Axiomatic Set Theory

— Lecture Note of “数理論理学特論”, Spring Semester 2013

渕野 昌 (Sakaé Fuchino)

29/07/2013 (18:29JST)

1 Introduction

intro

[This will be written later.]

2 Naïve Axiomatic Set Theory

naive

In this section, we introduce the axiom systems: Zermelo set theory (Z) and Zermelo-Fraenkel set theory (ZF) as well as these axiom systems with Axiom of Choice (AC) which are denoted by ZC and ZFC respectively. These Axiomatic frameworks encompass “the whole mathematics”. Most of the the classical mathematics can be developed in Z. While ZC is sufficient to develop the 19th century or even the early 20th century mathematics (excluding Cantor’s theory of transfinite numbers), their extension ZFC is needed to develop the full-fledged modern mathematics including the theory of transfinite numbers inside the axiomatic framework.

These axiom systems as treated in this section are still “naïve” in that they are formulated without the precision first obtained by formulating them in the language of first order logic which we introduce in Section 4.

In the following development of axiom systems, we take the standpoint that our mathematical world consists of sets (and nothing else). In particular, if we say “for all x” or “there exists x”, the range of the variable x is thought to be all possible “sets”. We describe the fundamental properties of these sets in the following axioms in terms of the relation "x is an element of y" which is denoted by $x \in y$. If x is not an element of y, this is denoted by $x \notin y$.

The statement "*our mathematical world consists of sets (and nothing else)*" also imply that the elements of a set are also sets. Hence, with regard to mathematical contents, there is no difference whether we say "a set" or "a family of sets". Still, I shall sometimes make this difference for the sake of intelligibility. Since we are talking only about sets anyway, we drop often the explanation that something be a set. As this is the case in the formulation of our first axiom asserting that the elements of a set decide this set completely:

Axiom I. (Extensionality). Any two x, y are equal (denoted by $x = y$) if and only if x and y have exactly the same elements. That is, $x = y$ if and only if, for any u, we have $u \in x$ iff $u \in y$.

Axiom II. (Empty Set). There is a set x which has no element. That is, there is an x such that $u \notin x$ for any u.

By Axiom I, the set x whose existence is guaranteed by Axiom II is unique. We shall denote this unique set by $\emptyset$ and call it the empty set. Any set which is not $\emptyset$ is said to be a non-empty set.

Axiom III. (Pairing). For any x, y there is a u which consists of x and y. That is, for any x, y, there is a u with the property that, for any v, $v \in u$ iff $v = x$ or $v = y$.

Here, again by Axiom I, the set u is unique for given x and y. So we shall denote this set u by $\{x, y\}$. If $x = y$ the set u consists of the unique element x $(= y)$. In this case we denote this u by $\{x\}$ (or $\{y\}$).

Axiom IV. (Union). For any x there is a u which consists of elements of some element of x. That is, for any x, there is a u with the property that, for any v, $v \in u$ iff there is $y \in x$ such that $v \in y$.

The set u as above is again unique for a given x. So we shall denote such u by $\bigcup x$. If $x = \{a, b\}$ then we also denote $\bigcup x$ by $a \cup b$. For any v, $v \in a \cup b$ iff $v \in a$ or $v \in b$.

y is said to be a subset of x (notation: $y \subseteq x$), if $u \in y$ implies $u \in x$ for any u.

For any x, we have $x \subseteq x$ and $\emptyset \subseteq x$.

Axiom V. (Power Set). For any x, there is y which consists of all subsets of x. That is, for any x, there is y with the property that, for any u, $u \in y \Leftrightarrow u \subseteq x$.

The set y in Axiom V is again unique for a given x. So we shall denote this by $\mathcal{P}(x)$ and call it the power set of x. For any x, we have $\emptyset$, $x \in \mathcal{P}(x)$.

psof-subset

Lemma 2.1 *If $x \subseteq y$ then $\mathcal{P}(x) \subseteq \mathcal{P}(y)$.*

Proof. Suppose $x \subseteq y$. If $u \in \mathcal{P}(x)$ (or equivalently $u \subseteq x$) then $u \subseteq y$ (or equivalently $u \in \mathcal{P}(y)$). This shows that $\mathcal{P}(x) \subseteq \mathcal{P}(y)$. ❑

Axiom VI. (Infinity). There is x such that

(2.1) $\emptyset \in x$ and, infty-0

(2.2) for any $y \in x$, $y \cup \{y\} \in x$. infty-1

Intuitively this axiom declares that there is an infinite set. Since the set x as in Axiom VI contains

$$\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\},...$$

as elements. In set theory these sets are considered to be the numbers 0, 1, 2, 3,... respectively.

