Equivalents, Applications, and Consistency of the Axiom of Choice

Xi Yao

Supervised by Professor Greg Hjorth

xiyao17@gmail.com

Department of Mathematics and Statistics
The University of Melbourne

November 2010
To Hui-Ying Hu
Acknowledgments

I would like to particularly thank my supervisor Professor Greg Hjorth. As a previous biology student who had never done any pure mathematics before and somehow decided to do a pure mathematics Master thesis in set theory and logic, and as an international student who had to work a lot to make a living, I have experienced a hard time. Professor Greg Hjorth has been very patient and considerate, helping me through many of the elementary aspects of pure mathematics.

I also want to thank Raj A. Dahya, Leigh Humphries, and Toby Meadows for the discussions about \( \mathbb{L} \), Dr Lawrence Reeves for the discussions about the Banach–Tarski Paradox, Dr Paul Norbury for the discussions on the proof of \( \text{Vec} \Rightarrow \text{AC} \), and Bowen Plug for some general discussions about AC and correcting some of my grammatic mistakes.
Abstract

Given any family of non-empty sets, the Axiom of Choice (AC) gives us the power to choose one element from each member from the family. AC appears and is used in most modern mathematics and sciences, more often than not, implicitly. For example, when we say “choose coset representatives”, we are indeed using AC.

Upon its formal formulation after unconscious uses, AC had been receiving heavy criticisms for years due to its non-constructive nature. However, AC indeed gives us a richer mathematical (and scientific) world.

We first provide, as motivation, some equivalents of AC and results requiring AC from various branches of mathematics. Then we prove the consistency of AC in ZF. Some results from AC, e.g. the Banach–Tarski paradox, are discussed.
Contents

1 Introduction 1

2 Preliminaries 3
   2.1 Zermelo–Fraekel Set Theory . . . . . . . . . . . . . . . . . . . . . 3
   2.2 Ordinals . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
   2.3 Transfinite Induction . . . . . . . . . . . . . . . . . . . . . . . . 9
   2.4 Cardinals . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
   2.5 The universe of sets . . . . . . . . . . . . . . . . . . . . . . . . . . 13

3 Some equivalents of AC 14

4 Some results requiring AC 20
   4.1 Set theory . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
   4.2 General topology . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
   4.3 Algebra . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
   4.4 Group theory . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
   4.5 Measure theory . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
   4.6 Functional Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . 21

5 The consistency of AC in ZF 22
   5.1 Overview of the consistency and its proof . . . . . . . . . . . . . . . . 22
   5.2 A quick review of model theory . . . . . . . . . . . . . . . . . . . . . 22
   5.3 Absoluteness . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26
   5.4 Constructible sets . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
   5.5 Gödel operations . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
   5.6 Con(ZF + V = L) . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
   5.7 Con(ZFC) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
   5.8 ZF + ¬AC . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

6 The Sun and a pea 45

7 Conclusion 47

References 48
1 Introduction

The Axiom of Choice is necessary to select a set from an infinite number of socks, but not an infinite number of shoes.

Bertrand Russell

Along with the eight Axioms in the Zermelo–Fraenkel axiomatic set theory (ZF hereafter), together called ZFC, the Axiom of Choice (AC hereafter) forms a natural and simple but indispensable foundation for most modern mathematics and sciences. It penetrates into most, if not all, branches of mathematics, including set theory, order theory, algebra, topology, group theory, and analysis, just to name a few; and it has many equivalents and applications.

In the history of mathematics, only through extensive use, did assumptions become clean formulas. According to Moore 1982, the development of AC can be divided into four stages. The rudimentary version of AC — picking out an arbitrary element from a set — can be traced back at least to 2300 years ago in Euclid’s Elements. In the second stage, around the start of the nineteenth-century, mathematicians constructed the rule of choice when dealing with infinity. But gradually, mathematicians, like Cauchy, started to leave the rule implicit. The last stage can probably marked by that in 1871, Cantor once used some choice functions that no rule existed and he did not recognize this. After the four stages, AC was first formulated by Ernst Zermelo when he was seeking a proof for the well-ordering principle.

AC often appears implicitly as an assumption in many proofs and can easily be overlooked. The proof of Theorem 3.1 later provides an example of the oversight of AC.

In the first 30 years after AC was formulated, whether or not AC could be taken as meaningful and acceptable was very controversial. However, AC is now widely accepted and used by the majority of mathematicians.

AC states:

The Axiom of Choice 1 Every family of non-empty sets has a choice function.

This choice function maps each set in the family to an element of the set. AC has many forms. We list a few below and it is not hard to see they are essentially the same.
The Axiom of Choice 2 For every set $x$, there exists a function

$$f : \mathcal{P}(x) \setminus \{\emptyset\} \rightarrow x$$

such that $f(s) \in s$.

The Axiom of Choice 3 The product of any family of non-empty sets is non-empty.

The Axiom of Choice 4 Every family $x$ of non-empty sets has a choice function on $x$ such that for all $s \in x$, $f(s)$ is a non-empty finite subset of $s$.

If the sets are finite or are ordered in some way, e.g. natural numbers, we do not need AC. A choice function is required if the sets are infinite and there is no clear rule to select elements from them. It is important to point out that the essence of AC is the existence of a choice function, not the choice that the function gives. Therefore, we do not require the knowledge about the choice function, and thus no construction of the choice function is needed; that is, the choice function used need not be well-defined.

In constructivism, all existence proofs are required to be totally explicit. That is, one must be able to construct, in an explicit and canonical manner, anything that is proven to exist. Due to its non-constructive nature, AC received strong criticisms, especially after Banach and Tarski proved in 1924 a paradox using AC. We will discuss this in section 6.

The criticisms did not stop until in 1938 Kurt Gödel proved the consistency of AC in ZF. This verified why we can use AC and displaced any suspicion that the use of AC could lead to contradiction.

In this paper, we first provide, as motivation, some equivalents of AC and results requiring AC from various branches of mathematics. Then we prove the consistency of AC in ZF.
2 Preliminaries

The essential idea on which the Axiom of Choice is based constitutes a general logical principle which ... is necessary and indispensable.

David Hilbert

We assume that the reader has some knowledge in the formal language of logic. See Kunen 1980 [12] for a quick review or Mendelson 1987 [15] for a formal treatment.

2.1 Zermelo–Fraekel Set Theory

We list, formulate, and command on all the axioms in ZF. Some familiar mathematics objects are defined using these axioms.

0. Axiom of Existence.

\[ \exists x (x = x). \]

This axiom assume a non-empty universe. It is listed here for emphasis but can be omitted since it is derivable from the logical axioms of formal logic [12]. (Thus it is not counted as one of ZF.) But what is a set and what is not a set? It is not easy to give definitions to them without self-reference. Although in ZF, classes do not exist, i.e. the only objects are sets [9], sometime it is unavoidable to work on classes.

Definition 2.1 (§4 [15]). A class is a collection \( \{ x : P(x) \} \), with some property \( P \). A class \( x \) is a set if it is a member of a class; it is a proper class if it is not a set.

We would like this definition to capture all the legitimate mathematical objects. For instance, we do not accept a \( P(x) \) that says \( x \) is a dog.

1. Axiom of Extensionality.

\[ \forall x \forall y \left( \forall u \in x (u \in y) \land \forall u \in y (y \in x) \implies x = y \right). \]

This axiom says a set is defined by its members.
2. Axiom of Foundation.

\[ \forall x (x \neq \emptyset \Rightarrow \exists u \in x (u \cap x = \emptyset)). \]

This axiom says any set non-empty set contains an \( \in \)-least element. It prevents \textit{ill-founded sets} from existing, e.g. a set \( x \) with \( x \in x \), and thus solve the famous Russell’s Paradox:

\[ \exists u \left( u = \{ x : x \notin x \} \right). \]

Since there exist no sets containing themselves, \( u \) contains all the sets and thus is not a set itself by the Axiom of Separation below.

3. Axiom of Pairing.

\[ \forall x \forall y \exists u (x \in z \land y \in z \land \forall u \in z (u = x \lor u = y)). \]

Given \( x \) and \( y \), define the pair \( u \) to be \( \{ x, y \} \). Easy to see \( u \) is unique by the Axiom of Extensionality. If \( y = x \), then \( u = \{ x, x \} = \{ x \} \), and there exists a set \( z \) with

\[ z = \{ x, u \} = \{ x, \{ x \} \}. \]

Thus we define the \textit{ordered pair} \( (x, y) \) of \( x \) and \( y \) to be \( \{ \{ x \}, \{ x, y \} \} \). Similarly, define \textit{tuples} with more than two entries to be

\[ (x, y, z) = ((x, y), z) = (x, (y, z)) \]
\[ (x, y, z, u) = ((x, y, z), u) \]
\[ \vdots \]

and so on.

4. Axiom of Union.

\[ \forall x \exists y \forall t (t \in y \leftrightarrow \exists s (s \in x \land t \in s)). \]

This axioms states for any set \( x \), the collection \( \bigcup x \) of all the elements in the members in \( x \) is a set. Using pairing we define

\[ x \cup y = \bigcup \{ x, y \} \]
\[ x \cup y \cup z = (x \cup y) \cup z \]
\[ \vdots \]

and so on.
5. Axiom of Power Set.

$\forall x \exists y \forall u (u \subset x \iff u \in y)$.

This axiom says for any set $x$, the collection of all the subset of $x$ is a set. Using ordered pairing and power set, we define the product $x \times y$ of $x$ and $y$ to be $\{(s, t) : s \in x \land t \in y\}$. Similarly, define

$$x \times y \times z = ((x \times y) \times z) = (x \times (y \times z))$$
$$x \times y \times z \times u = (x \times y \times z) \times u$$

and so on. We also called a set of ordered tuples an relation and define the domain of a relation $r$ to be

$$\text{dom}(r) = \{ x : \exists y ((x, y) \in r) \},$$

and the range of $r$ to be

$$\text{ran}(r) = \{ y : \exists x ((x, y) \in r) \}.$$

If $f$ is a relation such that

$$(x, y) \in f \land (x, z) \in f \Rightarrow y = z,$$

then it is a function. If $(x, y) \in f$, we write

$$f(x) = y.$$

For a function $f$, if $\text{dom}(f) = x$ and $\text{ran}(f) = y$, we write

$$f : x \to y,$$

and say $f$ maps $x$ to $y$. The image of a set $x$ under a function $f$, or simply the image of $f$, is

$$f[x] = \text{img}(f) = \{ t : \exists s \in x (t = f(s)) \}.$$

A function $f$ is one-to-one or injective if

$$(x, z) \in f \land (y, z) \in f \Rightarrow x = y;$$

it is onto or surjective if

$$\text{img}(f) = \text{ran}(f);$$

5
and it is *bijective* if it is both one-to-one and onto. The *pre-image* of \( t \in \text{ran}(f) \) under a function \( f \) is

\[
f^{-1}(t) = \{ s \in \text{dom}(f) : (s, t) \in f \}.
\]

If \( f \) is a one-to-one function, its inverse \( f^{-1} \) is also a function. By definition, \( \text{ran}(f) \) contains and could be strictly larger than the image of \( \text{dom}(f) \). The *restriction* \( f|_x \) of a function \( f \) to a set \( x \) is

\[
f|_x = \{(s, t) \in f : s \in x \}.
\]

Easy to see the terms defined above are sets.

6. **Axiom of Infinity.**

\[
\exists x (\emptyset \in x \land \forall s \in x (s \cup \{s\} \in x)).
\]

This axiom says there exists a set with infinitely many members. If a set \( x \) is finite, then there exists a natural number \( n \) such that \( x \) has exactly \( n \) members. Instead of formulating the property of infinity, in the formula of this axiom we formulate that of *inductivity*. In fact, the natural numbers is the smallest inductive set.

**Proposition 2.2.** Any inductive set contains the natural numbers.

**Proof.** Let \( x \) be an inductive set, i.e.

\[
\emptyset \in x \land \forall s \in x (s \cup \{s\} \in x).
\]

Suppose \( x \) does not contain all the natural numbers. Let \( n \) be the smallest natural number that is not contained in \( x \). Then

\[
\forall m < n (m \in x \land m \cup \{m\} \in x).
\]

Now since \( n \) is not zero, there exists a natural number \( l \) such \( l \cup \{l\} = n \). If \( l < n \), we are done. But clearly \( l \not\in n \).

7. **Axiom Schema of Replacement.** For each formula \( \varphi \),

\[
\forall x \forall y \forall z \left( \varphi(x, y) \land \varphi(x, z) \Rightarrow y = z \right) \Rightarrow \forall x \exists y \forall u \left( u \in y \Leftrightarrow \exists s \in x (\varphi(s, u)) \right).
\]

First note, since there are infinitely many formulas and thus infinitely many axioms of Replacement, this is an axiom schema; and similarly for the axioms of Separation. In short, this axiom schema says the image of a set under a function is a set.
8. Axiom Schema of Separation. For each formula $\varphi$,

$$\forall x \forall p_1 \ldots \forall p_n \exists y \forall s \left( s \in y \iff s \in x \land \varphi(s, p_1, \ldots, p_n) \right).$$

This schema states given a formula $\varphi$, a set $x$, and parameters $p = \{p_1, \ldots, p_n\}$, there exists a set $u = \{s \in x : \varphi(s, p)\}$; and such $u$ is unique by Extensionality. It is important to note that the statement that there exists a set $u = \{s : \varphi(s, p)\}$ is false. Then it follows that

$$\emptyset = \{x : x \neq x\}$$

is a set and that the universal class

$$\{x : x = x\}$$

is not.

