of definition 2.4 are known as **Atomic Formulas**. They are the formulas that are most basic and build up other formulas.

Now we have to narrow down the types of variables in formulas as either free or bound variables. To not get confused by excessive notation, free variables are ones without a quantifier ($\exists$, $\forall$) and bound variables are with a quantifier.

**Definition 2.5. Sentence**
A formula $\phi$ is called a sentence if it has no free variables.

Now that we have our notation established for model theory we are going to move on to the actual "modeling" aspect.

**Definition 2.6. Language**
A language $\mathcal{L}$ is a set containing all logical symbols and quantifiers and an arbitrary number of constants, variables, functions, symbols and relation symbols.

A language is different from an alphabet as it has the functions and relations used to model a certain thing, but not necessarily all of them. The language for modeling the real numbers might have much different relations and function symbols than the language used for modeling addition.

**Definition 2.7. Model**
A model $\mathfrak{A}$ for a language $\mathcal{L}$ is an ordered pair $< \mathbf{A}, \mathcal{I} >$ where $\mathbf{A}$ is a nonempty set and $\mathcal{I}$ is an interpretation function with domain of the set of all constant, function and relation symbols of $\mathcal{L}$ such that:

1. If c is a constant symbol, then $\mathcal{I}(c) \in \mathbf{A}$; $\mathcal{I}(c)$ is called a constant

2. If $F$ is a m arity function symbol, then $\mathcal{I}(F)$ is an m arity function on $\mathbf{A}$.

3. If R is an n airty relation symbol, then $\mathcal{I}(R)$ is an n airty relation on $\mathbf{A}$.

$\mathbf{A}$ is called the universe of the model $\mathfrak{A}$.

The importance of model theory comes from the fact that mathematical objects can be cast as models for a language. For example, the real numbers with the usual ordering under $<$, the standard arithmetic operations $(\times +)$ and the special numbers 0 and 1 can together be described as a model. Let $\mathcal{L}$ contain one 2 placed relation symbol $R_0(<)$, two 2 placed function symbols $F_1, F_2(\times, +)$, and two constant symbols $c_1, c_2(0, 1)$. We build the model by having $\mathbf{A}$ be the set of real numbers. The interpretation function maps each relation function and constant to its standard one in parentheses. So finally, $< A, \mathcal{I} >$ is an example of a model described by the language $\{R_0, F_1, F_2, c_0, c_1\}$.

**Definition 2.8. Valuation**
The value $t[x_0, ..., x_q]$ of a term $t(v_0, ..., v_q)$ in the universe $\mathbf{A}$ of a model $\mathfrak{A}$ is defined as follows:

1. If $t$ is $v_i$ then $t[x_0, ..., x_q]$ is $x_i$

2. If $t$ is the constant symbol c, then $t[x_0, ..., x_q]$ is $\mathcal{I}(c)$, the interpretation of c in $\mathbf{A}$

3. if t is $F(t_1, ..., t_m)$ where $F$ is an m placed function symbol and $t_1, ..., t_m$ are terms, then $t[x_0, ..., x_q]$ is $G(t_1[x_0, ..., x_q], ..., t_m[x_0, ..., x_q])$ where G is the m placed function $\mathcal{I}(F)$, the interpretation of $F$ in $\mathbf{A}$.

Now after defining valuation we define how formulas are satisfied by models

**Definition 2.9. Satisfaction** Suppose $\mathfrak{A}$ is a model for a language $\mathcal{L}$. The sequence $x_0, ..., x_q$ of elements of $\mathbf{A}$ satisfies the formula $\psi(v_0, ..., v_q)$ all of whose free and bound variables aree among $v_0, ..., v_q$ in the model $\mathfrak{A}$, written $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ provided we have:

1. if $\psi(v_0, ..., v_q)$ is the formula $(t_1 = t_2)$, then $\mathfrak{A} \vDash (t_1 = t_2)[x_0, ...x_q]$ means that $t_1[x_0, ..., x_q]$ equals $t_2[x_0, ..., x_q]$

2. if $\psi(v_0, ...v_q)$ is the formula $(R(t_1, ..., t_n))$ where R is a n placed relation symbol, then $\mathfrak{A} \vDash (R(t_1, ..., t_n))[x_0, ..., x_q]$ means $S(t_1[x_0, ..., x_q], ..., t_n[x_0, ..., x_q])$ where S is the n placed relation $\mathcal{I}(R)$, is the interpretation of R in $\mathbf{A}$

3. if $\psi$ is $\neg\theta$ then, $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$means not $\mathfrak{A} \vDash \theta[x_0, ..., x_q]$