In contrast to the other axioms introduced so far, the set x in Axiom VI is not determined uniquely. But we will define the canonical infinite set ω by using this Axiom VI and the next Axiom VII.

Axiom VII. (Separation). For any set u and any definite property $\varphi(x)$ about the set x which may contain some parameters other than x, there is a set v which consists of elements of u which satisfy $\varphi(x)$. That is, there is a set v such that, for any y, $y \in v$ iff $y \in u$ and $\varphi(y)$.

The set v above is unique up to u and $\varphi(x)$. We denote this set by $\{y \in u : \varphi(y)\}$.

Axioms I $\sim$ VII correspond to the axioms introduced by E. Zermelo (1871, Berlin – 1953, Freiburg im Breisgau) in his 1908 paper [3] where the first attempt of axiomatization of the set theory was done. Although the treatment of the axiom of infinity in [3] slightly differs from our Axiom VI and permits also objects which are not sets, we shall call the axiom system consisting of Axioms I $\sim$ VII the Zermelo set theory and denote it by Z.

"Z" in parenthesis in the following theorem indicates that the assertion of the theorem is provable in the theory Z.

Theorem 2.2 (Z) (a) *For any x, y, there is z which consists of common elements of x and y.* intersection

(b) *for any non-empty x, there is y which consists of common elements of all elements of x.*

(c) *Let P be a definite property (about a set) such that there is at least one set which satisfies P. Then there is a set y which consists of elements which are contained in all sets having the property P.*

Proof. (a): $z = \{u \in x : u \in y\}$ is such a set. Note that the set z exists by Axiom VII.

(b): Fix a $u \in x$. Then $y = \{v \in u : v \in u' \text{ for all } u' \in x\}$ is as desired.

(c): just the same argument as the proof for (b) works. Let u be a set which satisfies the property P. Then $y = \{v \in u$: for all u' if u' satisfies P then $v \in u'\}$ will do. ❑

The set z in Theorem 2.2, (a) is denoted by $x \cap y$. and y in Theorem 2.2, (b) by $\bigcap x$.

Theorem 2.3 (Z) *There is a set x_0 such that* existence-of-omega

(2.3) $\emptyset \in x_0$; omega-1

(2.4) *for any u, if $u \in x_0$ then $u \cup \{u\} \in x_0$.* omega-2

and such that x_0 is minimal with respect to $\subseteq$ among sets having properties (2.3) *and* (2.4).

Proof. By Theorem 2.2, (c). ❑

The unique set x_0 in Theorem 2.3 is considered to be the set of all natural numbers

(2.5) $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, \{\emptyset\}\}$, $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$,...

and denoted by ω (or $\mathbb{N}$).

E-1

Exercise 2.4 *Show that any x is an element of ω if and only if*

(2.6) *$x = \emptyset$ or $\emptyset \in x$;* omega-3

(2.7) *for all $u \in x$ and $v \in u$ we have $v \in x$;* omega-3-0

(2.8) *for all u, $u' \in x$, either $u = u'$ or $u \in u'$ or $u' \in u$;* omega-4

(2.9) *for all $y \subseteq x$ there is $u \in y$ such that $u \cap x = \emptyset$;* omega-4-0

(2.10) *for all $u \in x$, if $u \neq \emptyset$ then there is $u' \in x$ such that $u = u' \cup \{u'\}$;* omega-5

(2.11) *if $x \neq \emptyset$, then there is a $u \in x$ such that $x = u \cup \{u\}$.* omega-6

Hint. It is enough to show that (0): if x satisfies (2.6)$\sim$(2.11) and $u \in x$, then u satisfies (2.6)$\sim$(2.11). (1): for all x satisfying (2.6)$\sim$(2.11), we have $x \in \omega$. (2): $\{x \in \omega$: x satisfies (2.6) $\sim$ (2.11)$\}$ satisfies (2.1) and (2.2). ❑ (Exercise 2.4)

E-1-0

Exercise 2.5 (1) *For any m, $n \in \omega$, we have $m = n$ or $m \in n$ or $n \in m$.*

(2) *For any m, $n \in \omega$, $m \in n$ if and only if $m \subseteq n$.*

Hint. Use Exercise 2.4. ❑ (Exercise 2.5)

For x, y, we define the ordered pair $\langle x, y\rangle$ of x and y by

(2.12) $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$.