2.2 Ordinals

Definition 2.3. A partial ordered set $(x, <)$ consists of a set $x$ and a binary relation $<$, called a partial ordering, on $x$, such that

i) $\forall s \in x(\neg(s < s))$;

ii) $\forall r, s, t \in x(r < s \land s < t \Rightarrow r < t)$.

Definition 2.4. Let $(x, <)$ be a partial ordered set, $\emptyset \neq t \subset x$, and $s \in x$. We say $s$ is

i) maximal in $t$ if $s \in t$ and $\neg(\exists u \in t(s < u))$;

ii) minimal in $t$ if $s \in t$ and $\neg(\exists u \in t(u < s))$;

iii) largest or greatest in $t$ if $s \in t$ and $\forall u \in t(u < s \lor u = s)$;

iv) least or lowest in $t$ if $s \in t$ and $\forall u \in t(s < u \lor s = u)$.

v) an upper bound of $t$ if $\forall u \in t(u < s \lor u = s)$;

vi) a lower bound of $t$ if $\forall u \in t(s < u \lor s = u)$.

Definition 2.5. A linear ordered set $(x, <)$ is a partial ordered set such that

$$\forall s, t \in x(s = t \lor s < t \lor t < s).$$

Remark 2.6. A partial ordered set can have more than one maximal elements but only one largest element, whereas in a linear ordered set, the maximal element and a largest one coincide; similar for minimal and least elements.
Remark 2.7. If we order all the elements in a linear ordered set $x$, they look like a chain. So we call a subset $z$ of a partial ordered set $(y, <)$ a chain if $(z, <_z)$ is a linear ordered set, where $<_z$ is the restriction of $<$ to $z$.

Notation 2.8. Let $x$ and $y$ be two sets. If there exists a one-to-one mapping from $x$ to $y$, we write $|x| \leq |y|$. If $|x| \leq |y|$ and $|y| \leq |x|$, we write $|x| = |y|$. If $|x| \leq |y|$ and $|y| \notin |x|$, we write $|x| < |y|$.

Definition 2.9. Two linearly ordered sets $(x, <_1)$ and $(y, <_2)$ are order isomorphic, denoted $(x, <_1) \simeq (y, <_2)$, if there exists a map $f$ from $x$ to $y$ such that $f$ is bijective and order-preserving, i.e.

$$\forall s, t \in x (s <_1 t \Rightarrow f(s) <_2 f(t)) \land \forall s, t \in y (s <_2 t \Rightarrow f^{-1}(s) <_1 f^{-1}(t)).$$

Definition 2.10. A well-ordered set $(x, <)$ is a linear ordered set such that every non-empty subset of $x$ has a $<$-least element. An initial segment $x(t)$ of $(x, <)$ is the set $\{s \in x : s < t\}$ for some $t \in x$.

Definition 2.11. A set $x$ is transitive if

$$\forall s \in x (s \in x \Rightarrow s \subset x).$$

Definition 2.12. A set $\alpha$ is an ordinal if and only if $(\alpha, \epsilon)$ is a transitive well-ordered set. We write lowercase Greek letters $\alpha$, $\beta$, ... to represent ordinals. Denote the class of ordinals by $\text{Ord}$. For $\alpha, \beta \in \text{Ord}$, define

$$\alpha < \beta \iff \alpha \in \beta.$$

It is easy to see the following properties of ordinals:

Lemma 2.13. Let $\alpha, \beta \in \text{Ord}$.

1. $\gamma \in \alpha \Rightarrow \gamma \in \text{Ord}$.
2. $\beta \neq \alpha \land \beta \subset \alpha \Rightarrow \beta \in \alpha$.
3. $\alpha \subset \beta \lor \beta \subset \alpha$.
4. $\alpha \cup \{\alpha\} \in \text{Ord}$.
5. $\beta < \alpha \iff \beta \in \alpha \iff \beta \subset \alpha$. 

8
6. $<$ linearly orders $\text{Ord}$.

7. $\alpha = \beta \iff (\alpha, <) \simeq (\beta, <)$. \hfill $\square$

**Definition 2.14.** An ordinal $\alpha$ is a *successor* if there exists an ordinal $\beta$ such that

$$\alpha = \beta \cup \{\beta\} = \beta + 1;$$

it is a *limit* if it is not a successor, that is, it has no last element.

**Remark 2.15.** If $\alpha$ is a limit, then either

i) $\alpha = 0$, or

i) $\alpha > 0$, and $\alpha = \bigcup_{\beta < \alpha} \beta$.

**Examples 2.16** (of ordinals).

- $0 = \emptyset$,
- $1 = \{\emptyset\}$,
- $2 = \{\emptyset, \{\emptyset\}\}$,
- $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$,
- $\vdots$
- $\omega = \{\text{the finite ordinals}\} = \{\text{the natural numbers}\}$,
- $\omega_1 = \{\text{the countable ordinals}\}$.

### 2.3 Transfinite Induction

In the light of proofs by induction on natural numbers, we extend them to their counterparts on ordinals.

**Theorem 2.17** (Proof by transfinite induction, Theorem 1.7.3 [19]).

Let $\{\varphi_s : s \in x\}$ be a set of mathematical expressions or statements labeled by a well-ordered set $(x, <)$. Suppose

$$\forall s \in x (\forall t \in x (t < s \Rightarrow \varphi_t) \Rightarrow \varphi_s),$$

then

$$\forall s \in x (\varphi_s).$$

**Proof.** Suppose, instead, $\forall s \in x (\varphi_s)$ is not true. Let $u$ be the least element in $x$ such that $\varphi_u$ fails. Then

$$\forall t \in x (t < u \Rightarrow \varphi_t).$$

But by the assumption,

$$\forall t \in x (t < u \Rightarrow \varphi_t) \Rightarrow \varphi_u.$$

Thus $\varphi_u$ is true. This contradicts the definition of $u$. \hfill $\square$
Transfinite induction can also be used to define objects.

**Theorem 2.18** (Definition by transfinite induction, Theorem 1.7.4 [19]). Let \( x \) be a set, \((y, <)\) be a well-ordered set, and define

\[
\Omega = \{ f : t \to x \mid \exists u \in y (t = y(u) = \{v \in y : v < u\}) \},
\]

Let \( G : \Omega \to x \) be a function. Then there exists a unique function \( F \) on \( y \) such that for each \( u \in y \),

\[
F(u) = G(F\restrictedto y(u)).
\]

**Remark 2.19.** In \( \Omega \), each \( f \) is a sequence (ordered tuples) with an ordinal length of elements in \( x \) labeled by the elements in \( y \). For the purpose of demonstration, say \( y \) is an ordinal. Then an \( f \) is in the form of

\[
\langle s_0, s_1, \ldots, s_n, \ldots, s_\alpha \rangle,
\]

where \( \langle \ldots \rangle = (\ldots) \) denotes a sequence, \( \alpha \in \text{Ord} \) and each \( s_i \in x \).

Intuitively, Theorem 2.18 works because we get a sequence \( F \) of elements in \( x \) from the given \( G \) by induction as follows:

- At \( \alpha = 0 \), since \( \emptyset \) is a sequence (of length 0), let \( F(0) = G(\emptyset) \), say, \( = s_0 \);
- at \( \alpha = 1 \), since \( \langle s_0 \rangle \) is a sequence (of length 1), let \( F(1) = G(\langle s_0 \rangle) \), say, \( = s_1 \);
- at \( \alpha = n \in \omega \), since \( \langle s_0, s_1, \ldots, s_{n-1} \rangle \) is a sequence (of length \( n \)), let \( F(n) = G(\langle s_0, s_1, \ldots, s_{n-1} \rangle) \); and so on.

So each time when we get an element from \( G \), we glue it to the end of the old sequence. This is exactly the familiar definition by induction. But Theorem 2.17 allows us to proceed to the limit cases.

**Proof of Theorem 2.18.** For each \( u \in y \), let \( \varphi_u \) be the statement that there exists a unique function \( f_u : u \to x \) such that

\[
\forall v < u (f_u(v) = G(f_u\restrictedto v)).
\]

Let \( w \in y \) such that for all \( u < w \), \( \varphi_u \) holds. (\( \varphi_\emptyset \) is obviously the trivial one that holds.) We have the unique set \( \Omega_w = \{ f_u : u < w \} \), where \( f_u \) is as in the statement \( \varphi_u \). It is easy to see that for each \( u < w \), if \( p < u \), \( f_u\restrictedto p \) also satisfies

\[
\forall v < p (f_u\restrictedto p(v) = G((f_u\restrictedto p)\restrictedto v)).
\]

Therefore by the uniqueness of \( \Omega_w \), if \( p < u < w \), then \( f_p = f_u\restrictedto p \). So \( h = \bigcup_{u<w} f_u \) is a common extension for \( \{ f_u : u < w \} \). Now if \( w \) is a limit, let \( f_w = h \); if \( w \) is a successor, let \( f_w\restrictedto y(w-1) = h \) and \( f_w(w-1) = G(h) \), where \( w-1 \) is the element whose successor is \( w \). Since \( \Omega_w \) is unique, so is \( f_w \). Then we have

\[
\forall u \in y (\forall v < y(v < u \Rightarrow \varphi_v) \Rightarrow \varphi_u).
\]
So by Theorem 2.17,

$$\forall u \in y(\varphi_u).$$

Similarly, one can see

$$H = \bigcup_{u \in y} f_u$$

is the unique common extension for \( \{ f_u : u \in y \} \). Again, if \( y \) is a limit, let \( F = H \); if it is a successor, let \( s \) be the largest element in \( y \) and take \( F|y(s-1) = H \) and \( F(s) = G(H) \). And similarly, \( F \) is unique.

Since induction is just a special case of transfinite induction, we collectively call induction and transfinite induction just induction hereafter.

**Theorem 2.20** (Theorem 2.12 [9]). Every well-ordered set is order isomorphic to a unique ordinal.

**Proof.** Let \((x, <)\) be a well-ordered set. Using Theorem 2.18, we define by induction

$$\Omega = \{ f : \alpha \to x \mid \alpha \in \text{Ord} \}.$$ 

Since \( \alpha \) and \( x \) are sets, the elements in \( \Omega \) are well-defined. \( \Omega \) is the set of the sequence of elements in \( x \) (labeled by ordinals); in particular, \( \Omega \) contains all the \( f_t \)'s with \( t \in x \) such that

i) the domain of \( f_t \) is an ordinal,

ii) the image of \( f_t \) is the initial segment \( x(t) = \{ s \in x : s < t \} \), and

iii) \( f_t \) is an order isomorphism.

Now let \( G : \Omega \to x \) be such that \( G \) maps an initial segment \( x(s) \) of \( x \) to \( s \), where \( s \in x \). Then by Theorem 2.18, we get a unique \( F \). And \( F \) is an is an order isomorphism from some ordinal \( \alpha \) to \( x \). This \( \alpha \) is also unique. \qed

**Remark 2.21.** From Theorem 2.20, we see that an important idea of ordinals is that they provide the canonical well-orderings for all sets (by assuming AC). Later we will see in section 3, every set can be well-ordered. On the other hand, since no set contains a copy (up to order isomorphim) of itself, Ord is a proper class. Using this fact, we prove the following result.

**Theorem 2.22** (Hartogs' Theorem, Proposition 4.28 [15]).

$$\forall x \exists \alpha \in \text{Ord} \forall y (y \subset x \Rightarrow |y| \neq |\alpha|).$$
Proof. Suppose $x$ is a set such that
\[ \forall \alpha \in \text{Ord} \exists y(y \subset x \land |y| = |\alpha|). \]
Now for each pair of $y$ and $\alpha$ such that $|y| = |\alpha|$, we let $f_{y,\alpha} : y \to \alpha$ be the bijection and define a relation $<_{y,\alpha}$ on $y$ by
\[ \forall s, t \in y (s <_{y,\alpha} t \iff f_{y,\alpha}(s) < f_{y,\alpha}(t)). \]
Thus $f_{y,\alpha} : (y, <_{y,\alpha}) \to (\alpha, <)$ is the natural extension of $f_{y,\alpha}$ and is an obvious order isomorphism. For each $f_{y,\alpha}$, since $y \subset x$ and $|y| = |\alpha|$, we have
\[ f_{y,\alpha} \subset x \times (x \times x) \times x. \]
Define a function $F$ mapping each ordinal $\alpha$ to the set
\[ u_\alpha = \{ f_{y,\beta} : \beta < \alpha \}. \]
As we see
\[ u_\alpha \subset \mathcal{P}(x) \times (\mathcal{P}(x) \times \mathcal{P}(x)) \times \mathcal{P}(x), \]
$u_\alpha$ is a set. So it follows that the image of $F$ is also a set since
\[ F[\text{Ord}] \subset \mathcal{P}\left(\mathcal{P}(x) \times (\mathcal{P}(x) \times \mathcal{P}(x)) \times \mathcal{P}(x)\right). \]
Now since $F$ is a one-to-one function, $F$ has an inverse $F^{-1}$. $F^{-1}$ maps $F[\text{Ord}]$ to Ord. So, by the Axiom of Replacement, Ord is a set. This contradicts the fact that Ord is a proper class. \qed

2.4 Cardinals

Cardinality is a familiar concept seen in all branches of mathematics. It counts the number of the elements of a set. Cardinality and cardinal are the same thing, but intuitively, one may think of cardinality as a property and cardinal as a number. Now we define cardinals formally using ordinals.