4. if $\psi$ is $(\theta \wedge \phi)$, then $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ means both $\mathfrak{A} \vDash \theta[x_0, ..., x_q]$ and $\mathfrak{A} \vDash \phi[x_0, ..., x_q]$

5. if $\psi$ is $(\theta \vee \phi)$ then $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ means either $\mathfrak{A} \vDash \theta[x_0, ..., x_q]$ or $\mathfrak{A} \vDash \phi[x_0, ..., x_q]$

6. if $\psi$ is $\theta \to \phi$ then $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ means $\mathfrak{A} \vDash \theta[x_0, ..., x_q]$ implies $\mathfrak{A} \vDash \phi[x_0, ..., x_q]$
7. if $\psi$ is $\theta \leftrightarrow \phi$ then $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ means $\mathfrak{A} \vDash \theta[x_0, ..., x_q]$ if and only if $\mathfrak{A} \vDash \phi[x_0, ..., x_q]$
8. if $\psi$ is $\forall v_i \theta$ then $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ means for every $x \in \mathbf{A}$, $\mathfrak{A} \vDash \theta[x_0, ..., x_{i-1}, x_i, x_{i+1}, ..., x_q]$
9. if $\psi$ is $\exists v_i \theta$ then $\mathfrak{A} \vDash \psi[x_0, ..., x_q]$ means for some $x \in \mathbf{A}$, $\mathfrak{A} \vDash \theta[x_0, ..., x_{i-1}, x_i, x_{i+1}, ..., x_q]$

Satisfaction and valuation are what define truth in model theory, a logically valid sentence is true in every language, these are things such as tautologies. One such example is the law of contraposition:

$$(A \to B) \Leftrightarrow (\neg B \to \neg A)$$

however a sentence is logically satisfiable if it can be made true by some model. The process of satisfaction can be thought of as if a sentences can be built up from base sentences in a model in a sequential fashion to show that the original is true. Consistent sentences and theories are ones that no contradiction can be logically deduced, no sentence can have both itself and its negation be satisfied.

# 3 Set Theory and the ZFC axioms

As one does mathematics, rarely are all principles derived directly from axioms. To prove any reasonably complex statement with only axioms would be immensely tedious. Thus, it is not critically important for every mathematician to be familiar with the exact statements of common axioms. However, since many of the results proved about the CH are directly related to the axioms of set theory, we will list them and give some explanation for the meaning of each one as well as the associated statement in first order logic.

## 3.1 List and Motivation of the Zermelo-Fraenkel Axioms

**Definition 3.1.** The Actual Axioms of Zermolo-Fraenkel

1. **Axiom of Extensionality**: If X and Y have the same elements, then X=Y.

   $$\forall u(u \in X \leftrightarrow u \in Y) \to X = Y$$

   The motivation for the axiom of extentionality is to define sets strictly by their elements. Additionally the converse statement if $X = Y$ then $u \in X \leftrightarrow u \in Y$is an axiom of predicate calculus and thus creates an if and only if statement about what makes a set.

2. **Axiom of Pairing**: For any a and b there exists a set $\{a, b\}$ that contains exactly a and b.

   $$\forall a \forall b \exists c \forall x(x \in c \leftrightarrow x = a \vee x = b)$$

We define $\{a, b\}$ as the unique c that follows the above property. Sets with duplicate elements such as $\{a, a\}$ are equal to just $\{a\}$. Order does not matter in a set so we can define the ordered pair $(a, b)$ to satisfy the condition that $(a, b) = (c, d)$ if and only if $a = c, b = d$. With this set theoretic definition we can show the ordered pair to be equal to $(a, b) = \{a, \{a, b\}\}$

3. **The Axiom Scheme of Separation or Axiom of Specification**: If P is a property (with parameter p), then for any X and p there exists a set $Y = \{u \in X : P(u, p)\}$ that contains all those $u \in X$ that has property P.

$$\forall X \forall p \exists Y \forall u (u \in Y \leftrightarrow u \in X \wedge \varphi(u, p))$$

Note that the statement here is not that for some property, there exists a set with that property. Instead, it states that for elements that are in sets that fulfill a property, we can specify a set that is made of those elements. This prevents things like Russell's Paradox, since you cannot declare things that might not exist like "the set of all sets that don't contain themselves."

4. **The Axiom of Union**: For any X there exists a set $Y = \bigcup X$, the union of all elements of X

$$\forall X \exists Y \forall u (u \in U \leftrightarrow \exists z (z \in X \wedge u \in z))$$

There is not a axiom of intersection due to the fact that we can use the axiom of separation to narrow down a set A to $A \cap B$. The existence of "building" up a set of arbitrary unions cannot be proven from the other axioms such as power sets so it requires its own axiom.

