The Continuum Hypothesis and Set-Theoretic Forcing

Tristan Knight
University of Connecticut

May 21, 2019

1 Introduction

In this paper, I discuss the statement and history of the continuum hypothesis, its importance in mathematics, and the impact of its solution. I then prove an overview of the proof of the independence of the continuum hypothesis from the axioms of ZFC. I first build a model that satisfies the generalized continuum hypothesis, then build a model that does not. The proofs assume familiarity with some model theory and set theory; the proofs of those theorems either too basic to spend time addressing or too complex to meaningfully contribute to understanding of the topic are omitted. The omitted proofs can be found in most set-theory books that discuss forcing, and a few helpful texts are included in the references.[**7**][**8**]

2 History of the continuum hypothesis

The continuum hypothesis originates with the mathematician Georg Cantor (1845-1918), the founder of modern set theory. Cantor's work was responsible for bringing the concept of infinity into mainstream mathematical discourse, and most importantly the concept of cardinality and varying degrees of infinity. Cantor famously discovered that the set of real numbers (the continuum) has a strictly greater cardinality than the set of natural numbers.

Cantor first posits the continuum hypothesis in his 1877 paper "Ein Beitrag zur Mannigfaltigkeitslehre"[**1**]. In its more modern restatement, it reads: "There is no set with cardinality strictly between that of the natural numbers and the real numbers." Cantor spent much of his life attempting to prove this hypothesis true, to no avail. Despite this, he believed wholeheartedly in its veracity.

Although Cantor's work was considered controversial by many important mathematicians when it was introduced, by the turn of the century it had garnered significant support. Esteemed mathematician David Hilbert (1862-1943) included the continuum hypothesis in his 1900 Paris lecture "Mathematische Probleme"[**2**] as the first among twenty-three questions that would become famously known as the "Hilbert problems."

Around this same time period, a crisis was developing in the philosophy of mathematics. There had arisen a desire for mathematics to be placed upon solid foundational principles; namely, for a collection of axioms from which the totality of mathematics could be derived. The early 1900s brought an influx of attempts at such a system; each iteration seemed to invariably fall to some paradox or error. Hilbert was a major force in the search for such an axiomatic system. Hilbert's program, as this search came to be known, "calls for a formalization of all of mathematics in axiomatic form, together with a proof that this axiomatization of mathematics is consistent."[**3**] The most famous, and perhaps most successful axiomatic system created during this time, is the Zermelo-Fraenkel set theory.

Unfortunately, this program was doomed to fail from the beginning; in 1931, Kurt Gödel (1906-1978) proved his incompleteness theorems.[**4**] These theorems essentially stated that any system powerful enough to be a candidate for the foundation of mathematics must necessarily be either self-contradictory or be unable to prove the truth of all its true statements. However, not all is lost. From this situation rose a new problem: that of which axioms are, so to speak, proper to assume. the axioms of ZFC (the Zermelo-Fraenkel set theory with the axiom of choice added) rose as a popular base theory to work with. From this questionv arose the importance of consistency and independence proofs. Mathematicians were eager to find important theorems and problems which were independent of or contradictory to ZFC; among these was the ever-present continuum hypothesis.

It was proven much earlier, again by Gödel in 1940,[**5**] that the continuum hyothesis was consistent with the axioms of ZFC using inner models. It wasn't until 1963 that Paul Cohen (1934-2007) was able to show that there were models which did not satisfy the continuum hypothesis. By showing that there were models satisfying ZFC that both did and did not satisfy the continuum hypothesis, it is shown to be an independent axiom.

Cohen's proof[**6**] introduces a novel mathematical technique known as forcing, which can be thought of in terms somewhat similar to the process of extending a field. Forcing allowed for an unprecedented degree of control over the construction of certain models, and proved very impactful in the study of model theory. Still, the most popular and well-known application of Cohen's technique is its role in finally proving the independence of the continuum hypothesis.

3 A model that satisfies CH

In this section, we define a model of ZFC in which the continuum hypothesis is true. This construction predates Cohen's notion of forcing; although one can use forcing to make such a model, we will not.