Theorem 2.6 (Z) *For any x, y, a, b, we have $\langle x, y\rangle = \langle a, b\rangle$ if and only if $x = a$ and $y = b$.* ordered-pair

Proof. If $x = a$ and $y = b$ then clearly we have $\langle x, y\rangle = \langle a, b\rangle$. So assume $\langle x, y\rangle = \langle a, b\rangle$ and we show that $x = a$ and $y = b$.

Case 1. $x = y$. Then $\langle x, y\rangle = \{\{x\}, \{x\}\} = \{\{x\}\}$. Thus, since $\langle x, y\rangle = \langle a, b\rangle$, we have $\{\{a\}, \{a, b\}\} = \{\{x\}\}$. It follows that $a = b = x$ and thus $x = a$ and $y = b$.

Case 2. $x \neq y$. In this case $\langle x, y\rangle$ contains two distinct elements and so $\langle a, b\rangle$ should also have two distinct elements and thus we have $a \neq b$. Since $\{\{x\}, \{x, y\}\} = \{\{a\}, \{a, b\}\}$, the unique element $\{x\}$ of $\{\{x\}, \{x, y\}\}$ consisting of one single element must be equal to the unique element $\{a\}$ of $\{\{a\}, \{a, b\}\}$ consisting of unique element. It follows that $x = a$ and $\{x, y\} = \{a, b\}$. Since $x \neq y$ and $a \neq b$ it follows that $y = b$. ❑

E-2

Exercise 2.7 *Show that $\{x, y\}$ in general does not have the property corresponding to Theorem 2.6.* ❑

By Theorem 2.6 $\langle x, y\rangle$ codes the pair of x and y including the information about the first and second components being x and y. $\langle x, y\rangle$ is thus called the ordered pair of x and y.

We have $\{x\}$, $\{x, y\} \subseteq \{x, y\}$ and thus $\{x\}$, $\{x, y\} \in \mathcal{P}(\{x, y\})$. It follows that $\langle x, y\rangle = \{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(\{x, y\})$ or

(2.13) $\quad \langle x, y\rangle \in \mathcal{P}(\mathcal{P}(\{x, y\}))$. pair-0

If $x \in X$ and $y \in Y$ then

(2.14) $\quad \mathcal{P}(\mathcal{P}(\{x, y\})) \subseteq \mathcal{P}(\mathcal{P}(X \cup Y))$ pair-1

by Lemma 2.1. By (2.13) and (2.14), it follows that $\langle x, y\rangle \in \mathcal{P}(\mathcal{P}(X \cup Y))$ for all $x \in X$ and $y \in Y$. Thus the set

(2.15) $\quad X \times Y = \{u \in \mathcal{P}(\mathcal{P}(X \cup Y))$: u consists of a singleton of an element of X and a two element set consisting of an element of X and an element of Y $\}$.

consists of all pairs $\langle x, y\rangle$ for $x \in X$ and $y \in Y$. $X \times Y$ is called the Cartesian product of $X \times Y$.

A subset f of $X \times Y$ is said to be a mapping from X to Y if, for any $x \in X$, there is a unique $y \in Y$ such that $\langle x, y\rangle \in f$. If f is a function from X to Y, this is denoted by $f : X \to Y$. For f with $f : X \to Y$ and $x \in X$ the unique $y \in Y$ with $\langle x, y\rangle \in f$ is denoted by $f(x)$.

For $f : X \to Y$ we say that X is the domain of f (notation: $X = \mathrm{dom}(f)$) and Y is the range of f. The range of a function f is not unique since $f : X \to Y$ and $Y \subseteq Y'$ implies $f : X \to Y'$. For $f : X \to Y$ and $X' \subseteq X$, the image of X' by f is the set

(2.16) $\quad f''X' = \{y \in Y :$ there is $x \in X'$ such that $f(x) = y\}$.

$f''\,\mathrm{dom}(f)$ is called the range of f and is also denoted by range(f).

Usual notions about functions can be defined also here naturally. For example a function $f : X \to Y$ is one to one (or injective) if, for any two distinctive x, $x' \in X$, we have $f(x) \neq f(x')$. f is onto (or surjective) if $f''X = Y$. f is one to one onto (or bijective, in this case we say also frequently that f is a bijection) if f is both one to one and onto.