Definition 2.23 (§3 [9]). A cardinal is an ordinal $\kappa$ such that
\[ \forall \alpha \in \text{Ord} (\alpha < \kappa \Rightarrow |\alpha| < |\kappa|). \]
We say the cardinality of a set $x$ is $\kappa$ if $|x| = |\kappa|$.

Examples 2.24 (of cardinals).

1. All the natural numbers are the finite cardinals.
2. The infinite cardinals are called alephs, written $\aleph_\alpha$‘s. The ordinal $\omega$ is the first infinite cardinal, written $\aleph_0$. The first uncountably infinite cardinal is $\omega_1$, written $\aleph_1$.
3. The ordinal $\omega + 2$ is not a cardinal, since $\omega < \omega + 2$ and $|\omega + 2| = |\omega| = \aleph_0$. 

12
2.5 The universe of sets

**Definition 2.25** (Definitions § III 2.1, 2.2 [12]). We define by induction the universe \( V \) of sets,

i) \( V_0 = \emptyset \),

ii) \( V_\alpha = \begin{cases} \mathcal{P}(V_{\alpha-1}) & \text{if } \alpha \text{ is a successor} \\ \bigcup_{\beta<\alpha} V_\beta & \text{if } \alpha \text{ is a limit} \end{cases} \), and

iii) \( V = \bigcup_{\alpha \in \text{Ord}} V_\alpha \).

Since the Axiom of Foundation says there are no ill-founded sets, i.e. any (non-empty) set has an \( \in \)–minimal element, it follows that any (non-empty) class has the same property (see Lemma 6.2 [9]). Using this fact, we can show \( V \) contains all the (well-founded) sets. (Thus \( V \) is not a set itself but a proper class.)

**Lemma 2.26** (Lemma 6.3 [9]).

\[ \forall x (x \in V). \]

*Proof.* Let \( y = \{ x : x \notin V \} \). If \( y = \emptyset \), we are done; otherwise, let \( x \) be \( \in \)-minimal in \( y \). Then we have

\[ \forall s \in x \ (s \in V); \]

if not, \( x \) is not \( \in \)-minimal. So \( x \subset V \) and therefore there exists an obvious one-to-one map from \( x \) to \( V \). Then since \( x \) is a set, we use the Axiom of Replacement to get an ordinal \( \beta \) with \( x \subset \bigcup_{\alpha<\beta} V_\alpha \). Then

\[ x \subset V_\beta \]

\[ \Rightarrow x \in \mathcal{P}(V_\beta) = V_{\beta+1} \]

\[ \Rightarrow x \in V. \]

\[ \square \]

**Definition 2.27.** Let \( \text{rank}(x) = \inf\{ \beta \in \text{Ord} : x \in V_{\beta+1} \} \) be the *rank* for \( x \in V \).

It is not hard to check the properties of \( V \) and \( V_\alpha \)'s listed in the following lemma.

**Lemma 2.28.** Let \( \alpha \in \text{Ord} \).

1. \( V_\alpha \) is transitive (Lemma § III 2.3 (a) [12]).
2. \( \forall \gamma \in \text{Ord} (\gamma \subset \alpha \Rightarrow V_\gamma \subset V_\alpha) \) (Lemma § III 2.3 (b) [12]).
3. \( V_\alpha \cap \text{Ord} = \alpha \) (Lemma § III 2.7 (b) [12]).
4. \( V_\alpha = \{ x \in V : \text{rank}(x) < \alpha \} \) (Lemma § III 2.5 [12]).

\[ \square \]
3 Some equivalents of AC

The Axiom of Choice is obviously true, the Well-ordering Principle obviously false, and who can tell about Zorn’s lemma?

Jerry Bona

The quote at the beginning of this section is indeed a joke, as we will show the three things in the quote are equivalent.

AC has a large number of equivalents; they live in different branches of mathematics, including mathematical logic, set theory, order theory, algebra, general topology, functional analysis, and so on [17]. Here we only list and prove a few of them. For a more comprehensive review of AC’s equivalents, the interested reader is referred to Rubin & Rubin’s 1963 monograph [17].

We will show the following statements are equivalent:

1. AC (the Axiom of Choice).
2. WOP (Zermelo’s Well-ordering Principle): every set can be well-ordered.
3. Tri (Trichotomy): \( \forall x \forall y (|x| < |y| \lor |y| < |x| \lor |x| = |y|) \).
4. Zorn (Zorn’s lemma): if every chain in a non-empty partial ordered set \( (x, \prec) \) has an upper bound, then \( x \) has a maximal element.
5. Vec: Every vector space has a basis.
6. Tych (Tychonoff’s Theorem): A product of compact spaces is compact.

Before starting the proof of the equivalences, we would like to point out that since the use of AC often occurs in a very subtle way, sometimes it is easy to make mistakes in the proof of AC from an equivalent of it. The wrong proof below from a standard textbook is an example.
\textbf{Theorem 3.1.} Zorn ⇒ AC.

\textit{Proof.} This is a \textbf{wrong} proof adapted from §1.3 [19].

Let \( \{u_i : i \in I\} \) be a family of non-empty sets, for some index set \( I \). Let \( \mathcal{F} \) denote the set of all the choice functions for

\[ \{u_i : i \in J \wedge J \subset I\}. \]

We want to show the set \((\mathcal{F}, \subset)\) satisfies the assumption of Zorn’s lemma.

Let \( \{f_s : s \in S\} \) be a \( \subset \)–chain in \( \mathcal{F} \), for some \( S \subset I \). Let

\[ f = \bigcup_{s \in S} f_s \]

be the common extension for \( \{f_s : s \in S\} \). Clearly, \( \forall s \in S(f_s \subset f) \). So \( f \) is an upper bound of \( \{f_s : s \in S\} \) in \( \mathcal{F} \). So by Zorn’s lemma, \( \mathcal{F} \) has a maximal element, say \( g \).

It remains to show \( g \) is a choice functions for \( \{u_i : i \in I\} \). Suppose not. Since \( g \in \mathcal{F} \), \( g \) is a choice function for, say, \( \{u_i : i \in J\} \) for some \( J \subseteq I \). Let \( v \in I - J \) and \( b \in u_v \). Define \( h \supseteq g \) such that \( h(u_v) = b \). Then \( h \in \mathcal{F} \), contradicting the maximality of \( g \).

This proof of Zorn ⇒ AC by Srivastava [19] seems to be a routine exercise using Zorn’s lemma. However, when the author chooses some \( v \in I - J \) and \( b \in u_v \), he implicitly but indeed uses AC. Without assuming AC or any knowledge of \( I - J \) and \( u_v \), it could happen that we have no way to choose elements from \( I - J \) and \( u_v \).

Now we prove the equivalences listed above. We will prove the following relations:

\begin{center}
\begin{tikzcd}
\text{Tri} & & \\
WOP & AC & \text{Tych} \arrow[leftrightarrow]{u} \arrow[leftrightarrow]{l} \arrow[leftrightarrow]{r}
\end{tikzcd}
\end{center}

The proof of Tychonoff’s Theorem using ultrafilter, a maximal set, and the proof that every vector space has a basis using Zorn’s lemma, are omitted here since they can be found, respectively, in most standard textbooks for general topology (e.g. Theorem 7.4.2 [24]) and algebra (e.g. Theorem § 5.2 [7]).

\textbf{Proposition 3.2.} WOP ⇒ AC.

\textit{Proof.} Let the minimal element be the choice. \qed
Theorem 3.3 (Proposition 4.37 [15]). \( AC \Rightarrow Zorn \)

*Proof.* Let \((x, \prec)\) be a non-empty partial ordered set such that every \( \prec \)-chain in \(x\) has an upper bound in \(x\). We want to show \(x\) has a maximal element.

Suppose \(f\) is a choice function for \(x\). Let \(b = f(x)\). By induction we define a function \(F : \text{Ord} \rightarrow x\) by

\[
F(\alpha) = \begin{cases} 
  b & \text{if } \alpha = 0, \\
  f(u_\alpha) & \text{otherwise},
\end{cases}
\]

where

\[
u_\alpha = \{ s \in x : s \notin F[\alpha] \land \forall t \in F[\alpha] (t \prec s) \}.
\]

Note that

i) \(u_\alpha\) is the set of upper bounds of \(F[\alpha]\) in \(x\) and \(u_\alpha \cap F[\alpha] = \emptyset\); and

ii) \(F(\cdot)\) is the image of a point, whereas \(F[\cdot]\) is the image of a set.

Now there is an ordinal \(\alpha\) such that \(u_\alpha = \emptyset\); if not, by the argument similar to that in Theorem 2.22, \(\text{Ord}\) would be a set. Let \(\gamma\) be the least such ordinal and \(g = F[\gamma]\). Observe, for each \(\alpha > 0\), that \(g(\alpha)\) is an upper bound of \(g[\alpha]\) in \(x\), and that \(g(\alpha) \notin g[\alpha]\). So \(g\) is a one-to-one function. Then \(g[\gamma]\) is a \(\prec\)-chain in \(x\) and thus has an upper bound in \(x\), say \(p\). Since \(u_\gamma = \emptyset\), which states that there exists no upper bounds of \(g[\gamma]\) that are in \(x - g[\gamma]\), we see \(p \in g[\gamma]\) and is unique.

We finish the proof by claiming that \(p\) is \(\prec\)-maximal in \(x\). Suppose not. Then there exists an \(s \in x\) such that \(s \neq p\) and \(p \prec s\). Then \(s \notin F[\gamma]\) and \(\forall t \in F[\gamma] (t \prec s)\). So \(u_\gamma \neq \emptyset\), leading to contradiction.

\[\square\]

Lemma 3.4 (Proposition 4.37 [15]). \( Zorn \Rightarrow WOP \).

*Proof.* Let \(x\) be a set. Let \(\mathcal{F}\) be all the sequences with an ordinal length of elements in \(x\), as described in Remark 2.19, but with no repeated element of \(x\). \(\mathcal{F} \neq \emptyset\), since the empty sequence \(\emptyset \in \mathcal{F}\). By the argument similar to that in Theorem 2.22, \(\mathcal{F}\) is a set. Clearly, \(\mathcal{F}\) is partial ordered by \(\subset\) and the union of any \(\subset\)-chain in \(\mathcal{F}\) is an upper bound. So by Zorn’s lemma, there exists a maximal element in \(\mathcal{F}\), say \(f\). Clearly, \(f\) is a sequence consisting of all the elements in \(x\). Let \(\alpha \in \text{Ord}\) be the length of \(f\) and define a well-ordering for \(f\) similarly to Theorem 2.22.

\[\square\]

Proposition 3.5 (Proposition 4.37 [15]). \( WOP \Rightarrow \text{Tri.} \)

*Proof.* Follows Theorem 2.20 and Lemma 2.13.

\[\square\]
Proposition 3.6 (Proposition 4.37 [15]). Tri ⇒ WOP.

Proof. By Hartogs’ Theorem (2.22) and Trichotomy,
\[ \forall x \exists \alpha \in x (|x| < |\alpha|). \]
Then we can define a well-ordering for \( x \) by the argument similar to that in Theorem 2.22. \qed

Theorem 3.7 ([13]). Tych ⇒ AC.

Proof. Let \( \{x_i : i \in I\} \) be a family of non-empty sets, for some index set \( I \). We want to show the Form 3 of AC, i.e.
\[ \prod_{i \in I} x_i \neq \emptyset. \]
Without loss of generality, assume \( \{x_i : i \in I\} \) are pairwise disjoint.
Let \( \alpha \) be an ordinal such that \( \alpha \notin x_i \) for all \( i \in I \). Such \( \alpha \) exists by Hartogs’ Theorem. Now for each \( i \in I \), let
\[ y_i = x_i \cup \{\alpha\} \]
and equip it with the topology
\[ \{\emptyset, \{\alpha\}, x_i, y_i\}. \]
Each \( y_i \) is clearly compact. Let
\[ y = \prod_{i \in I} y_i \]
equipped with the product topology.
For each \( i \in I \), let
\[ \pi_i : y \to y_i \]
be the projection map and
\[ u_i = \pi_i^{-1}(\{\alpha\}). \]
Then \( u_i \)'s are open sets, since \( \pi_i \)'s are continuous and \( \{\alpha\} \) is a basic open set. Now for any finite \( J \subseteq I \),
\[ \bigcup_{j \in J} u_j \]
does not contain the element
\[ (\ldots, \{\alpha\}, \ldots), \]
where \( \{\alpha\} \) is the \( i \)-th entry with \( i \in I \setminus J \). It follows \( \bigcup_{j \in J} u_j \) does not cover \( y \) and so \( \{u_i : i \in I\} \) has no finite subcover of \( y \). Therefore \( \bigcup_{i \in I} u_i \) too does not cover \( y \); otherwise, it has a finite subcover of \( y \), since \( y \) is compact by Tychonoff’s Theorem. Now since all \( y_i \)'s are different, \( \prod_{i \in I} x_i \) must be non-empty, as required. \qed
**Theorem 3.8** (Theorem 2 [2]). $\text{Vec} \Rightarrow AC$.

*Proof.* Let $\{x_i : i \in \alpha\}$ be a family of non-empty sets, where $\alpha$ is an ordinal. We use $\{x_i\}$ to construct a vector space by the idea of the field of rational fractions, and then find a finite set $u_i \in x_i$ for each $i \in \alpha$. For the background of the field of rational fractions, one can consult any standard algebra textbook (e.g. p124 [7]). Without loss of generality, assume $\{x_i\}$ are pairwise disjoint. Let $x = \bigcup_{i \in \alpha} x_i$.