3.1 Building the constructible universe

Definition 3.1. Let A be a set. Define the *definable powerset of A* by

$$\mathcal{D}(A) = \{\{a \in A : \phi^A(a, b_1, b_2, \ldots, b_n)\} : \phi \text{ is first-order and } b_1, b_2, \ldots, b_n \in A\}.$$

We use the above notion to define the constructible universe using transfinite recursion as follows.

Definition 3.2. We define the following sets for each ordinal α:

- $L(0) = \emptyset$
- $L(\alpha + 1) = \mathcal{D}(L(\alpha))$
- If α is a limit ordinal, $L(\alpha) = \bigcup_{\beta<\alpha} L(\beta)$

We define the *constructible universe* $\mathbb{L}$ by

$$\mathbb{L} = \bigcup_{\alpha \in \mathbb{ON}} L(\alpha)$$

Theorem 3.1. Let κ be an infinite cardinal. Then $|L(\kappa)| = \kappa$.

Theorem 3.2 (Mostowski Collapse Lemma)**.** Let R be a binary relation on a class X that satisfies the following:

- For every $x \in X$, $\{y : yRx\}$ is a set (set-like);
- Every nonempty $Y \subset X$ has an R-minimal element (well-founded);
- For every $x, y \in X$, if $x \neq y$, then $\{z : zRx\} \neq \{z : zRy\}$ (extensional).

Then there is a unique transitive class S such that $(S, R) \cong (X, R)$.

Definition 3.3. The *axiom of constructability*, denoted as $\mathbb{V} = \mathbb{L}$, is the statement that every set is an element of $\mathbb{L}$. This axiom is consistent with ZFC; $\mathbb{L}$ is a model of the theory.

Theorem 3.3. If $\mathbb{V} = \mathbb{L}$, then for all $\alpha \geq \omega$, $\mathcal{P}(L(\alpha)) \subseteq L(\alpha^+)$.

Before continuing with the proof, we need the following theorem, along with the Reflection Principle:

Theorem 3.4. Let $\Gamma \subset$ ZFC $+\mathbb{V} = \mathbb{L}$ be the finitely many axioms needed such that:

- the notions of ordinal, rank, and $L(\alpha)$ are absolute for transitive models of Γ;
- the statement that says there is no greatest ordinal can be proven from Γ;
- the axiom $\mathbb{V} = \mathbb{L} \in \Gamma$;

- the axiom of extensionality is in Γ.

If M is a transitive set such that $M \models \Gamma$, then $M = L(o(M))$.

Theorem 3.5 (Reflection Principle)**.** Let Γ be a finite set of formulas of ZFC. If $\mathbb{V} = \mathbb{L}$, then for each ordinal δ, there is an ordinal $\beta > \delta$ such that Γ is absolute for $L(\beta)$.

We now begin the proof of Theorem 3.3.

Proof. Let α be an ordinal, and fix $B \subset L(\alpha)$. By the definition of $\mathbb{L}$, we know there is some $\delta > \alpha$ such that $B \in L(\delta)$.

By Theorem 3.5, there is a $\beta > \delta$ such that Γ from Theorem 3.4 is absolute for $L(\beta)$; then $L(\beta) \models \Gamma$. Define the set $X = L(\alpha) \cup \{B\} \subset L(\beta)$.

By the Downward Lowenheim-Skolem Theorem, there exists $A \subset L(\beta)$ such that $(A, \in) \preceq (L(\beta), \in)$, $X \subseteq A$, and $|A| = |X|$. Notably, this tells us that $|A| = |\alpha| < \alpha^+$ and $(A, \in) \models \Gamma$; and since the axiom of extensionality is in Γ, we know $\in$ is extensional on A.

Let $(M, \in)$ be the Mostowski collapse of A. Since $\in$ is extensional on A, $(M, \in) \cong (A, \in)$. Then $M \models \Gamma$, so by Theorem 3.4, $M = L(o(M))$. But we know that $|M| = |A| = |\alpha|$, so $o(M) < \alpha^+$. Since $L(\alpha)$ is transitive, and $L(\alpha) \subset A$, we have $L(\alpha) \subset M$. Note that $B \subseteq M$, and the Mostowski collapse function is the identity on B. Thus, $B \in M$, and we get $B \in L(\alpha^+)$. Since B was arbitrary, we have $\mathcal{P}(L(\alpha)) \subseteq L(\kappa^+)$. □

Theorem 3.6. If $M \models (\text{ ZFC} + \mathbb{V} = \mathbb{L})$, then $M \models \text{CH}$.