5. **The Axiom of Power Set**: For any X, there exists a set $Y = P(X)$ the set of all subsets of X.

$$\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$$

This axiom states that for each set x, there exists another set y that contains all subsets of x. This axiom is not covered by the axiom of replacement, since we cannot define the power set as the range of a function.

6. **The Axiom of Infinity**: There exists an infinite set, or there exists a set containing 0 and containing the successor of each of its elements.

$$\exists x (\emptyset \in x \& \forall y (y \in x \rightarrow \cup y \in x)).$$

This is a very powerful axiom since it asserts the existence of a set by its property, namely that there is a set that it does not contain a largest element. The set in question specified by the infinity axiom is called the inductive set because it is the set upon which induction depends.

7. **Axiom Schema of Replacement**: If a class F is a function then for any X there exists a set $Y = F(X) = \{F(X) : x \in X\}$

$$\forall t_1, ...t_k(\forall x \exists! y \ A_n(x, y; t_1, ..., t_k) \rightarrow \forall u \exists v \ B(u, v)) \ where$$

$$B(u, v) = \forall r \ (r \in v \leftrightarrow \exists s(s \in u \& A_n(s, r; t_1, ..., t_k)))$$

The replacement schema states that for any function that can be written in a first order logic if we have some set in the domain of the function then we can take the image of that function and it is also a set. The importance of this is that given a set, even if we have a crazy function that itself is not a proper set, the image of a proper set is still a set.

8. **The Axiom of Regularity**: Every non-empty set contains an element that is disjoint from itself. (This prevents things such as Russel's Paradox)

$$\forall x \exists y \ (x = \emptyset \vee (y \in x \& \forall z(z \in x \implies \neg z \in y)))$$

This axiom is somewhat artificial even as far as axioms go. It states that each set has a minimal element with respect to element inclusion ($\in$). In a more intuitive sense, this means that we cannot have infinite descending chains with respect to inclusion so that each of our sets is "built up" from the empty set.

9. **The Axiom of Choice**: Every family of non-empty sets has a choice function or, more confusingly, the Cartesian product of a collection of non-empty sets is non-empty.

$$\text{If } \alpha \rightarrow A_\alpha \neq \emptyset \text{ is a function defined for all } \alpha \in x,$$

$$\text{then there exists another function } f(\alpha) \text{ for } \alpha \in x, \text{ and } f(\alpha) \in A_\alpha$$

The most famous of all the ZFC Axioms, the Axiom of Choice is also the most controversial. Some mathematicians do not consider it an intuitive part of ZFC and many theorems are often considered with and without the axiom. The axiom itself allows us to make an arbitrary number of "choice functions" when specifying things and properties. One intuitive way to imagine the axiom at work is take an infinite number of socks, normally indistinguishable. We want to take all of the pairs of socks and specify one of them to be the "left" sock and the other to be the "right" sock. Normally this is not able to be done without using some property of the sock to allow it to be distinct. The axiom of choice allows us, as if by fiat, to know that a choice function exists. With the Axiom of choice we no longer have to manually create a property to specify what our function would depend on.

These axioms form the basis of set theory. Each one has a meaning and creates restrictions. Some axioms have stronger or weaker versions that create new implications.

# 4 Ordinals and Cardinals

The continuum hypothesis is very closely tied to idea of infinity so it is important that we make rigorous our claims about how we interact with it. Ordinals are how we can create an order-motivated definition of infinity and cardinals are how we create a size-motivated definition of infinity.

## 4.1 Orderings and Ordinal numbers

**Definition 4.1. Partial Ordering** A binary relation $<$ on a set $P$ is a partial ordering of $P$ if:

- $p \not< p$ for any $p \in P$
- if $p < q$ and $q < r$ then $p < r$

$(P, <)$ is a called a partially ordered set. A partial ordering $<$ of P is called a linear ordering if additionally

- $p < q$ or $p = q$ or $p > q$ for all $p, q \in P$

The difference between a linear and partially ordered set is that in a linear ordered set every element is comparable while this might not be true in a partially ordered set.

**Definition 4.2. Well Ordering** A linear ordering of a set P is a well ordering if every non-empty subset of P has a least element.

Now we move to define ordinal numbers. Our idea is that

$$\alpha < \beta \text{ if and only if } \alpha \in \beta, \quad \text{and} \quad \alpha = \{\beta : \beta < \alpha\}$$

**Definition 4.3. Transitive set** A set T is transitive if every element of T is subset of T. (This is equivalent to $\cup T \subset T, or T \subset P(T)$)

**Definition 4.4. Ordinal Number** A set is an ordinal number if it is transitive and well ordered by $\in$. Ordinals are denoted as lower case Greek letters. For the successor to an ordinal, we define $\alpha + 1 = \alpha \cup \{\alpha\}$

**Lemma 4.1.** *If $(W, <)$ is a well-ordered set and $f : W \to W$ is an increasing function, then $f(x) \geq x$ for each $x \in W$*

*Proof.* Assume the set $X = \{x \in W : f(x) < x\}$ is non-empty and let x be the least element of X. If $\omega = f(z)$, then $f(\omega) < \omega$, a contradiction. □

**Lemma 4.2.** *No well ordered set is isomorphic to an initial segment of itself.*

*Proof.* If $range(f) = \{x : x < u\}$, then $f(u) < u$, contrary to lemma 4.1 □

Now we can prove a theorem relating well-ordered sets to ordinals.