E-2-0

Exercise 2.8 *Suppose that $\mathcal{F}$ is a set consisting of functions such that for any f, $g \in \mathcal{F}$ we have always either $f \subseteq g$ or $g \subseteq f$. Then $\bigcup \mathcal{F}$ is again a function from X to Y where $X = \bigcup\{\mathrm{dom}(f) \,:\, f \in \mathcal{F}\}$ and $Y = \bigcup\{\mathrm{range}\, f \,:\, f \in \mathcal{F}\}$.* ❑

Example 2.1 *In Section **??**, we see how the arithmetical functions like addition, multiplication etc. can be defined on ω and how the basic mathematics can be formulated in Z using these functions. The first step toward the definition of such functions is definition of the successor function S on ω where $S : \omega \to \omega$ is defined by $S = \{\langle n, n \cup n\rangle \in \omega \times \omega \,:\, n^2\}$.*

For $n \in \omega$, $S(n)$ is sometimes also denoted as $n + 1$.

E-3

Exercise 2.9 (1) *$S : \omega \to \omega$ is one to one.*

(2) *$S''\omega = \omega \setminus \{\emptyset\}$.*

Hint. By Exercise 2.4. ❑ (Exercise 2.9)

For any sets X, Y, there is the set ${}^XY$ consisting of all functions from X to Y:

(2.17) $\quad {}^XY = \{f \in \mathcal{P}(X \times Y) \,:\, f : X \to Y\}.$

Exercise 2.10 *Answer which sets the following are:* (a) *${}^{\emptyset}Y$ for any Y;* (b) *${}^X\emptyset$ for nonempty X;* (c) *${}^X\{\emptyset\}$.*

Exercise 2.11 (1) *For X, Y and Y' with $Y \subseteq Y'$ we have ${}^XY \subseteq {}^{X'}Y$.*

(2) *For X, X' with $X \neq X'$ and Y we have ${}^XY \cap {}^{X'}Y = \emptyset$.*

The following Axiom VIII enables certain arguments indispensable in modern mathematics.

Axiom VIII. (Axiom of Choice). For any x with $\emptyset \notin x$ there is a mapping $f : x \to \bigcup x$ such that$f(u) \in u$ for all $u \in x$.

This Axiom is also called AC for short. The mapping f as above is called a choice function on x.

The system of axioms consisting of Axioms I $\sim$ VII is (the modern version of) the Zermelo set theory[1] and denote it by Z. The system of axioms Z + Axiom VIII is the Zermelo set theory with the Axiom of Choice (AC) and denoted by ZC.

The next Axiom IX is rather of technical nature. Its meaning will be clear in later sections. Intuitively, the Axiom IX asserts that any set has the property (2.9) of subsets of natural numbers.

Axiom IX. (Regularity). For any nonempty set x there is $y \in x$ such that $y \cap x = \emptyset$.

We shall call Axiom IX also Axiom of Regularity and denote it by AR. AR excludes some pathological sets.

Theorem 2.12 (Z + AR) *There is no x such that $x \in x$.* x-in-x

Proof. Suppose otherwise and let x be such that $x \in x$. Let $x_0 = \{y \in x : y \in y\}$. Note that, since $x \in x$, we have $x \in x_0$, and in particular $x_0 \neq \emptyset$. By Regularity, there is $y \in x_0$ such that $y \cap x_0 = \emptyset$. But since $y \in x_0$, we have $y \in y$ and hence $y \in y \cap x_0$. This is a contradiction. ❑ (Theorem 2.12)

Under AC, AR is equivalent to the non-existence of descending $\in$-chain:

Theorem 2.13 (ZC) AR *is equivalent to the assertion:*

(2.18) *there is no function f on ω such that $f(n+1) \in f(n)$ for all $n \in \omega$.* ar-0

Proof. [2] Suppose first that (2.18) fails, that is, that there is a function f on ω such that $f(n+1) \in f(n)$ for all $n \in \omega$. Then $x = f''\omega$ is a counterexample to AR (Exercise).