For $s \in x$, let $\mathbb{R}[s]$ and $\mathbb{R}(s)$ be the field of polynomials and that of rational fractions, respectively, with variable $s$ and coefficients in $\mathbb{R}$. Similarly let $\mathbb{R}[x]$ and $\mathbb{R}(\langle x \rangle)$ denote the field of polynomials and that of rational fractions, respectively, with variables in $x$ and coefficients in $\mathbb{R}$.

For an $f \in \mathbb{R}[x]$, we say $f$ has $[i]$–degree $n$ if the sum of the exponents of elements in $x_i$ in each term of $f$ is $n$. For an $h \in \mathbb{R}(\langle x \rangle)$, where $h = f/g$ with $f, g \in \mathbb{R}[x]$, we say $h$ has $(i)$–degree $n$ if

$$[i]$$–degree of $f = n + [i]$–degree of $g.$

For examples, if $s, t \in x_1$ and $p, q \in x_2$, then

$$f = 2sp^{-2} + \frac{1}{2}t^2p^3q^{10} - t$$

has $[1]$–degree 4 and $[2]$–degree 11, and

$$h = \frac{f}{t^2q^{15}}$$

has $(1)$–degree 2 and $(2)$–degree $-4$.

Now let

$$k = \{h \in \mathbb{R}(\langle x \rangle) : h \text{ has } (i)\text{–degree } 0, i \in \alpha\}.$$ 

It is a routine to check that $k$ is a subfield of $\mathbb{R}[x]$. Then $\mathbb{R}[x]$ is a vector space over $k$. So by our assumption, the subspace $y$ spanned by $x$ has a basis, say $B$. Now for $i \in \alpha$, $s \in x_i$, and $s \neq 0$, by the definition of a basis, we can write

$$s = \sum_{b \in B_s} a(s, b) \cdot b,$$

where $B_s \subset B$ is finite, and $a(s, b) \in k$ is non-zero and depends on $s$ and $b$. Now if $t \in x_i$, we have

$$t = \sum_{b \in B_t} a(t, b) \cdot b = \sum_{b \in B_s} a(s, b) \frac{t}{s} \cdot b.$$ 

Since $a(t, b)$ and $a(s, b)t/s$ are just coefficients, we see $B_s = B_t$ and the $s$ in $a(s, b)/s$ could be any element in $x_i$. Since $a(s, b) \in k$ so has $(i)$–degree 0;
therefore $a(s, b)/s$ has $(i)$-degree $-1$. Then for $i \in \alpha$, $s \in x_i$, and $s \neq 0$, we can write

$$s = \sum_{b \in B_i} a(b) \cdot b \cdot s,$$

where $B_i \subset B$ is finite and only depends on $i$, and each $a(b)$ has $(i)$-degree $-1$. $B_i$ is finite, so there are finitely many $a(b)$’s. Then we finish the proof by letting $u_i$ be the finitely many elements in $x_i$ appearing in the denominators of $a(b)$’s. □
4 Some results requiring AC

Mathematics would not be an international science if its theorems did not possess an objective content independent of the language in which we express them.

Ernst Zermelo

In this chapter we list some results requiring AC in various branches of mathematics. We omit the proofs since they are all classic results and can all be found in standard textbooks; we provide references instead. For a more comprehensive list of the applications of AC, see Appendix 2 of Moore 1982 [16] and Chapter 2 of Jech 1973 [8].

4.1 Set theory

Theorem 4.1 (Schröder–Bernstein Theorem, Proposition 4.21 (d) [15], Theorem 1.2.3 [19]).

\[ |x| \leq |y| \land |y| \leq |x| \Rightarrow |x| = |y|. \]

\[ \square \]

Theorem 4.2 (Lemma §I 10.21 [12]). Let \( \kappa \) be a cardinal. Then

\[ \kappa \geq \omega \land \forall \alpha \leq \kappa (|x_\alpha| \leq \kappa) \Rightarrow \left| \bigcup_{\alpha \leq \kappa} x_\alpha \right| \leq \kappa. \]

\[ \square \]

Corollary 4.3. Any union of finitely or countably many countable sets is itself countable.

\[ \square \]

4.2 General topology

Theorem 4.4 (Theorem 11.3.7 [24]). A uniform space is compact if and only if it is complete and totally bounded.

\[ \square \]
4.3 Algebra

Theorem 4.5 (Theorem §IV 4.4 [7]). Every field has an algebraic closure.

4.4 Group theory

Theorem 4.6 (Nielsen–Schreier Theorem, 1.5 [10]). Every subgroup of a free group is free.

4.5 Measure theory

Theorem 4.7 (Theorem 3.23 [23]). Let \( \{x_n : n \in \omega\} \) be countable disjoint measurable sets and \( \mu \) be the Lebesgue measure. Then

\[
\mu\left(\bigcup_{n \in \omega} x_n\right) = \sum_{n \in \omega} \mu(x_n).
\]

Theorem 4.8 (Vitali Theorem, 3.38 [23]). There exist set that are not Lebesgue measurable.

4.6 Functional Analysis

Theorem 4.9 (Hahn–Banach Theorem, 12.4.1 [24]). Let \( x \) be a normed space and \( y \subset x \) a vector subspace. Then every bounded linear functional on \( y \) can be extended to a linear functional on \( x \) with the same norm.

Theorem 4.10 (Lemma 3.20 [18]). Every Hilbert space has an orthonormal basis.
5 The consistency of AC in ZF

With the increasing recognition that there are questions in mathematics which cannot be decided without this axiom [of choice], the resistance to it must increasingly disappear.

Ernst Steinitz

5.1 Overview of the consistency and its proof

The proof of the consistency of AC in ZF, denoted \( \text{Con}(\text{ZF} + \text{AC}) \) or \( \text{Con}(\text{ZFC}) \), was first outlined in Kurt Gödel’s 1938 paper [5] and was fully given in his 1940 monograph [6].

To prove \( \text{Con}(\text{ZFC}) \), we in fact prove \( \text{Con}(\text{ZF}) \Rightarrow \text{Con}(\text{ZFC}) \).

By the Gödel Second Incompleteness Theorem, ZF cannot prove its own consistency. So we take \( \text{Con}(\text{ZF}) \) as an assumption.

By the Gödel Completeness Theorem, if ZFC is consistent, it has a model. It follows that if ZFC has a model, it is consistent. Therefore, while \( \text{Con}(\text{ZFC}) \) can be proved within model theory [9], the unusual approach people take is to find a model for ZFC.

We adopt Gödel’s approach.

5.2 A quick review of model theory

**Definition 5.1** (Definition 1.1.1 [14]). A *language* \( \mathcal{L} = \{ F, R, C \} \) consists a set \( F \) of functions, a set \( R \) of relations, and a set \( C \) of constants.

**Definition 5.2** (Definition 1.1.2 [14]). A *model* \( \mathcal{M} \) for a language \( \mathcal{L} \) is a pair \( (M, \mathcal{I}) \), where \( M \) is a set, called the universe of \( \mathcal{M} \), and \( \mathcal{I} \) is the interpretation of \( \mathcal{L} \) in the universe \( M \). We write \( \mathcal{M} = (M, F^\mathcal{L}, R^\mathcal{L}, C^\mathcal{L}) \).
Examples 5.3.

1. The language for set theory is just \( \{\in\} \). \( \{\bigvee_\alpha, \in\} \) is a model for \( \{\in\} \), for \( \alpha \in \text{Ord} \).

2. The language for additive groups is \( \{+ , e\} \), where + is a binary function and e is constant. A set \( g \) is a group if it satisfies the axioms

\[
\forall x(e + x = x + e = x), \\
\forall x \forall y \forall z(x + (y + z) = (x + y) + z), \\
\forall x \exists y(x + y = y + x = e).
\]

A group \( g \) is Abelian if it satisfies the additional axiom

\[
\forall x \exists y(x + y = y + x).
\]

The set \( \mathcal{F} \) of all the formulas in set theory is an Abelian group with the following group structure:

i) let \( e \) be \( \emptyset \), the empty formula;
ii) let \( \varphi \land \psi \) be \( \varphi + \psi \);
iii) let \( \neg \varphi \) be the inverse of \( \varphi \).

In addition, one can define multiplicative groups in a similar fashion.

3. The language for ring theory is \( \{+ , - , \cdot , 0 , 1\} \), where +, −, and \( \cdot \) are binary functions and 0 and 1 are constants. An additive Abelian group \( R \) is a ring if it satisfies axioms

\[
\forall x(x \cdot 0 = 0 \land x \cdot 1 = 1 \cdot x = x), \\
\forall x \forall y \forall z \left( x \cdot (y \cdot z) = (x \cdot y) \cdot z \right) \\
\quad \land x - y = z \iff x = y + z \\
\quad \land x \cdot (y + z) = (x \cdot y) + (x \cdot z) \\
\quad \land x \cdot (y + z) = (x \cdot y) + (x \cdot z).
\]

If \( x \) is a set, define

i) \( \mathcal{P}(x) = \{s \subset x\} \), the power set of \( x \);
ii) \( 0^e = \emptyset \);
iii) \( 1^e = x \);
iv) for \( s , t \subset x \), \( s + t = (s \cup t) - (s \cap t) \);
v) \( s = -s \);
vi) $s \cdot t = s \cap t$.

Then $(\mathcal{P}(x), +, -, \cap, \emptyset, x)$ is a model for ring theory.

**Definition 5.4** (Definition 1.1.4 [14]). Given $\mathcal{L} = \{F, R, C\}$, the set of $\mathcal{L}$-terms is the smallest set $\mathcal{T}$ such that

i) $C \subset \mathcal{T}$

ii) $\{v_i \text{ is a variable} : i \in \mathbb{N}\} \subset \mathcal{T}$

iii) if $x_1, \ldots, x_n \in \mathcal{T}$ and $f \in F$, then $f(x_1, \ldots, x_n) \in \mathcal{T}$.

**Definition 5.5** (Definition 1.1.5 [14]). Given $\mathcal{L} = \{F, R, C\}$, if $r \in R$ and $t_1, \ldots, t_n \in \mathcal{T}$, then $\varphi$ is an atomic $\mathcal{L}$-formula if it is in the form of $t_1 = t_2$ or $r(t_1, \ldots, t_n)$. The set of $\mathcal{L}$-formulas is the closure of atomic $\mathcal{L}$-formulas under the $\neg$, $\wedge$, $\vee$, $\exists$, and $\forall$ operations.

A variable $v$ in a formula $\varphi(\ldots, v, \ldots)$ is *bounded* if it is inside a quantifier $\exists v$ or $\forall v$; otherwise, it is *free*. A quantifier $\exists v$ or $\forall v$ in $\varphi(\ldots, v, \ldots)$ is *restricted* if $v$ has a domain $x$, i.e. $\exists v \in x$ or $\forall v \in x$; otherwise, it is *unrestricted*. A formula is a *sentence* if it has no free variables.

A sentence in a model is always true or false. A formula with free variables shows a property of the free variables rather than anything true or false. For example, $\exists x \in \mathbb{R}(x > 0)$ is a true sentence while $x > 0$ shows a property of $x$.

**Definition 5.6** (Definition 1.1.6 [14]). By induction on the length of a term $t \in \mathcal{T}$ of a model $\mathcal{M}$, we define its *value* $t^\mathcal{M}(a)$ over a finite sequence $a = (a_1, \ldots, a_n)$ of $\mathcal{M}$ as follows.

i) if $t$ is the constant $c$, then $t^\mathcal{M}(a) = c^\mathcal{M}$;

ii) if $t$ is the variable $v_i$, then $t^\mathcal{M}(a) = a_i$;

iii) if $t$ is the term $f(t_1, \ldots, t_m)$, where $f \in F$ is a function and $t_1, \ldots, t_m \in \mathcal{T}$ are terms, then $t^\mathcal{M}(a) = f^\mathcal{M}(t_1^\mathcal{M}(a), \ldots, t_m^\mathcal{M}(a))$. 

24
**Definition 5.7.** By induction on the complexity of a formula $\varphi$ of a model $M$, we define its satisfaction $\varphi^M$ over a finite sequence $a = (a_1, \ldots, a_n)$ of $M$ as follows:

\[
\begin{align*}
(t = s)^M & \text{ iff } t^M(a) = s^M(a) \\
(r(t_1, \ldots, t_m))^M & \text{ iff } (t_{1}^M(a), \ldots, t_{m}^M(a)) \in r^M \\
(\neg \psi)^M & \text{ iff } \neg(\psi^M) \\
(\psi \land \phi)^M & \text{ iff } \psi^M \land \phi^M \\
(\psi \lor \phi)^M & \text{ iff } \psi^M \lor \phi^M \\
(\exists x(\psi))^M & \text{ iff } \exists b \in M(\psi(a_1, \ldots, a_n, b))^M \\
(\forall x(\psi))^M & \text{ iff } \forall b \in M(\psi(a_1, \ldots, a_n, b))^M
\end{align*}
\]

where $t, s \in T$, $r \in R$, and $\psi$ and $\phi$ are formulas.

If $\varphi^M$, we say $\varphi$ holds or is true in $M$. We also write $M \models \varphi(a)$ or $\models_M \varphi(a)$.

In the rest of this paper, we restrict our interest to $\{\in\}$, the language of set theory.