**Theorem 4.3.** *Every well-ordered set is isomorphic to a unique ordinal*

*Proof.* Uniqueness follows from lemma 4.2. Given a well ordered set W we find an isomorphic ordinal as follows: Define $F(x) = \alpha$ if $\alpha$ is isomorphic to an initial segment of W given by x. If such an $\alpha$ exists then it is unique. By the replacement axioms, F(W) is a set. For each $x \in W$ such an $\alpha$ exists (if not consider the least x for which such an $\alpha$ does not exist). If $\gamma$ is the least $\gamma \notin F(W)$, then $F(W) = \gamma$ and we have an isomorphism of $W$ onto $\gamma$ □

**Definition 4.5. Limit Ordinals** If $\alpha = \beta + 1$ then $\alpha$ is a successor ordinal. If $\alpha$ is not a successor ordinal, then $\alpha = \sup\{\beta : \beta < \alpha\} = \cup\alpha$ and $\alpha$ is called a limit ordinal. 0 is also a limit ordinal since $0 = \sup \emptyset$.

The least nonzero limit ordinal is denoted as $\omega$ (or N). The ordinals less than $\omega$ are the natural numbers or finite ordinals.

## 4.2 Ordinal Arithmetic

Here we will define addition, multiplication and exponentiation of ordinal numbers.

**Definition 4.6. Addition** For all ordinal numbers $\alpha$

- $\alpha + 0 = \alpha$
- $\alpha + (\beta + 1) = (\alpha + \beta) + 1$, This applies only to successor ordinals, limit ordinals are defined by the third property only
- $\alpha + \beta = \lim_{\xi \to \beta}(\alpha + \xi)$ for all limit $\beta > 0$

**Definition 4.7. Multiplication** For all ordinal numbers $\alpha$

- $\alpha * 0 = \alpha$
- $\alpha * (\beta + 1) = \alpha * \beta + \alpha$
- $\alpha * \beta = \lim_{\xi \to \beta}(\alpha * \xi)$ for all limit $\beta > 0$

Keep in mind that addition and multiplication of ordinals are both associative but not commutative. For example:

$$1 + \omega = \omega \neq \omega + 1, \quad 2 * \omega = \omega \neq \omega * 2 = \omega + \omega$$

**Definition 4.8. Exponentiation** For all ordinal numbers $\alpha$

- $\alpha^0 = 1$
- $\alpha^{\beta+1} = \alpha^\beta * \alpha$
- $\alpha^\beta = \lim_{\xi \to \beta}(\alpha^\xi)$ for all limit $\beta > 0$

### 4.3 Cardinal Numbers

Two sets have the same cardinality $|X| = |Y|$ if there exists a one-to-one mapping of $X$ onto $Y$. The cardinality operation is an equivalence relation so we can assign to each set X its cardinal number $|X|$ to represent its cardinality or size. When denoting arbitrary cardinals, we generally use the letters $\kappa$ and $\lambda$

**Definition 4.9. Arithmetic on cardinals.** let $|A| = \kappa, |B| = \lambda$ Arithmetic on cardinals is defined as

- $\kappa + \lambda = |A \cup B|$ Where A and B are disjoint.
- $\kappa * \lambda = |A \times B|$
- $\kappa^{\lambda} = |A^B|$

**Theorem 4.4. Cardinality of Power sets:** *For every set X,* $|X| < |P(X)|$

*Proof.* For sake of contradiction let $f$ be a function from X onto P(X). The set

$$Y = \{x \in X : x \notin f(x)\}$$

is clearly not in the range of f. if there existed a $z \in X$ such that $f(z) = Y$, then $z \in Y$ if and only if $z \notin Y$, this is a contradiction. Thus, $f$ is not a function of X onto $P(X)$ and $|P(X)| \neq |X|$. □

**Definition 4.10. Alephs** An ordinal $\alpha$ is called a cardinal number (or just cardinal) if $|\alpha| \neq |\beta|$ for $\beta < \alpha$. If W is a well-ordered set, then there exists an ordinal $\alpha$ such that $|W| = |\alpha|$. Thus we can take

$$|W| = \text{ the least ordinal such that } |W| = |\alpha|$$

$|W|$ is a cardinal number. Take note that all infinite cardinals are limit ordinals. Cardinals are denoted by $\aleph_\alpha$ where $\alpha$ is an ordinal.