Suppose now that x is a counterexample to AR. Let $c : \mathcal{P}(x) \setminus \{\emptyset\} \to x$ be a choice function. Let

(2.19) $$\begin{aligned}\mathcal{F} = \{h \in \mathcal{P}(\omega \times x) \,:\, &(\alpha)\ \ h : n \to x \text{ for some } n;\\ &(\beta)\ \ \text{if } 0 \in n \text{ then } h(0) = g(x);\ \text{and}\\ &(\gamma)\ \ \text{for all } m \in n\\ &\qquad \text{if } m+1 \in n \text{ then } h(m+1) = g(h(m) \cap x)\}\end{aligned}$$ ar-1

We show that $f = \bigcup \mathcal{F}$ is a counterexample to (2.18).

C-ar-1

Claim 2.13.1 *$\mathcal{F} \neq \emptyset$ and $\mathcal{F}$ contains a non-empty set.*

$\vdash$ $\{\langle \emptyset, g(x) \rangle\} \in \mathcal{F}$ by (2.19)(β). $\dashv$ (Claim 2.13.1)

C-ar-2

---

[1] The main differences between what we call here Z and the system Zermelo introduced in 1908 paper [3] is that (1) the axioms in [3] do not exclude the existence of so called urelements which are objects without elements but different from the empty set; and (2) the Axiom of Infinity in [3] asserts the existence of a set different from the set x in our Axiom VI.

[2] The formulation of this proof is somewhat awkward since we are proving the theorem directly without using the theorem on recursive definition of functions on ω (Theorem **??**) which is introduced later in Section **??**. See also Exercise **??**.

Claim 2.13.2 *For all h, $h' \in \mathcal{F}$ either $h \subseteq h'$ or $h' \subseteq h$ holds.*

$\vdash$ Suppose that h, $h' \in \mathcal{F}$. By Exercise 2.5 we may assume that $\mathrm{dom}(h) \subseteq \mathrm{dom}(h')$. It is enough to show that $h \subseteq h'$. Suppose that is not the case then, by Exercise 2.4, (2.9), there is $n \in \mathrm{dom}(h)$ such that $h(n) \neq h'(n)$ and such that n is minimal with respect to $\in$ among such n with this property.

$n \neq \emptyset$, by (2.19), (β). So by Exercise 2.9 (2), there is $m \in \mathrm{dom}(h)$ such that $n = m+1$. By the minimality of n we have

$$\begin{aligned} h(n) &= g(h(m) \cap x) \\ &= g(h'(m) \cap x) \\ &= h'(n) \end{aligned}$$

This is a contradiction. $\dashv$ (Claim 2.13.2)

By Claim 2.13.2 and Exercise 2.8, $f = \bigcup \mathcal{F}$ is a function.

Claim 2.13.3 $\mathrm{dom}(f) = \omega$.

$\vdash$ Suppose that $\mathrm{dom}(f) \neq \omega$. Then $\mathrm{dom}(f) = n$ for some $n \in \omega$ (This can be shown by using Exercise 2.4, (2.9)). $n \neq \emptyset$ by Claim 2.13.1. Hence, by Exercise 2.9, (2), and ExerciseofE-1, (2.10), there is $m^* \in n$ such that $n = m^* + 1$. $f(m^*) \cap x \neq \emptyset$ since x is a counterexample to AR. Thus $f' = f \cup \{\langle m^*, g(f(m^*) \cap x)\rangle\}$ is well defined and $f' \in \mathcal{F}$. Since $\mathrm{dom}(f) \subsetneq \mathrm{dom}(f')$, this is a contradiction to the definition of f. ❏ (Theorem 2.13)

The last Axiom we consider here asserts that for any mapping the image of a set by the mapping is a set.

Axiom X. (Replacement). For any set u and any definite property $\varphi(x, y)$ of sets which may contain some parameters other than x and y, and such that,

(2.20) for any a, there is a unique b such that $\varphi(a, b)$, repl-0

there is a set v such that $b \in v$ if and only if there is $a \in x$ such that $\varphi(a, b)$.

Thus Axiom X asserts that there is the set $\{b : \text{there is } a \in u \text{ such that } \varphi(a, b)\}$ if φ satisfies (2.20).

The Axiom of Replacement (Axiom X) was suggested by A. Fraenkel. It is easy to see that Axiom of Separation (Axiom VII) follows from this axiom (Exercise). The axiom system consisting of Axioms I $\sim$ VII + IV + X is called the Zermelo-Fraenkel set theory and denoted with ZF. The axiom system ZF + Axiom VIII is called the Zermelo-Fraenkel set theory with Axiom of Choice and denoted with ZFC. Later we shall prove that ZF (ZFC resp.) is strictly stronger than Z (ZC resp.)