**Definition 5.8** (Definition 12.6 [9]). Let $M = (M, \in)$ be a model for the language of set theory, where $M$ is a class, and let $\varphi(v_1, \ldots, v_n)$ be a formula for it. We now define $\varphi^M$, the Relativization or the counterpart of satisfaction for classes of $\varphi$ to $M$, by induction on the complexity of $\varphi$ as follows:

\[
\begin{align*}
(v_i \in v_{i_2})^M & \text{ iff } v_i \in v_{i_2} \\
(v_i = v_{i_2})^M & \text{ iff } v_i = v_{i_2} \\
(\neg \psi)^M & \text{ iff } \neg(\psi^M) \\
(\psi \land \phi)^M & \text{ iff } \psi^M \land \phi^M \\
(\psi \lor \phi)^M & \text{ iff } \psi^M \lor \phi^M \\
(\exists x(\psi))^M & \text{ iff } \exists b \in M(\psi^M) \\
(\forall x(\psi))^M & \text{ iff } \forall b \in M(\psi^M)
\end{align*}
\]

where $v_i$ and $v_{i_2}$ are variables, and $\psi$ and $\phi$ are formulas.

Since there is no confusion, we write $M$ instead of $M$. So similarly, we can write $M \models \varphi$ or $\models_M \varphi$.
5.3 Absoluteness

Definition 5.9 (Definition §IV 3.5 [12]). A \( \Delta_0 \)-formula is a formula \( \varphi \) (in ZF) in the form of

i) an atomic formula, or

ii) \( \neg \psi, \psi \land \phi, \psi \lor \phi, \psi \Rightarrow \phi, \) or \( \psi \Leftrightarrow \phi, \) or

iii) \( \exists x \in y(\psi) \) or \( \forall x \in y(\psi), \)

where \( \psi \) and \( \phi \) are \( \Delta_0 \)-formula. We say \( \varphi \) is \( \Delta_0 \).

First we notice that a \( \Delta_0 \)-formula has no unrestricted quantifiers. Then given a formula \( \varphi \) with unrestricted quantifiers, say \( \exists x(\psi) \) with \( \psi \Delta_0, \)

\[ (\exists x(\psi))^M = \exists x \in M(\psi^M) \]

is \( \Delta_0 \). So we have proved

Proposition 5.10. The relativization of a formula is \( \Delta_0 \). \qed

Definition 5.11. A formula is a \( \Sigma_1 \)-formula or \( \Pi_1 \)-formula if it can be written, respectively, in the form of \( \exists x(\psi) \) or \( \forall x(\psi), \) where \( \psi \) is a \( \Delta_0 \)-formula. A formula is a \( \Delta_1 \)-formula if it is both \( \Sigma_1 \) and \( \Pi_1 \).

Proposition 5.12. A \( \Delta_0 \)-formula \( \varphi \) is \( \Delta_1 \).

Proof.

\[ \varphi \Leftrightarrow \exists x(x \in \omega \land \varphi) \]

and

\[ \varphi \Leftrightarrow \forall x(x \in \omega \Rightarrow x \in \mathbb{R} \land \varphi). \]

Lemma 5.13 (Lemma 13.10 [9]). Let \( \varphi(v, \ldots) \) and \( \psi(v, \ldots) \) be two \( \Delta_0 \)-formulas. Then

1. the following formulas are \( \Sigma_1 \):

\[ \neg \forall x(\varphi(v, x, \ldots)), \]

\[ \exists x_1 \ldots \exists x_n(\varphi(x_1, \ldots, x_n, v, \ldots)), \]

\[ \exists x(\varphi(v, x, \ldots)) \land \exists y(\psi(v, x, \ldots)), \text{ and} \]

\[ \exists x(\varphi(v, x, \ldots)) \lor \exists y(\psi(v, x, \ldots)); \]
2. the following formulas are \( \Pi_1 \):

\[
\neg \exists x (\varphi(v, x, \ldots)), \\
\forall x_1 \ldots \forall x_n (\varphi(x_1, \ldots, x_n, v, \ldots)), \\
\forall x (\varphi(v, x, \ldots)) \land \forall y (\psi(v, x, \ldots)), \text{ and} \\
\forall x (\varphi(v, x, \ldots)) \lor \forall y (\psi(v, x, \ldots)).
\]

Proof.

1. 

\[\neg \forall x (\varphi) \iff \exists x (\neg \varphi),\]

\[\exists x_1 \ldots \exists x_n (\varphi) \iff \exists s \exists t \in s \exists x_1 \in t \ldots \exists x_n \in t (s = (x_1, \ldots, x_n) \land \varphi),\]

\[\exists x (\varphi) \land \exists y (\psi) \iff \exists x \exists y (\varphi \land \psi),\]

\[\exists x (\varphi) \lor \exists y (\psi) \iff \exists x \exists y (\varphi \lor \psi).\]

2. 

\[\neg \exists x (\varphi) \iff \forall x (\neg \varphi),\]

\[\forall x_1 \ldots \forall x_n (\varphi) \iff \forall x_1 \ldots \forall x_{n-1} (\neg \exists x_n (\neg \varphi))\]

\[\quad \iff \forall x_1 \ldots \forall x_{n-2} (\neg \exists x_{n-1} \exists x_n (\neg \varphi))\]

\[\vdots\]

\[\iff \neg \exists x_1 \ldots \exists x_n (\neg \varphi)\]

\[\iff \forall s (\exists t \in s \exists x_1 \in t \ldots \exists x_n \in t (s = (x_1, \ldots, x_n) \land \neg \varphi)\]

\[\iff \forall s (\exists t \in s \exists x_1 \in t \ldots \exists x_n \in t (s = (x_1, \ldots, x_n) \land \neg \varphi)),\]

\[\forall x (\varphi) \land \forall y (\psi) \iff \forall x \forall y (\varphi \land \psi),\]

\[\forall x (\varphi) \lor \forall y (\psi) \iff \forall x \forall y (\varphi \lor \psi).\]

We will work in transitive models and show \( \text{Con}(\text{ZFC}) \). The relation between a \( \Delta_0 \)-formula \( \varphi \) and a transitive model \( M \), to which we wish to relativize \( \varphi \), will be essential in our proof, as we will see in lemma 5.18. Since all the quantifiers in \( \varphi \)
are unrestricted, so each bounded variable has a domain sitting in $M$. Therefore, thanks to the transitivity of $M$, all the variables in $\varphi^M$ also sit in $M$.

$\Delta_0$-formulas are defined purely by logic symbols. The following two lemmas show that some familiar operations or expressions are $\Delta_0$.

**Definition 5.14.** An expression is a $\Delta_0$-expression or simply $\Delta_0$ if it can be written as a $\Delta_0$-formula. An operation $T$ is a $\Delta_0$-operation or simply $\Delta_0$ if $z = T(\ldots)$ is a $\Delta_0$-expression. Similarly we define $\Sigma_1$, $\Pi_1$, and $\Delta_1$-expressions and -operations.

**Lemma 5.15** (Lemma 12.10 [9], Theorem §IV 3.9 [12]). $\{\cdot, \cdot\}$, $(\cdot, \cdot)$, $\subset$, $\times$, $-$, $\cap$, $\cup$, $\bigcap$, $\bigcup$, $\text{dom}$, $\text{ran}$, $f(\cdot)$, and $f\upharpoonright$ are $\Delta_0$-operations, where $f$ is a function.

**Proof.** Let $x$, $y$ be sets. Since $x = y$ iff $x \subset y \land y \subset x$, then we have

- $z = \{x, y\}$ iff $x \in z \land y \in z \land \forall u \in z(u = x \lor u = y),$
- $z = (x, y)$ iff $\exists u \in z(u = \{x, x\}) \land \exists u \in z(u = \{x, y\}) \land \forall u \in z(u = \{x, x\} \lor u = \{x, y\}),$
- $x \subset z$ iff $\forall u \in x(u \in z),$
- $z = x \times y$ iff $\forall u \in z \exists s \in x \exists t \in y(u = (s, t)) \land \forall s \in x \forall t \in y \exists u \in z(u = (s, t)),$
- $z = x - y$ iff $z \subset x \land \neg(\exists u \in z(u \in y)) \land \forall u \in x(u \in y \lor u \in z),$
- $z = x \cap y$ iff $\forall u \in z(u \in x \land u \in y) \land \forall u \in x(u \in y \iff u \in z) \land \forall u \in y(u \in x \iff u \in z),$
- $z = x \cup y$ iff $\forall u \in z(u \in x \lor u \in y) \land x \subset z \land y \subset z,$
- $z = \bigcap x$ iff $\forall u \in z \forall s \in x(u \in s) \land \forall s \in x(z \subset s),$
- $z = \bigcup x$ iff $\forall u \in z \exists s \in x(u \in s) \land \forall s \in x(s \subset z),$
- $z = \text{dom}(x)$ iff $\forall u \in z \exists s \in x \forall t \in s(u \in t) \land \forall s \in x \forall t \in s \forall u, v \in t(s = (u, v) \Rightarrow u \in z),$

28
\[ z = \text{ran}(x) \text{ iff } \forall u \in z \exists s, t \in x \exists p \in t \left( p = \bigcap t \land \exists v \in p(s = (v, u)) \right) \]
\[ \land \forall s \in x \exists t \in x \forall p \in t \exists u \in z \left( p = \bigcap t \Rightarrow \exists v \in p(s = (v, u)) \right), \]
\[ z = f(x) \text{ iff } \exists u \in f(u = (x, z)), \text{ and} \]
\[ g = f \upharpoonright x \text{ iff } g \subset f \land \text{dom}(g) = \text{dom}(f) \cap x. \]

\[ \square \]

**Lemma 5.16** (Lemma 12.10 [9], Theorems §IV 3.11, 5.1 [12]). *The following expressions are $\Delta_0$: $x = \emptyset = 0$, $x$ is transitive, $\alpha \in \text{Ord}$, $\alpha$ is a limit ordinal, $z = \omega$, $n$ is a natural number, $r$ is a relation, and $f$ is a function.*

**Proof.**

\[ x = \emptyset = 0 \text{ iff } \neg(\exists u \in z(u = u)), \]
\[ x \text{ is transitive iff } \forall s \in x(s \subset x), \]
\[ \alpha \in \text{Ord} \text{ iff } \alpha \text{ is transitive } \land \forall s, t \in x(s \in t \lor s = t \lor t \in s) \]
\[ \land \forall s, t, u \in s(s \in t \land t \in u \Rightarrow s \in u), \]
\[ \alpha \text{ is a limit ordinal iff } \alpha \in \text{Ord} \land \forall s \in x \exists t \in x(s \in t), \]
\[ z = \omega \text{ iff } \neg(z = 0) \land z \text{ is a limit ordinal } \land \forall u \in z(\neg(u \text{ is a limit ordinal})), \]
\[ n \text{ is a natural number iff } n \in \omega, \]
\[ r \text{ is a relation iff } \forall u \in r \exists s \in \bigcup u \exists t \in \bigcup u(u = (s, t)), \text{ and} \]
\[ f \text{ is a function iff } f \text{ is a relation} \]
\[ \land \forall s, t, u \in \bigcup \bigcup f((s, t) \in f \land (s, u) \in f \Rightarrow t = u). \]

\[ \square \]
Definition 5.17. A formula \( \varphi \) is absolute for a model \( M \) if for all \( v_1, \ldots, v_n \),
\[
\varphi^M(v_1, \ldots, v_n) \Leftrightarrow \varphi(v_1, \ldots, v_n).
\]

Lemma 5.18 (Lemma 12.9 [9]). A \( \Delta_0 \)-formula is absolute for a transitive \( M \).

Proof. We prove by induction on the complexity of a formula \( \varphi \). It is straightforward if \( \varphi \) is atomic, \( \neg \psi, \psi \land \phi, \psi \lor \phi, \psi \Rightarrow \phi \), or \( \psi \Leftrightarrow \phi \), assuming the result holds for \( \psi \) and \( \phi \).

Now assuming \( \psi = \psi(y, \ldots) \) is a \( \Delta_0 \)-formula, if \( \varphi \) is \( \exists x \in y(\psi) \), then
\[
\exists x \in y(\psi) \\
\Leftrightarrow \exists x (x \in y \land \psi) \\
\Leftrightarrow \exists x (x \in y \land \psi^M) \\
\Leftrightarrow \exists x (x \in y \land x \in M \land \psi^M) \\
\Leftrightarrow (\exists x (x \in y \land \psi))^M \\
\Leftrightarrow (\exists x \in y(\psi))^M.
\]
\( \forall x \in y(\psi) \) is similar.

Lemma 5.19. A \( \Delta_1 \)-formula \( \varphi \) is absolute for a transitive \( M \).

Proof. First write \( \varphi \) in the form of \( \exists x(\psi) \), where \( \psi \) is a \( \Delta_0 \)-formula. Then
\[
\varphi^M \Leftrightarrow (\exists x(\psi))^M \\
\Leftrightarrow \exists x \in M(\psi)^M \\
\Leftrightarrow \exists x \in M(\psi) \\
\Rightarrow \exists x(\psi) \\
\Leftrightarrow \varphi.
\]
For the other direction, write \( \varphi \) in the form of \( \forall x(\psi) \), where \( \psi \) is a \( \Delta_0 \)-formula. Then
\[
\varphi \Leftrightarrow \forall x(\psi) \\
\Rightarrow \forall x \in M(\psi) \\
\Leftrightarrow \forall x \in M(\psi)^M \\
\Leftrightarrow (\forall x(\psi))^M \\
\Leftrightarrow \varphi^M.
\]
5.4 Constructible sets

Definition 5.20. A subset $x$ of a set $M$ (as an $\{\in\}$-model) is *definable* over $M$ if

$$x = \{s \in M : \varphi^M(s, a)\},$$

where $a = (a_1, \ldots, a_n)$ is a finite sequence over $M$.