Proof. Let M be a model satisfying the above, κ be an infinite cardinal, and $X \subseteq \kappa$. Since $\kappa \subseteq L(\kappa)$, we know that $X \subseteq L(\kappa)$ and that $X \in \mathcal{P}(L(\kappa))$. By Theorem 3.3, it follows that $X \in L(\kappa^+)$. Since $X \subseteq \kappa$ was arbitrary, $\mathcal{P}(\kappa) \subseteq L(\kappa^+)$. By Theorem 3.1, $|L(\kappa^+)| = \kappa^+$, so $|\mathcal{P}(\kappa)| \leq \kappa^+$. □

4 A model that does not satisfy CH

In this section, we use forcing to construct a model that does not satisfy the continuum hypothesis. This direction, to the best of my knowledge, requires forcing. The manner in which we do so here does so by way of a complete Boolean algebra that can be thought of as coding possible truth values of first order statements in a base model M. By defining a filter G on B with certain properties, we extend M into a new model $M[G]$. The filter G can be thought of as collapsing the "possible" truth values in B into true and false statements in the new model; by controlling particular aspects of M and B, we can fine-tune $M[G]$ to satisfy certain properties (in our case, one that contradicts the continuum hypothesis).

4.1 Partial orders

Definition 4.1. A *partially ordered set* (also called a *poset*) is a set P equipped with a binary relation $\leq$ such that the following properties hold:

- (reflexivity) $p \leq p$
- (antisymmetry) if $p \leq q$ and $q \leq p$, then $p = q$
- (transitivity) if $p \leq q$ and $q \leq r$, then $p \leq r$

Two elements $p, q \in P$ are *incompatible*, written $p \perp q$, if there is no $r \in P$ such that $r \leq p$ and $r \leq q$. A poset P is *atomless* if for all $p \in P$, there are $q, r \leq p$ where $q \perp r$.

Definition 4.2. Let P be a poset. We say $A \subset P$ is an *antichain* if, for any $p, q \in A$, $p \perp q$.

Further, we say that P satisfies the *countable chain condition* (or c.c.c.) if every antichain in P is countable.

For the purposes of this paper, we will be primarilly discussing posets with a greatest element. This element will normally be denoted by 1.

4.2 Boolean algebras and filters

Definition 4.3. A *Boolean algebra* is a set B equipped with binary operations $\wedge$ and $\vee$, a unary operation $\neg$, and two elements $0, 1 \in B$ such that the following properties hold:

- (commutativity) $u \wedge v = v \wedge u$ and $u \vee v = v \vee u$
- (associativity) $u \wedge (v \wedge w) = (u \wedge v) \wedge w$ and $u \vee (v \vee w) = (u \vee v) \vee w$
- (distributivity) $u \wedge (v \vee w) = (u \wedge v) \vee (u \wedge w)$ and $u \vee (v \wedge w) = (u \vee v) \wedge (u \vee w)$
- $u \wedge u = u$ and $u \vee u = u$
- $u \wedge (u \vee v) = u$ and $u \vee (u \wedge v) = u$
- $u \vee 0 = u$ and $u \wedge 0 = 0$
- $u \vee 1 = 1$ and $u \wedge 1 = u$
- $u \wedge (\neg u) = 0$ and $u \vee (\neg u) = 1$
- $\neg(\neg u) = u$
- $\neg(u \wedge v) = (\neg u) \vee (\neg v)$ and $\neg(u \vee v) = (\neg u) \wedge (\neg v)$

Define $\leq$ on B by $a \leq b$ if and only if $a \wedge (\neg b) = 0$. This relation forms a partial order on B, with 0 and 1 as the greatest and least element, respectively. B is *complete* if every set of elements in B has a least upper bound.