Sets whose cardinally is $\aleph_0$ are called "countable." Infinite sets that are not countable are called "uncountable."

**Definition 4.11. Cofinality** Let A be a set and let $\leq$ be a binary relation on A. Then a subset B of A is said to be cofinal if for every $a \in A$, there exists some $b \in B$ such that $a \leq b$.

The cofinality (cf(A)) of a partially ordered set A is the least of the cardinalities of all the cofinal subsets of A.

## 5 Forcing

Forcing is a difficult concept to work with for mathematical and metamathematical reasons. The mathematical difficulties come from using concepts like partial orders, dense sets and filters in strange ways. The more abstract and metamathematical difficulties come from any result about proving consistency

of a system. Any method to prove consistency of a theory using the same theory is doomed to fail from Gödels incompleteness theorem. We will state it here without proof for reference.

**Theorem 5.1. Gödels incompleteness theorem** Assume F is a consistent formalized system which contains elementary arithmetic. Then F cannot prove the consistency of F.

Given the additional difficulty of proving the consistency of ZFC, like many consistency proofs the strategy for proving the independence of $CH$ is to show that given the consistency of ZFC that both $CH$ and $\neg CH$ are also consistent. The first part of showing $Cons(ZFC) \rightarrow Cons(CH)$ was done by Gödel in 1940. Now we use forcing to show $Cons(ZFC) \rightarrow Cons(\neg CH)$. Gödels original method involved using the idea of constructable sets to create a model for ZFC where $CH$ was true. One of the main ideas of forcing is not to create a model exactly as Gödel had done, but to take an existing model of ZFC and adjoin a larger set to it that allows new statements to be true, but still models the same object as our original model.

## 5.1 Tools of Forcing

Before we sketch out our final proof we need some definitions that will be used in forcing.

**Definition 5.1. Constructable Sets** Let X be a set. The set X' is defined as the union of X and the set of all sets Y for which there is a formula $A(z, t_1, ..., t_k)$ in ZF such that if $A_X$ denotes A with all bound variables restricted to X, then for some $t_1$ in X,

$$y = \{z \in X | A_X(z, t_1, ..., t_k)\}$$

Now we can define on $\alpha$, $M_\alpha$ to be $M_0 = \emptyset$ and $M_\alpha = (\bigcup_{\beta<\alpha} M_\beta)'$. We can think of $M_\alpha$ as all of the constructable set that are "built up" in $\alpha$ steps. A set x is constructable if $\exists\alpha$, on $\alpha$ and $x \in M_\alpha$

Constructable sets are important since they are used to define our standard model of ZF that we can play around with to try and create the properties that we want.

**Definition 5.2. Labeling and Label Space** A labeling is a mapping defined in ZF which assigns to each ordinal $0 < \alpha < \alpha_0$, a set $S_\alpha$, the "label space". Functions $\psi_\alpha$ defined in $S_\alpha$ such that the sets $S_\alpha$ are disjoint if $c \in S_\alpha$, $\psi_\alpha(c)$ is a formula $A(x)$ which has all its bound variables restricted to $X_\alpha$ and which may have elements of $S_\beta$ with $\beta < \alpha$ appearing as constants. The function $\psi_\alpha$ must put $S_\alpha$ into one to one correspondence with the set of all such functions. The set $S_0$ is defined as the set $\omega \cup \{a\}$ where a is a formal symbol. We will write $S = \cup_\alpha S_\alpha$. Take note that each $c \in S$ is in a unique $S_\alpha$. When we are

done and finally choose the particular set $a \subseteq \omega$, we shall be able to define for each $c \in S_\alpha$ a set $\bar{c}$ as follows. For c in $S_\alpha$, $\alpha > 0$, let $\psi_\alpha(c) = A(x, c_1, ..., c_m)$, $c_i \in S_{\beta_i}$, $\beta_i < \alpha$ where $A(x, t_1, ..., t_m)$ is a formula in ZF with all bound variables restricted to $X_\alpha$. For c in $S_0$ $\bar{c}$ is obviously defined.