3 Axiom of Choice and cardinality in Z

AC-in-Z

For any set X a set $R \subseteq X^2 = X \times X$ is said to be a binary relation on X and, for $\langle x, y \rangle \in X^2$, "$\langle x, y \rangle \in R$" is also denoted by "$x \, R \, y$".

A binary relation $\sqsubseteq$ on X is a partial ordering on X if, for any x, y, $z \in X$, we have

(3.1) $x \sqsubseteq x$; ac-a

(3.2) if $x \sqsubseteq y$ and $y \sqsubseteq x$ then $x = y$; and ac-0

(3.3) if $x \sqsubseteq y$ and $y \sqsubseteq z$ then $x \sqsubseteq z$. ac-1

If $\sqsubseteq$ is a partial ordering on X, the pair $\langle X, \sqsubseteq \rangle$ is said to be a partially ordered set. If it is clear from the context which $\sqsubseteq$ is meant, we shall also say that X is a partially ordered set.

Functions from X to itself are binary relations on X. Similarly to the composition and the inverse of functions we can define the composition and the inverse of binary relations:

For a binary relation $R, S \subseteq X^2$ on X, let

(3.4) $R^{-1} = \{\langle y, x \rangle \in X^2 : \langle x, y \rangle \in R\}$; ac-2

(3.5) $S \circ R = \{\langle x, z \rangle \in X^2 :$ there is $y \in X$ such that $\langle x, y \rangle \in R$ and $\langle y, z \rangle \in S\}$. ac-3

L-ac-1

Lemma 3.1 (1) *For $f : X \to X$, $f^{-1} : X \to X$ if and only if f is a bijection.*

(2) *For $f : X \to X$ and $g : X \to X$ we have $g \circ f : X \to X$.*

Proof. Exercise. ❑ (Lemma 3.1)

For a binary relation $R \subseteq X^2$ on a set X and $Y \subseteq X$ we define

(3.6) $R|Y = R \cap Y^2$. ac-4

L-ac-2

Lemma 3.2 (1) *For $f : X \to X$ and $Y \subseteq X$, $f|Y : Y \to Y$ if and only if Y is closed under f. If this is the case we have $f|Y = f \restriction Y$.*

(2) *If $\sqsubseteq$ is a partial ordering on X then, for any $Y \subseteq X$, $\sqsubseteq |Y$ is a partial ordering on Y.*

A partial ordering $\sqsubseteq$ on X is said to be a linear ordering on X is for any two distinct x, $x' \in X$ either $x \sqsubseteq x'$ or $x' \sqsubseteq x$ holds. Note that, in this case, $x \sqsubseteq x'$ and $x' \sqsubseteq x$ cannot hold simultaneously by (3.2). If $\sqsubseteq$ is a linear ordering on X, the pair $\langle X, \sqsubseteq \rangle$ is called a linearly ordered set. If it is clear from the context which $\sqsubseteq$ is meant, we drop $\sqsubseteq$ and simply say that X is a linearly ordered set.

For a partial ordering $\sqsubseteq$ on X, a subset C of X is said to be a chain (with respect to $\sqsubseteq$) if $\sqsubseteq |C$ is a linear ordering on Y.

For a partially ordered set $\langle X, \sqsubseteq \rangle$ and $Y \subseteq X$, $y \in Y$ is a *minimal element* of Y if there is no $y' \in Y$ such that $y' \sqsubseteq y$ and $y' \neq y$. $y \in Y$ is a *maximal element* of Y if there is no $y' \in Y$ such that $y \sqsubseteq y'$ and $y' \neq y$.

A chain C in X with respect to $\sqsubseteq$ is bounded if there is some $x \in X$ such that $y \sqsubseteq x$ holds for all $y \in C$. Such x is called an upper bound of C. C is strictly bounded if there is an upper bound x of C in $X \setminus C$.

If $C \subseteq X$ is a chain, a minimal element of C (if it exists) is the unique smallest element of C and a maximal element of C (if it exists) is the unique largest element of C.

The following assertion is called the Zorn's Lemma or Kuratowski-Zorn's Lemma after Max Zorn (1906–1993) and Kazimierz Kuratowski (1896–1980):

(ZL) For any partially ordered set $\langle X, \sqsubseteq \rangle$, if every chain C of X is bounded, then X has the maximal element.