Definition 4.4. A *filter* on a Boolean algebra B is a subset F such that the following properties hold:

- $1 \in F$ and $0 \notin F$
- if $u, v \in F$ then $u \wedge v \in F$
- if $u \in F$ and $u \leq v$ then $v \in F$

Further, F is an ultrafilter if it satisfies the following additional property:

- for all $u \in B$, $u \in F$ or $\neg u \in F$

4.3 Relating Boolean algebras and posets

Definition 4.5. Let P be a poset with a greatest element $1 \in P$. A *P-name* is a set n of ordered pairs (m, p) such that m is a P-name and $p \in P$.

Definition 4.6. Let B be a Boolean algebra (or poset) and A a subalgebra (or subset) of B. We say A is *dense* in B if for every nonzero $b \in B$, there is a nonzero $a \in A$ such that $a \leq b$.

Definition 4.7. Let G be a filter on a Boolean algebra B. G is *generic* if, for every dense $D \subset B$, we have $G \cap D \neq \emptyset$.

Definition 4.8. A *completion* of a Boolean algebra B is a complete Boolean algebra C such that B is dense in C.

Definition 4.9. A poset P is *separative* if, for $p, q \in P$ with $p \not\leq q$, there is $r \leq p$ where $r \perp q$.

Theorem 4.1. Let P be a separative poset. Then there is a complete Boolean algebra B, unique up to isomorphism, such that:

- $P \subset B \setminus \{0\}$, and the partial ordering of P agrees with the partial ordering of B;
- P is dense in B.

Definition 4.10. Let A, B be sets. We define the *set of finite partial functions from A to B*, denoted as $Fn(A, B)$ to be the set of all finite partial functions from A to B.

Theorem 4.2. Let $P = Fn(\kappa \times \omega, 2)$ for some infinite cardinal κ. Define $\leq$ on P such that, for $f, g \in P$, $a \leq b$ if and only if $a \supseteq b$. Then $(P, \leq)$ is a poset. Furthermore, it is separative and atomless.

Proof. It follows immediately from Definition 4.1 that $(P, \leq)$ is a poset. Let $f, g \in P$. Since f is finite, there is some $(\alpha, n) \in \kappa \times \omega$ such that $f(\alpha, n)$ is undefined. Let $h = f \cup \{((\alpha, n), 0)\}$. Then $h \leq f$, and $(P, \leq)$ is atomless. For separativity, suppose further that $f \not\leq g$. Then $g \not\subseteq f$: there must be some $x \in \kappa \times \omega$ such that either $f(x) \neq g(x)$ or $g(x)$ is defined and $f(x)$ is undefined. In the former case, we get $f \perp g$ by definition; thus, we assume the latter. Without loss of generality, suppose $g(x) = 0$. Let $f' \in P$ be the partial function $f \cup \{(x, 1)\}$. Then $f' < f$ and $f' \perp g$. Thus, $(P, \leq)$ is separative. □

Theorem 4.3. The poset $Fn(\kappa \times \omega, 2)$ satisfies the countable chain condition.

4.4 Building $M[G]$

Definition 4.11. Let M be a countable transitive model of ZFC, and let G be a set such that $G \notin M$. We let $M[G]$ denote the smallest transitive model of ZFC such that $M \subset M[G]$ and $G \in M[G]$.

Theorem 4.4. Let M be a countable transitive model of ZFC. Let κ be an ordinal such that $M \models (\kappa = \aleph_\alpha)$ for some $\alpha > 1$. Let $P = Fn(\kappa \times \omega, 2)$. Let B be the corresponding complete Boolean algebra with respect to P, and let $G \notin M$ be a generic ultrafilter over B. Then $M[G] \models 2^\omega \geq \kappa^M$.

Proof. Let $F_G = \bigcup(P \cap G)$. We want to show that $F_G : (\kappa \times \omega) \to 2$ is a function. We first show that F_G is well-defined. Seeking contradiction, suppose $((\alpha, n), 0), ((\alpha, n), 1) \in F_G$. Then there must be $p, q \in G$ such that $((\alpha, n), 0) \in p$ and $((\alpha, n), 1) \in q$. This entails that $p \perp q$ in P, so $p \wedge q = 0$ in B. Because G is a filter, $p \wedge q \in G$, so $0 \in G$. This contradicts the definition of a filter; thus, F_G is well-defined.