To indicate that a bound variable x is restricted to $X_\alpha$ we shall write $\forall_\alpha x$ or $\exists_\alpha(x)$. If we define $X_\alpha$ inductively as $\{\bar{x} | c \in S_\beta \& \beta < \alpha\}$, then $\bar{c} = \{x \in X_\alpha | A(x, \bar{c_1}, ..., \bar{c_m})\}$

In the most abstract sense, we will demonstrate the consistency of ZFC and $\neg CH$ by creating a model that models ZFC and $\neg CH$. The construction method for this model will start with our ground model M and then we will add an element $G$ (known as the generic) that will serve the roll of creating our set $\aleph_2$ that does not follow CH. Constructing G is one of the difficult parts of forcing; we have the tricky task of maintaining control of our ground model so that it continues to behave well, and the problem of actually creating a G that accomplishes our task and even knowing it exists. The solution to these problems comes in the form of the forcing condition and actual forcing operation.

**Definition 5.3. Limited Statement** A limited statement is a statement in ZF in which every quantifier is of the form $\forall_\alpha$ or $\exists_\alpha$ for some ordinal $\alpha < \alpha_0$ and which may involve elements of S as constants.

**Definition 5.4. Unlimited Statement** An unlimited statement is a statement in ZF which may involve elements of F as constants.

**Definition 5.5. Rank** If $A$ is a limited statement, let rank $A = (\alpha, i, r)$ where

- $\alpha$ is the least ordinal such that if $\forall_\beta$ or $\exists_\beta$ occur in $A$, then $\beta \leq \alpha$ and if $c \in S_\beta$ occurs in $A$, then $\beta < \alpha$.
- $r$ is the number of symbols in $A$.
- $i = 0$ if $\alpha$ is a successor ordinal, say $\alpha = \beta + 1$, and $\exists_\alpha$ and $\forall_\alpha$ do not occur in A, and no term of the form $c \in (\cdot), c = (\cdot)$, or $(\cdot) = c$, occurs in A where $c \in S_\beta$. Otherwise $i = 1$.

These definitions help us specify the complexity and type of statements we will be using. A limited statement can be thought of as a statement that has been built up in a finite number of steps from the label space. Rank represents the number of steps it takes to construct the statement where $\alpha$ is the steps. The condition relating to i is necessary since we can encounter terms that occupy similar spots in the label space and it becomes difficult to construct out forcing conditions.

**Definition 5.6. Forcing Condition** A forcing condition P is a finite set of limited statements of the form $n \in G$ or $\neg n \in G$ where $n \in \omega$ and $G$ and $n$ are regarded as belonging to $S_0$, and such that for given n, not both $n \in G$ and $\neg n \in G$ are in P.

We use our forcing conditions P to form a chain of $P_n$ where $P_n \subseteq P_{n+1}, ....$ This chain forms a consistent family of conditions such that for every property A, $P_n$ forces $A$ or $\neg A$. This allows us to determine our generic set G since it determines every property of G. Now that we have all the needed equipment, we can define our forcing operation.

**Definition 5.7. P forces A** We define P forces A, for limited statement A, by induction on rank A as follows:

1. P forces $\exists_\alpha x B(x)$ if some $c \in S_\beta, \beta < \alpha$, P forces $B(c)$.
2. P forces $\forall_\alpha x B(x)$ if for all $Q \supseteq P$ and $c \in S_\beta,\ \beta < \alpha$
3. P forces $\neg B$ if for all $Q \supseteq P, Q$ does not force B.
4. P forces $B \& C$ if P forces B and P forces C
5. P forces $B \wedge C$ if P forces B or P forces C
6. P forces $B \rightarrow C$ if P forces C or P forces $\neg B$
7. P forces $B \leftrightarrow C$ if P forces $B \rightarrow C$ and P forces $C \rightarrow B$
8. P forces $c_1 = c_2$, where $c_1 \in S_\alpha,\ c_2 \in S_\beta,\ \gamma = \max(\alpha, \beta)$ if either $\gamma = 0$ and $c_1 = c_2$ as elements of $S_0$ or $\gamma > 0$ and P forces $\forall_\gamma x(x \in c_1 \implies x \in c_2)$.
9. P forces $c_1 \in c_2$, where $c_1 \in S_\alpha,\ c_2 \in S_\beta,\ \alpha < \beta$ if P forces $A(c_1)$ where $A(x) = \psi_\beta(c_2)$(ie,. $A(x)$ is the formula defining $c_2$).
10. P forces $c_1 \in c_2$, where $c_1 \in S_\alpha,\ c_2 \in S_\beta, \alpha \geq \beta$ and not $\alpha = \beta = 0$, if for some $c_3 \in S_\gamma$, $\gamma < \beta$ if $\beta > 0$, $\gamma = 0$ if $\beta = 0$, P forces $\forall_\alpha(x \in c_1 \leftrightarrow x \in c_3) \& (c_3 \in c_2)$
11. P forces $c_1 \in c_2$, where $c_1, c_2 \in S_0$ if $c_1, c_2 \in \omega$ and $c_1 \in c_2$ (as elements of $S_0$) or $c_2 = a$ and the statement $c_1 \in G$ is in P.