Intuitively, a definable set $X$ is uniquely defined by some property. But its careful definition by a formula in the formal language avoids the following false paradox:

Define the first undefinable ordinal $\lambda$ by the property that $\lambda$ is a undefinable ordinal and $\forall \alpha < \lambda (\alpha$ is definable). Then we have defined an undefinable ordinal.

Notation 5.21. Let

$$\mathcal{D}(M) = \{X \subset M : X$ is definable over $M\}.$$ 

Definition 5.22 (Definitions §VI 1.4, 1.5 [12]). We now define by induction the *universe $\mathbb{L}$ of constructible sets*, in which we will show Con(ZFC):

i) $\mathbb{L}_0 = \emptyset$,

ii) $\mathbb{L}_\alpha = \begin{cases} \mathcal{D}(\mathbb{L}_{\alpha-1}) & \text{if } \alpha \text{ is a successor} \\ \bigcup_{\beta < \alpha} \mathbb{L}_\beta & \text{if } \alpha \text{ is a limit} \end{cases}$, and

iii) $\mathbb{L} = \bigcup_{\alpha \in \text{Ord}} \mathbb{L}_\alpha$.

Remark 5.23. Being constructible and definable are *not* the same thing. A set $x$ is constructible if and only if $x \in \mathbb{L}$. whereas definability can be only talked about in the context of a model.

Definition 5.24. The *Axiom of Constructibility* states $V = \mathbb{L}$, i.e.

$$\forall x \exists \alpha \in \text{Ord}(x \in \mathbb{L}_\alpha).$$

Our next task is to show $\mathbb{L}$ is a model for $ZF + V = \mathbb{L}$. After that, the job becomes straightforward: just to construct a well-ordering for all the constructible sets.

Think carefully the Axiom of Constructibility. It is trivial that any set in $\mathbb{L}$ is constructible. But as we have seen the definitions of $V$ and $\mathbb{L}$, which seem (intuitively) very different, it is unlikely that any arbitrary set is constructible. However, it is indeed a fact that $ZF + V = \mathbb{L}$ is consistent (if we can show $\mathbb{L}$ is model of them). To show this, we need to show “$x$ is constructible” is absolute for $\mathbb{L}$, i.e. $(x$ is constructible)$^\mathbb{L} \iff x$ is constructible. First we see some properties of $\mathbb{L}$ and $\mathbb{L}_\alpha$. 

31
Lemma 5.25 (Lemma §VI 1.3 [12]). Let $M$ be a transitive set.

1. If $x$ and $y$ are definable over $M$, so are $M - x$, $x \cap y$, and $x \cup y$.
2. $M \subset \mathcal{P}(M) \subset \mathcal{P}(M)$.
3. If $x \subset M$ is finite, then $x$ is definable.
4. (AC) If $M$ is infinite, then $|\mathcal{P}(M)| = |M|$. 

Proof. 

1. Let $x = \{s \in M : \psi^M(s, a)\}$ and $y = \{s \in M : \phi^M(s, b)\}$. Then
   $M - x = \{s \in M : \neg(\psi^M(s, a))\}$
   $x \cap y = \{s \in M : \psi^M(s, a) \land \phi^M(s, b)\}$, and
   $x \cup y = \{s \in M : \psi^M(s, a) \lor \phi^M(s, b)\}$. 

2. Let $y \in M$. Since $M$ is transitive, $y = \{x \in y\} = \{x \in M : x \in y\}$ and thus is definable. $\mathcal{P}(M) \subset \mathcal{P}(M)$ follows the definition of $\mathcal{P}(M)$. 

3. Follows from 1 and 2. 

4. By Theorem 4.2, $\forall n \in \omega(|M^n| = |M|)$. $|\mathcal{P}(M)| \geq |M|$ follows from $M \subset \mathcal{P}(M)$. 

Definition 5.26. Let $\rho(x) = \inf \{\beta \in \text{Ord} : x \in \mathbb{L}_{\beta + 1}\}$ be the $\mathbb{L}$-rank for $x \in \mathbb{L}$. 

Some properties of $\mathbb{L}$ and $\mathbb{L}_\alpha$ listed in Lemmas 5.27 and 5.28 can be routinely checked.

Lemma 5.27. Let $\alpha \in \text{Ord}$.

1. $\mathbb{L}_\alpha$ and $\mathbb{L}$ are transitive. 
2. $\mathbb{L}_\alpha \in \mathbb{L}_{\alpha + 1}$ (Lemma §VI 1.10 [12]). 
3. $\forall \beta < \alpha (\mathbb{L}_\beta \subset \mathbb{L}_\alpha)$ (Lemma §VI 1.6 [12]). 
4. $\mathbb{L}_\alpha = \{x \in \mathbb{L} : \rho(x) < \alpha\}$ (Lemma §VI 1.8 [12]). 
5. $\mathbb{L}_\alpha \cap \text{Ord} = \alpha$ (Lemma 13.2 [9]). 
6. $\alpha \subset \mathbb{L}_\alpha$ and $\rho(\alpha) = \alpha$ (Lemma §VI 1.9 (a) [12]).
Lemma 5.28.

1. \( \forall \alpha \in \text{Ord}(\mathbb{L}_\alpha \subset \mathbb{V}_\alpha) \) (Lemma §VI 1.11 [12]).

2. \( \forall n \leq \omega(\mathbb{L}_n = \mathbb{V}_n) \) (Lemma §VI 1.13 [12]).

\( \square \)

Lemma 5.29 (Exercise 12.16 [9]). Let \( \varphi \) be a formula. Then there exists an arbitrarily large ordinal \( \alpha \) such that for all \( v_1, \ldots, v_n \in \mathbb{L}_\alpha \),

\[ \varphi^\mathbb{L}(v_1, \ldots, v_n) \Leftrightarrow \varphi^{\mathbb{L}_\alpha}(v_1, \ldots, v_n). \]

Proof. We prove by induction on the complexity of \( \varphi \). It is straightforward if \( \varphi \) is atomic, \( \lnot \psi, \psi \land \phi, \psi \lor \phi, \psi \Rightarrow \phi \), or \( \psi \Leftrightarrow \phi \), assuming the result holds for \( \psi \) and \( \phi \). Since

\[ \exists x \in y(\varphi^\mathbb{L}) \Leftrightarrow \exists x \in \mathbb{L}(x \in y \land \varphi^\mathbb{L}) \]

and

\[ \forall x \Leftrightarrow \lnot \exists x \lnot, \]

it remains to show if the result holds for \( \varphi \), then it also holds for \( \exists x(\varphi) \). Thus we finish the proof by seeing, for arbitrarily large \( \alpha \) and \( v_1, \ldots, v_n \in \mathbb{L}_\alpha \),

\[ \exists x \in \mathbb{L}_\alpha(\varphi^\mathbb{L}_\alpha(x, v_1, \ldots, v_n)) \]

\[ \Leftrightarrow \exists x \in \mathbb{L}_\alpha(\varphi(x, v_1, \ldots, v_n)) \]

\[ \Leftrightarrow \exists x \in \mathbb{L}(\varphi(x, v_1, \ldots, v_n)). \]

The first equivalence is by the assumption of \( \varphi \), and the second one is by the arbitrary largeness of \( \alpha \). \( \square \)

Theorem 5.30 (Theorem 13.3 [9]). \( \mathbb{L} \) is a model for ZF.

Proof. We prove \( \sigma^\mathbb{L} \) holds for every axiom \( \sigma \) in ZF.

Extensionality. We want to show

\[ \mathbb{L} \models \forall x \forall y(\varphi(x, y)), \]

where \( \varphi \) is

\[(\forall u \in x(u \in y) \land \forall u \in y(y \in x)) \Rightarrow x = y.\]

But \( \varphi \) is a \( \Delta_0 \)-formula and thus absolute.

Foundation. We want to show

\[ \mathbb{L} \models \forall x(x \neq \emptyset \Rightarrow \exists u \in x(u \cap x) = \emptyset). \]

Since \( \emptyset^\mathbb{L} = \emptyset \), let \( \emptyset \neq x \in \mathbb{L} \). Then \( \exists u \in x(u \cap x) = \emptyset \). \( \mathbb{L} \) is transitive so \( u \in \mathbb{L} \), and \( u \cap x \) is \( \Delta_0 \).
**Pairing.** We want to show
\[ L \models \forall x \forall y \exists z (z = \{ x, y \}). \]

Let \( x, y \in L \) and define \( z = \{ x, y \} \). Then \( \exists \alpha \in \text{Ord}(x \in L_\alpha \land y \in L_\alpha) \). So \( z \in L_{\alpha+1} \) since \( \{ x, y \} \) is definable on \( L_\alpha \). And \( z = \{ x, y \} \) is \( \Delta_0 \).

**Union.** We want to show
\[ L \models \forall x \exists y (y = \bigcup x). \]

Let \( x \in L \) and define \( y = \bigcup x \). Since \( L \) is transitive, \( z \in L \) for all \( z \in x \).

Use the transitivity of \( L \) again, we have
\[ \forall z \in x (u \in z \Rightarrow u \in L). \]

Therefore \( y \subset L \). So there exists an ordinal \( \alpha \) such that \( x \in L_\alpha \) and \( y \subset L_\alpha \).

Let \( t \in \bigcup s \) is \( \Delta_0 \). So
\[ y = \{ u \in L_\alpha : u \in \bigcup x \} = \{ u \in L_\alpha : L_\alpha \models u \in \bigcup x \} \]

is \( L_\alpha \)-definable. Then
\[ \forall x \in L \exists y \in L (y = \bigcup x) \]
\[ \Leftrightarrow \forall x \in L \exists y \in L (y = \bigcup x)^L \]
\[ \Leftrightarrow (\forall x \exists y (y = \bigcup x))^L. \]

**Power Set.** We want to show
\[ L \models \forall x \exists y (\varphi(x, y)), \]
where \( \varphi \) is
\[ \forall u (u \subset x \Leftrightarrow u \in y). \]

Let \( x \in L \) and \( y = \mathcal{P}(x) \cap L \). There exists an \( \alpha \) such that \( y \in L_\alpha \). \( L_\alpha \) is transitive, so \( y \subset L_\alpha \). Then
\[ y = \{ u \in L_\alpha : u \subset x \}. \]

Clearly, \( \varphi \) is \( \Delta_0 \) and \( y \) satisfies \( \varphi \).

**Infinity.** It suffices to show \( L \) contains an (infinite) inductive set, i.e.
\[ L \models \exists x (\varphi(x, u)), \]
where \( \varphi \) is
\[ \emptyset \in x \land \forall u \in x (\{ u, \{ u \} \} \in x). \]

We have already shown above \( \emptyset^L = \emptyset \) and the Axiom of Pairing holds in \( L \). So \( \omega \in L \) satisfies \( \varphi \).
Replacement. We want to show if $\varphi$ is a formula such that
\[ \mathbb{L} \models \forall x \forall y \forall z (\varphi(x, y) \land \varphi(x, z) \Rightarrow y = z), \]
where we write $\varphi(x, y)$ instead of $\varphi(x, y, \ldots)$ for convenience, then
\[ \mathbb{L} \models \forall x \exists y \forall u \left( u \in y \iff \exists s \in x(\varphi(s, u)) \right). \]
Let $x \in \mathbb{L}$ and define $\chi(s, \alpha)$ to be
\[ \alpha \in \text{Ord} \land \exists t \in \mathbb{L}_\alpha(\varphi^L(s, t)) \land \forall \beta < \alpha \forall u \in \mathbb{L}_\beta(\neg \varphi^L(s, u)). \]
Note that for $\chi(s, \alpha)$,
\[ \alpha = \min \{ \rho(t) : \varphi^L(s, t) \}, \]
where $\rho(t)$ is the $\mathbb{L}$-rank of $t$. Then clearly, if $\chi(s, \alpha)$ holds for some $s$, $\alpha$ is unique. Now by the Axiom of Replacement, there exists $y$ such that
\[ \forall \alpha \left( \alpha \in y \iff \exists s \in x(\chi(s, \alpha)) \right). \]
Let $\gamma = \bigcup (y \cap \text{Ord})$. Then $\gamma$ is an ordinal, and
\[ \forall s \in x \forall \alpha \in \text{Ord} (\chi(s, \alpha) \Rightarrow \alpha < \gamma). \]
By the definition of $\chi$,
\[ \forall s \in x \forall t \in \mathbb{L}_\alpha \exists \alpha \in \text{Ord}(\varphi^L(s, t) \Rightarrow t \in \mathbb{L}_\alpha). \]
But $t \in \mathbb{L}_\alpha \Rightarrow t \in \mathbb{L}_\gamma$. So $\mathbb{L}_\gamma$ satisfies
\[ \mathbb{L} \models \forall u \left( u \in \mathbb{L}_\gamma \iff \exists s \in x(\varphi(s, u)) \right). \]

Separation. We want to show if $\varphi$ is a formula, then
\[ \mathbb{L} \models \forall x \forall p_1 \ldots \forall p_n \exists y \forall s \left( s \in y \iff s \in x \land \varphi(s, p_1, \ldots, p_n) \right). \]
Let $x, p_1, \ldots, p_n \in \mathbb{L}$ and define
\[ y = \{ s \in x : \varphi^L(s, p_1, \ldots, p_n) \}. \]
It is easy to see by lemma 5.29 that there exists an $\alpha$ such that
\[ y = \{ s \in \mathbb{L}_\alpha : s \in x \land \varphi^L(s, p_1, \ldots, p_n) \}. \]
Therefore $y \in \mathbb{L}$ and satisfies
\[ \forall s \in \mathbb{L} \left( s \in y \iff s \in x \land \varphi(s, p_1, \ldots, p_n) \right). \]
5.5 Gödel operations

Definition 5.31. A Gödel operation is one of or a composition of the following basic (Gödel) operations:

\[
g_1(x, y) = \{x, y\},
\]

\[
g_2(x, y) = x \times y,
\]

\[
g_3(x, y) = \{(u, v) \in x \times y : u \in v\},
\]

\[
g_4(x, y) = x - y,
\]

\[
g_5(x) = \bigcup x,
\]

\[
g_6(x) = \text{dom}(x),
\]

\[
g_7(x \times y) = y \times x.
\]

Proposition 5.32. The following operations are Gödel operations:

1. \(g_8(x, y) = x \cap y\).
2. \(g_9(x) = \text{ran}(x)\).
3. \(g_{10}(x \times y \times z) = x \times z \times y\).