We now show that the domain of F_G is $\kappa \times \omega$. Define $D_{\alpha,n}$ to be the set

$$D_{\alpha,n} = \{p \in P : p(\alpha, n) \text{ is defined}\}.$$

This set is dense in P. Fix $(\alpha, n) \in \kappa \times \omega$, and let $p \in P$. If $p(\alpha, n)$ is defined, then we are done. Otherwise, the function $p' = p \cup \{((\alpha, n), 0)\}$ is in $D_{\alpha,n}$, and therefore $D_{\alpha,n}$ is dense in P. Since P itself is dense in B, we have that $D_{\alpha,n}$ is dense in B as well; thus, $G \cap D_{\alpha,n} \neq \emptyset$ for any pair $(\alpha, n) \in \kappa \times \omega$. Therefore, F_G is defined on all of $\kappa \times \omega$.

For each $\alpha < \kappa$, let

$$X_\alpha = \{n \in \omega : F_G(\alpha, n) = 1\}.$$

Let $\alpha, \beta < \kappa$ such that $\alpha \neq \beta$. Define the set

$$A_{\alpha,\beta} = \{p \in P : \exists n[p(\alpha, n) \neq p(\beta, n)]\}.$$

Since $A_{\alpha,\beta}$ is dense in P, we have $G \cap A_{\alpha,\beta} \neq \emptyset$. It follows that $X_\alpha \neq X_\beta$ for any $\alpha \neq \beta$. Then each X_α denotes a distinct subset of ω in M; thus, we have $M[G] \models 2^\omega \geq \kappa^M$. $\square$

4.5 Forcing

Definition 4.12. Let P be a poset, B the associated complete Boolean algebra, and G a generic ultrafilter over B. For a P-name $n \in M$, we define *the interpretation of n in G* to be the set

$$n^G = \{m^G : (m, p) \in n \text{ and } p \in G\}.$$

Conversely, let $a \in G$. We define the set $\hat{a}$ as

$$\hat{a} = \{(\hat{b}, 1) : b \in a\},$$

where 1 denotes the maximal element of B.

Note that, by the definition above, $\hat{a}^G = a$ for any $a \in G$. Intuitively, this gives us a way to talk about elements of $M[G]$ while remaining within the scope of M.

Theorem 4.5. Given a countable transitive model M of ZFC, a poset $P \in M$ and its associated complete Boolean algebra B, and a generic ultrafilter G over B, define

$$N = \{n^G : n \in M \text{ and } n \text{ is a } P\text{-name}\}.$$

Then the following hold:

- $M \subset N$;
- $G \in N$;
- $N \models$ ZFC;
- N is transitive.

Thus, $M[G] \subseteq N$. On the other hand, for each P-name $n \in M$, $n^G \in M[G]$. Thus, $M[G] = N$.

Definition 4.13. Let ϕ be a first-order sentence, and M, P, and B be as defined previously. For $p \in P$, we say p *forces* ϕ (written "$p \Vdash \phi$") if $M[G] \models \phi$ in every generic ultrafilter G such that $p \in G$.

We also define the *forcing language over* B to be the first-order language with binary membership relation $\in$ and the set of P-names as its constants.

Theorem 4.6 (Truth Lemma)**.** For every sentence ϕ in the forcing language, there is a $[\![\phi]\!] \in B$ such that $p \Vdash \phi$ if and only if $p \leq [\![\phi]\!]$. In addition, the first-order logical operations are preserved in B: $[\![\neg\phi]\!] = \neg[\![\phi]\!]$ and $[\![\phi \wedge \psi]\!] = [\![\phi]\!] \wedge [\![\psi]\!]$.

Furthermore, for any sentence ϕ in the forcing language, $M[G] \models \phi$ if and only if there is a $p \in G$ such that $p \Vdash \phi$.