We define P forces A where A is an unlimited statement by induction on the number of symbols in A as follows:

1. P forces $\exists x B(x)$ if for some $c \in S$, P forces $B(x)$
2. P forces $\forall x\ B(x)$ if for all $c \in S,\ Q \supseteq P, Q$ does not force $\neg B(c)$
3. P forces $\neg B$ if for all $Q \supseteq P, Q$ does not force B.
4. P forces $B \& C$ if P forces B and P forces C
5. P forces $B \wedge C$ if P forces B or P forces C
6. P forces $B \rightarrow C$ if P forces C or P forces $\neg B$
7. P forces $B \leftrightarrow C$ if P forces $B \rightarrow C$ and P forces $C \rightarrow B$
8. P forces $c_1 \in c_2$ or $c_1 = c_2$ if it forces them as limited statements.

## 5.2 Useful Forcing Lemmas

Here we prove some lemmas that show that forcing is a well-defined operation that will suit our purposes.

**Lemma 5.2.** If P forces A and $Q \supseteq P$ then Q forces A.

*Proof.* We prove this first for limited A by induction on rank $A = (\alpha, i, r)$. In cases 4 to 11 in the definition of forcing do not require a discussion since P forces A is reduced to P forces B for some B with rank B less than rank A. Now our discussion of case one

1. Case one: If P forces $\exists_\alpha B(x)$ then P forces $B(c),\ c \in S_\beta, \beta < \alpha$, so by induction Q forces $B(c)$ so Q forces $\exists_\alpha x B(x)$.

2. If P forces $\forall_\alpha x B(x)$, and $Q \supseteq P$ then if $R \supseteq Q, R \supseteq P$ also so R does $\neg B(c)$ for any $c \in S_\beta$, $\beta < \alpha$ so Q forces $\forall_\alpha x B(x)$

3. If P forces $\neg B,\ Q \supseteq P$ and $R \supseteq Q,\ R \supseteq P$, also, so R does not force B so Q forces $\neg B$.

Now if A is unlimited, the same arguments apply for cases 1 to 7 and case 8 is handled by referring to the cases for limited arguments. □

**Lemma 5.3.** For all P and A there is a $Q \supseteq P$ such that either Q forces A or Q forces $\neg A$

*Proof.* This is a suprising fact about forcing, namely that every statement about the Generic G is decidable by finitely many statements of the form $n \in G$ or $\neg n \in G$. The proof is that if P does not force $\neg A$ by definition it is because for some $Q \supseteq P$, Q forces A. □

**Definition 5.8. Complete Sequence** A sequence $\{P_n\}$ of forcing conditions is a complete sequence if $P_n \supseteq P_{n+1}$ for all n and for every A, limiteed or unlimited, $\exists n$ such that either $P_n$ forces A or $P_n$ forces $\neg A$. The sequence $\{P_n\}$ (*i.e.*, $\{n, P_n\}$) is not assumed to be in M

Complete sequences are important since they determine every property of our Generic. The only question now to be answered is do they exist

**Lemma 5.4.** A complete sequence exists.

*Proof.* Here we use the countablity of our standard model M. Since M is countable we can enumerate all statements $A_n$. Define $P_n$ by induction as any forcing condition $Q \supseteq P_{n-1}$ such that either Q forces $A_n$ or Q forces $\neg A_n$ (Since the set of all P is countable, Axiom of choice is not needed).
If $\{P_n\}$ is complete then in particular for every k either some $P_n$ forces $k \in G$ or $\neg k \in G$. Let $G = \{k | \exists n (P_n f \text{ forces } k \in G)\}$. Then $G \subseteq \omega$ and as we remarked we can now define a map $c \to \bar{c}$ defined for all c in S which sends $G$ into $\bar{G}$ and we can then define N as $\cup\{M_\beta(G) | \beta < \alpha_0\}$. □

Our final lemma we state without proof since it is still important but the proof is too lengthy for the scope of this paper.

**Lemma 5.5.** The generic extension of M, A is true in M[G] if and only if for some $n$ $P_n$ forces A.

This final lemma allows us to connect the notion of forcing and truth in M[G]. The other lemmas show that forcing works to define our generic in a reasonable way and that the generic sets can actually be created.

# 6 Independence of the Continuum Hypothesis

Let $\aleph_\tau, \tau \geq 2$, be a fixed cardinal in M. Let S be defined as follows, For all $\alpha < \aleph_\tau$, $S_\alpha$ consists of one element $c_\alpha$ and it will turn out that $\bar{c_\alpha} = \alpha$. All these $c_\alpha \in G$. For $\alpha \in \aleph_\tau, S_\alpha$ consists of $\aleph_\tau$ elements, all in G which we denote by $a_\delta, \delta < \aleph_\tau$. For these we will have $\bar{a_\delta} \subseteq \omega$ and their presence will guarantee that the continuum is at least $\aleph_\tau$.