Proof.

1. \(g_8(x, y) = x \cap y = x - (x - y)\).
2. \(g_9(x) = \text{ran}(x) = \text{dom}(g_7(x))\).
3. Since \(x \times y \times z = x \times (y \times z)\), then
\[
g_{10}(x \times (y \times z)) = x \times (z \times y)
\]
\[
= \text{dom}\left(x \times (y \times z)\right) \times g_7\left(\text{ran}(x \times (y \times z))\right).
\]

Observe that operations \(g_1\) to \(g_{10}\) are in fact the set operations to generate new sets as a group of axioms of set theory [6]. Recall Lemma 5.15 that \(\{\cdot, \cdot\}\), \(\times\), \(-\), \(\cap\), \(\bigcup\), \(\text{dom}\), and \(\text{ran}\) are \(\Delta_0\)-operations. One may conjecture that all the Gödel operations are \(\Delta_0\). We prove this now.

Lemma 5.33 (Lemma 13.7 [9]). Any Gödel operation \(G\) is a \(\Delta_0\)-operation.

Proof. We prove by induction on the complexity of \(G\). To show

i) \(z = G(\ldots)\) is \(\Delta_0\),
it is equivalent to show
\[ \forall u \in z(u \in G(\ldots)) \land \forall u \in G(\ldots)(u \in z) \]
is \(\Delta_0\); thus it is also equivalent to show

ii) \(u \in G(\ldots)\) is \(\Delta_0\), and

iii) if \(\varphi\) is a \(\Delta_0\)-formula, then \(\forall u \in G(\ldots)(\varphi(u, \ldots))\) is \(\Delta_0\). (Similarly, \(\exists u \in G(\ldots)(\varphi(u, \ldots))\) is also \(\Delta_0\).)

First we show the basic Gödel operations \(g_1\) to \(g_7\) are \(\Delta_0\). Since we have already seen in Lemma 5.15 that \(g_1, g_2, g_4, g_5,\) and \(g_6\) are \(\Delta_0\), it remains to show \(g_3\) and \(g_7\) are also \(\Delta_0\). This is true since

\[ z = g_3(x, y) = \{ (u, v) \in x \times y : u \in v \} \]
iff \(\forall s \in z \exists u \in x \exists v \in y (u \in v \land s = (u, v)) \land \forall u \in x \forall v \in y \exists s \in z (u \in v \Rightarrow s = (u, v)) \),

and

\[ z = g_7(x) = \{ (u, v) : (v, u) \in x \} \] iff
\[ \forall p \in z \exists s \in x \exists u \in y (u \in v \land s = (u, v)) \land \forall s \in x \forall u, v \in \bigcup \bigcup x \exists p \in z (s = (v, u) \Rightarrow p = (u, v)) \].

Therefore, i), ii) and iii) hold for the basic cases.

We now proceed to the inductive steps. Assume i), ii) and iii) hold for \(G(\ldots)\) and \(H(\ldots)\), two Gödel operations. We would like to show the basic Gödel operations on \(G(\ldots)\) and \(H(\ldots)\) are \(\Delta_0\). For convenience, we drop \((\ldots)\) in \(G(\ldots)\) and \(H(\ldots)\).

\[ z = g_1(G, H) = \{ G, H \} \]
iff \(\forall u \in z(u = G \lor u = H) \land \exists u, v \in z(u = G \land v = H)\),

\[ z = g_2(G, H) = G \times H \]
iff \(\forall u \in z \exists s \in G \exists t \in H(u = (s, t)) \land \forall s \in G \forall t \in H \exists u \in z(u = (s, t))\),

\[ z = g_3(G, H) = \{ (u, v) \in G \times H : u \in v \} \]
iff \(\forall s \in z \exists u \in G \exists v \in H(u \in v \land s = (u, v)) \land \forall u \in G \forall v \in H \exists s \in z(u \in v \Rightarrow s = (u, v))\),
\[ z = g_4(G, H) = G - H \]
\[
\text{iff } \forall u \in z(u \in G) \land \neg(\exists u \in z(u \in H)) \land \forall u \in G(u \in H \lor u \in z),
\]

\[ z = g_5(G) = \bigcup G \]
\[
\text{iff } \forall u \in z \exists s \in G(u \in s) \land \forall s \in G(s \subset z),
\]

\[ z = g_6(G) = \text{dom}(G) \]
\[
\text{iff } \forall u \in z \exists s \in G \forall t \in s(u \in t) \land \forall s \in G \forall t \in s \forall u, v \in t(s = (u, v) \Rightarrow u \in z),
\]

\[ z = g_7(G) = \{(u, v) : (v, u) \in G\} \]
\[
\text{iff } \forall p \in z \exists s \in G \exists u, v \in \bigcup \bigcup G(s = (v, u) \land p = (u, v)) \land \forall s \in G \forall u, v \in \bigcup \bigcup G \exists p \in z(s = (v, u) \Rightarrow p = (u, v)).
\]

For \( g_1 \), we proved directly. For \( g_2 \) to \( g_7 \), we used inductive hypotheses. For \( g_7 \),
we are allowed to use \( \forall u \in \bigcup \bigcup G \) since we already proved the \( g_5 \) case. This completes the proof.

Having seen that any Gödel operation can be somehow expressed by a \( \Delta_0 \)-formula, we naturally ask if the reverse is also true. Next we show the set defined by a \( \Delta_0 \)-formula is a Gödel operation.

**Theorem 5.34** (Gödel's Normal Form Theorem, 13.7 [9]). Given a \( \Delta_0 \)-formula \( \varphi(v_1, \ldots, v_n) \) and a finite sequence \( x_1, \ldots, x_n \) of sets, there exists a Gödel operation \( G \) on \( x_1, \ldots, x_n \) such that \( G \) is defined by \( \varphi \), i.e.

\[ G(x_1, \ldots, x_n) = \{(v_1, \ldots, v_n) \in x_1 \times \ldots \times x_n : \varphi(v_1, \ldots, v_n)\}. \]

**Proof.** Let \( v = (v_1, \ldots, v_n) \) and \( x = x_1 \times \ldots \times x_n \). We prove by induction on the complexity of \( \varphi \). As we see the definition 5.9, the induction is done by the number of symbols \( \in, =, \land, \lor, \neg, \Rightarrow, \Leftrightarrow, \) and restricted quantifiers, \( \exists v_i \in v_j \) and \( \forall v_i \in v_j \) with \( i \neq j \).

We can assume \( \varphi \) contains only \( \in, \land, \neg, \) and restricted \( \exists \) since

\[ a \lor b \text{ iff } \neg(\neg a \land \neg b), \]
\[ \forall a \in b(\varphi) \text{ iff } \neg\exists a \in b(\neg\varphi), \]
\[ a = b \text{ iff } \forall u \in a(u \in b) \land \forall u \in b(u \in a), \]
\[ a \Rightarrow b \text{ iff } b \lor \neg a, \]
\[ a \Leftrightarrow b \text{ iff } a \Rightarrow b \land b \Rightarrow a. \]
Further, if \( \varphi \) is \( \exists v_s \in v_t(\psi(..., v_s, ..., v_t)) \), we can assume \( \varphi \) is

\[ \exists v_s \in v_t(\psi(..., v_t, ..., v_s)) \],

so that the quantified variable \( v_s \) is of the highest index, because this is just a matter of renaming.

**Case 1** We prove for an atomic formula \( \varphi \), i.e. \( v_i \in v_j \) where \( i \neq j \), by induction on the number of variables in \( \varphi \) (allowing dummy variables).

**Case 1a** If \( n = 2 \), then

\[ \{(v_1, v_2) \in x_1 \times x_2 : v_1 \in v_2\} = g_3(x_1, x_2) \]

and

\[ \{(v_1, v_2) \in x_1 \times x_2 : v_2 \in v_1\} = g_7(g_3(x_2, x_1)). \]

**Case 1b** If \( n > 2 \) and \( i, j \neq n \), suppose there exists a \( G \) such that

\[ G(x_1, \ldots, x_{n-1}) = \{(v_1, \ldots, v_{n-1}) \in x_1 \times \ldots \times x_{n-1} : v_i \in v_j\}. \]

Then

\[ \{v \in x : v_i \in v_j \land i, j \neq n\} = G(x_1, \ldots, x_{n-1}) \times x_n. \]

**Case 1c** If \( n > 2 \) and \( i, j \neq n - 1 \), by Case 1b we can assume there exists a \( G \) such that

\[ G(x_1, \ldots, x_n) \]

\[ = \{(v_1, \ldots, v_{n-2}, v_n, v_{n-1}) \in x_1 \times \ldots \times x_{n-2} \times x_n \times x_{n-1} : v_i \in v_j \land i, j \neq n - 1\} \]

\[ = \{((v_1, \ldots, v_{n-2}), v_n, v_{n-1}) \in (x_1 \times \ldots \times x_{n-2}) \times x_n \times x_{n-1} : v_i \in v_j \land i, j \neq n - 1\}. \]

Then

\[ \{v \in x : v_i \in v_j \land i, j \neq n - 1\} \]

\[ = \{((v_1, \ldots, v_{n-2}), v_n, v_{n-1}) \in (x_1 \times \ldots \times x_{n-2}) \times x_n \times x_{n-1} : v_i \in v_j \land i, j \neq n - 1\} \]

\[ = g_{10}(G(x_1, \ldots, x_n)). \]

**Case 1d** If \( n > 2 \) and \( i = n - 1, j = n \), then

\[ \{v \in x : v_{n-1} \in v_n\} = x_1 \times \ldots \times x_{n-2} \times g_3(x_{n-1}, x_n). \]

**Case 1e** If \( n > 2 \) and \( i = n, j = n - 1 \), then

\[ \{v \in x : v_n \in v_{n-1}\} = g_{10}((x_1 \times \ldots \times x_{n-2}) \times g_3(x_n, x_{n-1})). \]
Case 2 If \( \varphi(v) \) is \( \neg \psi(v) \), suppose there exists a \( G \) such that

\[
G(x_1, \ldots, x_n) = \{ v \in x : \psi(v) \}.\]

Then

\[
\{ v \in x : \varphi(v) \} = x - G(x_1, \ldots, x_n).
\]

Case 3 If \( \varphi(v) \) is \( (\psi_1 \land \psi_2)(v) \), suppose for \( i = 1, 2 \),

\[
\{ v \in x : \psi_i(v) \} = G_i(x_1, \ldots, x_n).
\]

Then

\[
\{ v \in x : \varphi(v) \} = G_1(x_1, \ldots, x_n) \cap G_2(x_1, \ldots, x_n).
\]

Case 4 If \( \varphi(v) \) is \( \exists u \in v_i(\psi(v,u)) \) with \( 1 \leq i \leq n \), suppose by inductive hypothesis

\[
\{ (v,u) \in x \times y : \psi(v,u) \} = G_1(x_1, \ldots, x_n, y)
\]

and by Case 1

\[
\{ (v,u) \in x \times y : u \in v_i \} = G_2(x_1, \ldots, x_n, y)
\]

Since

\[
u \in v_i \implies u \in \bigcup x_i,
\]

we have

\[
\varphi(v) \text{ iff } \exists u \in v_i(\psi(v,u))
\]

\[
\text{iff } \exists u(\psi(v,u) \land u \in v_i \land u \in \bigcup x_i)
\]

\[
\text{iff } v \in \text{dom}\{ (v,u) \in x \times \bigcup x_i : \psi(v,u) \land u \in v_i \}.
\]

Then

\[
\{ (v,u) \in x \times \bigcup x_i : \varphi(v,u) \}
\]

\[
=x \cap \text{dom}(G_1(x_1, \ldots, x_n, \bigcup x_i) \cap G_2(x_1, \ldots, x_n, \bigcup x_i)).
\]

Therefore, we have finished the proof. \( \square \)

**Notation 5.35.** Let \((\text{cl})_G(x)\) denote the closure of \( x \) under Gödel operations.

**Remark 5.36.** Think carefully \((\text{cl})_G(x)\) as it will play a crucial role in the rest of our proof. We construct it by induction on \( \omega \).