Theorem 4.7 (Definability Lemma)**.** Let M, B, P, G be as defined previously, and let $p \in B$. For every sentence ϕ in the forcing language, the statement $p \Vdash \phi$ is definable in M. Alternatively, there is a relation $\Vdash^* \in M$ such that, for P-names $p_1, \ldots, p_n \in M$ and sentence ϕ in the forcing language,

$$M[G] \models \phi(p_1, \ldots, p_n) \iff (p \Vdash^* \phi(p_1, \ldots, p_n))^M.$$

Theorem 4.8. Let M be a countable transitive model of ZFC. Let κ be an ordinal such that $M \models (\kappa = \aleph_\alpha)$ for some $\alpha > 1$. Let $P = Fn(\kappa \times \omega, 2)$. Let B be the corresponding complete Boolean algebra with respect to P, and let $G \notin M$ be a generic ultrafilter over B. Then $M[G] \models (\aleph_\alpha^M = \aleph_\alpha^{M[G]})$.

Proof. Let $\sigma, \sigma^+ \in M$ such that $|\sigma| = \aleph_\beta^M$ and $|\sigma^+| = \aleph_{\beta+1}^M$. From Definition 4.11, $M \subset M[G]$, so $\sigma, \sigma^+ \in M[G]$. We want to show that $M[G] \models (|\sigma| \neq |\sigma^+|)$; that is, that none of the new functions in $M[G]$ from G give a bijection between σ and σ^+.

Seeking contradiction, suppose there is some P-name $f \in M$ such that f^G defines a bijection in $M[G]$ between $\hat{\sigma}$ and $\hat{\sigma^+}$. By Theorem 4.6, there is some $p \in G$ that forces this property. From this, we define the function $F : \sigma \to \mathcal{P}(\sigma^+)$ as follows. We have

$y \in F(x)$ if and only if there is some $p_{x,y} \in B$ such that $p_{x,y} \leq p$ and $p_{x,y} \Vdash (f(\hat{x}) = \hat{y})$. By Theorem 4.7 and Comprehension in M, we see that F is definable in M. Also, since P is dense in B, we can further specify that $p_{x,y} \in P$.

Let $y, y' \in \sigma^+$ such that $y \neq y'$. Then $p_{x,y}$ forces f to be a different bijection than $p_{x,y'}$ does; thus, by Theorem 4.6, we have $p_{x,y} \perp p_{x,y'}$. By Theorem 4.3, any antichain in P is countable; thus, $F(x)$ is countable for each $x \in \sigma$. Since $M \models (|\sigma| < |\sigma^+|)$, we have $\bigcup_{x \in \sigma} F(x) \neq \sigma^+$. However, our assumption was that f^G is a bijection from σ to σ^+ in $M[G]$. Hence, for any $y \in \sigma^+$, there must be an $x \in \sigma$ and $p_{x,y} \in G$ such that $p_{x,y} \leq p$ and $p_{x,y} \Vdash (f(\hat{x}) = \hat{y})$. Thus, for all $y \in \sigma^+$, there is some $x \in \sigma$ such that $y \in F(x)$. This is a contradiction. Therefore, $M[G]$ does not collapse any cardinals in M, and so $M[G] \models (\aleph_\alpha^M = \aleph_\alpha^{M[G]})$. □

References

[1] Cantor, Georg (1878), "Ein Beitrag zur Mannigfaltigkeitslehre", *Journal für die Reine und Angewandte Mathematik*, **82**: 242-248

[2] Hilbert, David (1902), "Mathematical Problems", *Bull. Amer. Math. Soc.*, **8**: 437-479

[3] Zach, Richard (2015), "Hilbert's Program", *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/hilbert-program/

[4] Gödel, Kurt (1931), "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I", *Monatshefte für Mathematik und Physik*, **38**: 173-198

[5] Gödel, Kurt (1940), *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Princeton University Press

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

[7] Easwaran, Kenny (2007), "A Cheerful Introduction to Forcing and the Continuum Hypothesis" `arXiv:0712.2279v1 [math.LO]`

[8] Jech, Thomas (1971), *Set Theory*, Springer–Verlag, Heidelburg