For brevity's sake, here we present two lemmas without proof that will be used in the final proof of the continuum hypothesis.

**Lemma 6.1.** The continuum is not the sum of countably many smaller cardinals

$\aleph_{\alpha+1}$ is known as the successor cardinal to $\aleph_\alpha$

**Lemma 6.2.** In M, the number of countable subsets of $\aleph_\tau$ is $\aleph_\tau$ if $\aleph_\tau$ is not the countable sum of smaller cardinals, $\aleph_{\tau+1}$ otherwise.

These place some restrictions on the relationships between the continuum and its successor. The first lemma can also be rephrased as the continuum is a regular cardinal. Regular cardinals are equal to their own cofinality. An example of a non regular cardinal is $\aleph_\omega$ since its initial ordinal $\omega_\omega$ has order type $\omega$ making it singular and making the cardinal singular, or not regular. The second lemma explains what happens if $\aleph_\tau$ is regular or not.

Before the final proof a quick definition

**Definition 6.1. Minimal**
For any statement A, we say that P is minimal for A if P forces $\neg\neg A$ (or equivalently no $Q > P$ forces $\neg A$), and no $P < P'$ has this property.

**Theorem 6.3.** In M[G], $C = \aleph_\tau$ if $\tau$ is not cofinal with $\omega$ in M, $C = \aleph_{\tau+1}$ if $\tau$ is cofinal with $\omega$

*Proof.* For each $c \in S$ and $n \in \omega$, let $V(n, c)$ be the set of minimal P for $n \in c$. If $V(n, c) = V(n, c')$ for all n and $\bar{c} \subseteq \omega$, $\bar{c}' \subseteq \omega$, then $\bar{c} = \bar{c}'$. This follows since if $P_k$ is in the complete sequence and $P_k$ forces $n \in c$, there must be a minimal $P'$, $P' < P_k$. Then $P'$ and hence $P_k$ force $\neg\neg n \in \bar{c}'$ so $n \in \bar{c}'$. Thus $\bar{c} = \bar{c}'$.

For fixed n, the number of possible sets $V(n, c)$ is at most $|D|$ = the number of possible subsets of $\aleph_\tau$, as computed in M. Now we claim that if E is the set of

countable sequences in D, $|D| = |E|$. Each element of E gives rise to a countable subset of D and each subset comes from at most $2^{\aleph_\tau}$ sequences so we know that from lemma 6.2, we get $|E| \leq |D| * \aleph_1 = |D|$, and since $|D| \leq |E|$, $|D| = |E|$. Thus we have shown that in M[G] $|C| \leq |D|$. If $\tau$ is cofinal with $\omega$, we see that in M[G], $\aleph_\tau \leq C \leq \aleph_{\tau+1}$. Lemma 6.1 says that C cannot equal $\aleph_\tau$ so $C = \aleph_{\tau+1}$. If $\tau$ is not cofinal with $\omega$, we have $C \leq \aleph_\tau$ and since $C \geq \aleph_\tau, C = \aleph_\tau$. From this we see that the continuum hypothesis is false and that additionally we can construct C to be equal to many cardinalities such as $\aleph_2, \aleph_{\omega+1}, \aleph_{\omega^2+1}$ etc. □

# 7 Acknowledgments

I would like to first thank my mentor Diego Andres Bejarano Rayo for teaching me model theory and building up my logic knowledge to tackle forcing as well as answering my numerous questions along the way. I would also like to thank Professor Peter May for running the University of Chicago 2018 REU and making my learning experience possible and Daniil Rudenko for running the apprentice program and creating a wonderfully made class which combined so many fascinating mathematical topics. Finally I would like to thank the professors and lecturers who gave talks during the program and my numerous peers who all taught me more math than anybody could have expected to learn in one summer.

# References

[1] Chicago Style Citation Jech, Thomas J. Set Theory. Berlin ; New York: Springer, 2003.

[2] Kunen, Kenneth. Set Theory: An Introduction to Independence Proofs. Amsterdam ; New York: North-Holland Pub. Co., 1980.

[3] Chang, Chen Chung, and H. Jerome Keisler. Model Theory. Amsterdam, New York: North-Holland Pub. Co.; American Elsevier, 1973.

[4] Cohen, Paul J. Set Theory and the Continuum Hypothesis. New York: W. A. Benjamin, 1966.

[5] Smullyan, Raymond M., and Melvin Fitting. Set Theory and the Continuum Problem. New York: Clarendon Press, 1996.