At \( n = 0 \), let

\[
x_0 = x.
\]
At \( n = 1 \), apply once (arbitrary) basic Gödel operation over any \( u, v \in x_0 \), collect all the yielded sets, and along with \( x_0 \), define

\[
x_1 = x_0 \cup \{ g_i(u, v) : u, v \in x_0, i = 1, \ldots, 7 \}.
\]

Then, inductively, at \( n = m \), define

\[
x_{m+1} = x_m \cup \{ g_i(u, v) : u, v \in x_m, i = 1, \ldots, 7 \}.
\]

Finally, we have

\[
(\text{cl})_G(x) = \bigcup_{n \in \omega} x_n.
\]

**Proposition 5.37.** \((\text{cl})_G\) is a \( \Sigma_1 \)-operation.

**Proof.** In the light of Remark 5.36, we see

\[
y = (\text{cl})_G(x)
\]

iff \( \exists f \left( f \text{ is a function} \land \text{dom}(f) = \omega \land y = \bigcup \text{ran}(f) \land f(0) = x \land n \in \omega (f(n + 1) = f(n) \cup \{ g_i(u, v) : u, v \in f(n), i = 1, \ldots, 7 \}) \right) \).

\( \square \)

**Corollary 5.38.** For a transitive set \( M \),

\[
\mathcal{P}(M) = (\text{cl})_G(M \cup \{ M \}) \cap \mathcal{P}(M).
\]

**Proof.** By Proposition 5.10 and Gödel's Normal Form Theorem, we have

\[
\mathcal{P}(M) \subset (\text{cl})_G(M \cup \{ M \}) \cap \mathcal{P}(M).
\]

For the other direction, let \( x \subset M \) and \( x = G(\ldots) \) for some Gödel operation \( G \) over \( M \cup \{ M \} \). Since \( G \) is \( \Delta_0 \), there exists a \( \Delta_0 \)-formula \( \varphi \) such that

\[
x = \{ v \in M : \varphi(v, \ldots) \}.
\]

But \( \varphi \) is \( \Delta_0 \), so

\[
x = \{ v \in M : \varphi^M(v, \ldots) \} \in \mathcal{P}(M).
\]

\( \square \)

**Corollary 5.39.** Gödel operations are absolute for transitive models.

\( \square \)
5.6 Con(ZF + $\forall = L$)

Lemma 5.40 (Lemma 13.14 [9]). The function $\mathcal{D}$ is absolute for $L$.

Proof. We show $\mathcal{D}$ is $\Delta_1$. First it is clear that

$$z = \mathcal{D}(x) = (cl)_G(x \cup \{x\}) \cap \mathcal{D}(x)$$

iff $\exists y \left( y = (cl)_G(x \cup \{x\}) \land \forall u \in z(u \subseteq x \land u \in y) \right) \land \forall u \in y(u \subseteq x \Rightarrow u \in z)$.\]

Since $(cl)_G$ is $\Sigma_1$ by proposition 5.37 and the rest after the quantifier $\exists y$ is $\Delta_0$, the whole thing after $\exists y$ is $\Sigma_1$. So $\mathcal{D}$ is a $\Sigma_1$-function by Lemma 5.13 (1).

Now

$$z = \mathcal{D}(x) \text{ iff } x \in \text{dom(} \mathcal{D} \text{)} \land \forall u \left( \neg(u = z) \Rightarrow \neg(u = \mathcal{D}(x)) \right).$$

Expressions $x \in \text{dom(} \mathcal{D} \text{)}$ and $\neg(u = z)$ are $\Delta_0$, and $\neg(u = \mathcal{D}(x))$ is $\Pi_1$ by Lemma 5.13 (2). Then $z = \mathcal{D}(x)$ is also $\Pi_1$, again by Lemma 5.13 (2). So $\mathcal{D}$ is $\Delta_1$ and thus absolute for $L$ by Lemma 5.19. \hfill $\square$

Lemma 5.41. $\forall \alpha \in \text{Ord}(\text{ Ord}(L_\alpha)) = L_\alpha$.

Proof. We prove by contradiction. Suppose there exists an $\alpha$ such that $(L_\alpha)^L \neq L$. Let $\beta$ be the smallest such ordinal. First trivially $\beta \neq 0$. If $\beta > 0$ and is a limit, then

$$(L_\beta)^L = \left( \bigcup_{\alpha < \beta} L_\alpha \right)^L$$

$= \bigcup_{\alpha < \beta} (L_\alpha)^L$ by Lemma 5.15

$= \bigcup_{\alpha < \beta} L_\alpha$ by hypothesis

$= L_\beta$ contradicting the definition of $\beta$.

If $\beta$ is a successor, then

$$(L_\beta)^L = (\mathcal{D}(L_{\beta-1}))^L$$

$= \mathcal{D}(L_{\beta-1})^L$ by Lemma 5.40

$= \mathcal{D}(L_{\beta-1})$ by hypothesis

$= L_{\beta}$ contradicting the definition of $\beta$.

The result thus follows. \hfill $\square$
Corollary 5.42. \( L \models V = L \).

Proof. Since
\[
L \models V = L
\]
iff \( L \models \forall x \exists \alpha (\alpha \in \text{Ord} \land x \in L_\alpha) \)
iff \( \forall x \in L \exists \alpha \in L (\alpha \in (\text{Ord})^L \land x \in L_\alpha) \),
the result follows from Lemmas 5.27 (5) and 5.41.

Corollary 5.43. \( L \) is a model for \( \text{ZF} + V = L \).

5.7 Con(ZFC)

Theorem 5.44 (Definition §VI 4.1 [12], Theorem 13.18 [9]). There exists a well-ordering for \( L \).

Proof. By induction we construct a well-ordering \( \triangleleft_\alpha \) for each \( L_\alpha \). We require these well-orderings have the properties that whenever \( \alpha < \beta \),
\[
x \triangleleft_\alpha y \Rightarrow x \triangleleft_\beta y, \text{ and }
x \in L_\alpha \land y \in L_\beta - L_\alpha \Rightarrow x \triangleleft_\beta y.
\]

If \( \alpha \) is a limit, define
\[
\triangleleft_\alpha = \{(x, y) \in L_\alpha \times L_\alpha : \rho(x) < \rho(y) \lor (\rho(x) = \rho(y) \land (x, y) \in \triangleleft_{\rho(x)+1})\},
\]
where \( \rho \) is the \( L \)-rank function.

Recall, by Corollary 5.38 and Remark 5.36,
\[
L_{\alpha+1} = S(L_\alpha) = (\text{cl})_G(L_\alpha \cup \{L_\alpha\}) \cap S(L_\alpha) = \bigcup_{n \in \omega} x_{\alpha,n} \cap S(L_\alpha),
\]
where
\[
x_{\alpha,0} = L_\alpha,
x_{\alpha,n+1} = x_{\alpha,n} \cup \{g_i(u, v) : u, v \in x_{\alpha,n}, i = 1, \ldots, 7\}.
\]

So given \( \triangleleft_\alpha \) and \( s, t \in L_{\alpha+1} \), we use Gödel operations to construct \( \triangleleft_{\alpha+1} \) by induction on \( n \) (in \( x_{\alpha,n} \)) as following:

(i) If \( s, t \in L_\alpha \), then \( s \triangleleft_{\alpha+1} t \) iff \( s \triangleleft_\alpha t \).

(ii) If \( s \in L_\alpha \) and \( t \in L_{\alpha+1} - L_\alpha \), then \( s \triangleleft_{\alpha+1} t \).

(iii) If \( s \in x_{\alpha,n}, t \in x_{\alpha,m} \), then \( s \triangleleft_{\alpha+1} t \) iff \( n < m \).
(iv) If \( s,t \in x_{\alpha,n+1} - x_{\alpha,n} \), then \( s \triangleleft_{\alpha+1} t \) iff \( i_s < i_t \), where \( i_s \) is the least \( i \) such that

\[ \exists u,v \in x_{\alpha,n}(s = g_i(u,v)). \]

(v) If in (iv) \( i_s = i_t \), then \( s \triangleleft_{\alpha+1} t \) iff \( u_s \triangleleft_{\alpha+1} u_t \), where \( u_s \) is the \( \triangleleft_{\alpha+1} \)-least \( u \) such that

\[ \exists v \in x_{\alpha,n}(s = g_i(u_s,v)). \]

(vi) If in (v) \( u_s = u_t \), then \( s \triangleleft_{\alpha+1} t \) iff \( v_s \triangleleft_{\alpha+1} v_t \), where \( v_s \) is the \( \triangleleft_{\alpha+1} \)-least \( v \) such that

\[ s = g_i(u_s,v_s). \]

(vii) If in (vi) \( v_s = v_t \), it is easy to see \( s = t \).

Define

\[ \triangleleft_L = \bigcup_{\alpha \in \text{Ord}} \triangleleft_{\alpha}. \]

Then \( \triangleleft_L \) is a a well-ordering for \( L \). □

**Corollary 5.45.** \( L \models \forall = L \Rightarrow \text{AC}. \) □

**Corollary 5.46.** \( \text{Con}(ZF) \Rightarrow \text{Con}(ZFC) \) □

### 5.8 ZF + \( \neg \text{AC} \)

Our last words for this section is the following deep result which we only state here.

**Theorem 5.47** ([3], [4]). *There is a model of ZF in which \( \mathbb{R} \) cannot be well-ordered. Therefore*

\[ \text{Con}(ZF) \Rightarrow \text{Con}(ZF + \neg \text{AC}). \]

This theorem is proved by a powerful technique, called *forcing*, for consistency and independence proofs introduced by Paul J. Cohen in 1963 and 1964 ([3], [4]). It looks somewhat surprising and strange at the first glance, since we know how to order \( \mathbb{R} \) and the consistency of ZFC in our “normal” mathematics world might have let us believe AC was true in general. Along with \( \text{Con}(ZFC) \), we have:

**Corollary 5.48.** *AC is independent of ZF.* □
6 The Sun and a pea

To those who feel that the Banach–Tarski Paradox is absurd, it seems evident that the Axiom of Choice is the culprit.

Stan Wagon

In 1924, Banach and Tarski, using AC, proved a surprising and counterintuitive theorem:

**Theorem 6.1 (Banach–Tarski Theorem).** A unit 3-ball can be divided into a finite number of disjoint pieces, such that these pieces can be reassembled to create two unit 3-balls.

The proof is not included here but can be found in Chapter 3 of Wagon 1985 [21] or Chapter 5 of Wapner 2005 [22]. The idea of the proof is: assuming AC, there exist non-Lebesgue measurable sets (Theorem 4.8); divide a unit 3-ball into a finite number of non-Lebesgue measurable subsets; and these subset can somehow be put back to get two unit 3-balls.

Given Banach–Tarski Theorem, one can make an arbitrarily large ball using a unit ball; hence the title of this section.

Banach–Tarski theorem was one of the important reasons that AC was rejected by many mathematicians before Con(ZFC) was proved. This paradox raised the fear that AC could lead to contradiction. To remedy this, some weaker forms of AC were suggested. One of them is

**The Axiom of Dependent Choices** Give a binary relation $\prec$ on a set $x$, if

$$\forall s \in x \exists t \in x (t \prec s),$$

then there exist \( \{s_n \in x : n \in \omega\} \) such that

$$\forall n \in \omega (s_{n+1} \prec s_n).$$

It can be shown that there exists a model for ZF + DC in which any subset of \( \mathbb{R} \) is Lebesgue measurable; see Theorem 10.10 [8]. Therefore, ZF + DC seems to
appease both sides: there are choice functions but no longer the Banach–Tarski paradox.

However, as Wagon argued, if some forms of non-constructive choice functions are tolerated, why are not all the choice functions [21]? AC endows and inspire our mathematicians and scientists with a brave and liberal attitude. Without AC, mathematics world would have been a collection of algorithms. Those paradoxes led by AC, like the Banach–Tarski one, instead of raising any fear to AC, actually reinforce and reshape our views to mathematics.

Many fruitful results are known to be provable if and only if AC is assumed. One example is the existence of non-Lebesgue measurable sets. Another one is

**Theorem 6.2** (Theorem 10.11 [8]). *There exists a model for ZF in which there exists a vector space with no basis.*

In fact, the physical problems caused by AC happens entirely in abstract mathematical worlds, more specifically, in the area of Lebesgue measurable sets. Therefore, doubling the volume of a ball simply means the physical property, volume, is an undefinable term in Lebesgue measure theory.

Moreover, AC also helps to discover results in the world without assuming AC. For example, the proof of Schröder–Bernstein theorem can be without using AC, but is easier in ZFC [21]. Without AC, its discovery would be postponed. For another example, in response to the Banach–Tarski paradox, the concept of amenable groups was introduced to deal with paradoxical groups by von Neumann in 1929 [20].
7 Conclusion

No doubt [AC] will always be desirable despite the technical interest of various independence questions involving it and weaker principles.

Dana Scott

We have shown some equivalents of AC and results requiring AC from various branches of mathematics. We have also justified why AC can be used by our proof of the consistency of AC in ZF.

In most areas other than set theory, AC is used as a simple fact. However, in set theory, research on this topic often goes quite deep, e.g. large cardinals. This is especially powered by forcing, the technique invented by Cohen for his proof of Con(ZF + ¬AC). For the reader who would like more information about forcing, Cohen 1963 [3], Cohen 1964 [4], Kunen 1980 [12], and Jech 2002 [9] are recommended. For contents about large cardinals, one is referred to Jech’s 2002 textbook [9]. For more fundamental results, one may consult Marker 2002 [14] for model theory or Mendelson 1987 [15] for the Gödel Completeness Theorem and the two Gödel Incompleteness Theorems.
References