In the following we show that ZL is equivalent to AC over Z. In the proof of the equivalence, the equivalence of AC and ZL with the following Well-ordering Theorem (WO) is also established.

A linearly ordered set $\langle X, \sqsubseteq \rangle$ is a well-ordered set if, for any non empty $Y \subseteq X$, there is the minimal element of Y. Such $\sqsubseteq$ is said to be a well-ordering on X.

E-4

Exercise 3.3 *$E = \{\langle m, n\rangle \in \omega^2 : m \in n\}$ is a well-ordering on ω.*

Proof. Use Exercise 2.5 and Exercise 2.4. □ (Exercise 3.3)

(WO) For any set X, there is a well-ordering $\sqsubseteq$ on X.

L-ac-3

Theorem 3.4 AC*,* ZL *and* WO *are equivalent to each other over* Z*.*

For the proof of Theorem 3.4, we need the following Lemma on well-orderings. For a partially ordered set $\langle X, \sqsubseteq \rangle$, $Y \subseteq X$ and $x \in X$, we denote

$$Y \downarrow_{\sqsubseteq} x = \{x' \in Y : x' \sqsubseteq x\},$$
$$Y \not\downarrow_{\sqsubseteq} x = \{x' \in Y : x' \sqsubseteq x, x' \neq x\}.$$

If it is clear which partial ordering $\sqsubseteq$ is meant, we drop it and write $Y \downarrow x$ and $Y \not\downarrow x$.

$Y \subseteq X$ is an initial segment of X (with respect to $\sqsubseteq$) if, for any $y \in Y$ and $x \in X$, $x \sqsubseteq y$ implies $x \in Y$. $X \downarrow x$ and $X \not\downarrow x$ are initial segments of X for any $x \in X$. For two partially ordered sets $\langle X, \sqsubseteq_X \rangle$ and $\langle Y, \sqsubseteq_Y \rangle$, we say that $\langle Y, \sqsubseteq_Y \rangle$ is an initial segment of $\langle X, \sqsubseteq_X \rangle$ if $Y \subseteq X$, Y is an initial segment of X with respect to $\sqsubseteq_X$ and $\sqsubseteq_Y = \sqsubseteq_X |Y$.

L-ac-4

Lemma 3.5 (1) *Suppose that $\langle X, \sqsubseteq \rangle$ is a well-ordered set and x_0 a set with $x \notin X$. Then $\langle X', \sqsubseteq' \rangle$ is also a well-ordered set where $X' = X \cup \{x_0\}$ and $\sqsubseteq' = \sqsubseteq \cup (X' \times \{x_0\})$.*

(2) *Suppose that $\mathcal{A}$ is a family of well-ordered sets*[3] *such that for any $\langle X, \sqsubseteq_X\rangle$, $\langle Y, \sqsubseteq_Y\rangle \in \mathcal{A}$, one of them is an initial segment of the other. Then $\langle X_0, \sqsubseteq_0\rangle$ is also a well-ordered set where $X_0 = \bigcup\{X : \langle X, \sqsubseteq\rangle \in \mathcal{A}$ for some $\sqsubseteq\}$ and $\sqsubseteq_0 = \bigcup\{\sqsubseteq : \langle X, \sqsubseteq\rangle \in \mathcal{A}$ for some $X\}$.*[4]

Proof. Exercise. ❑ (Lemma 3.5)

Proof of Theorem 3.4:

(AC) $\Rightarrow$ (ZL): The following proof is taken from [2].

Assume that AC holds. Toward a contradiction, suppose that there is a partially ordered set $\langle X, \sqsubseteq\rangle$ such that

(3.7) every chain $C \subseteq X$ is bounded; but ac-5

(3.8) X does not have any maximal element. ac-6

First, it is easy to see that every chain $C \subseteq X$ has a strict upper bound: if $C \subseteq X$ is a chain, then there is a bound x of C by (3.7). Since x is not a maximal element in X by (3.8), there is $x' \in X$ such that $x \sqsubseteq x'$ and $x \neq x'$. This x' is a strict upper bound of C (Exercise).

Let $\mathcal{X} = \{C \in \mathcal{P}(X) : C$ is a chain in $X\}$. By AC, there is a mapping $f : \mathcal{X} \to X$ such that, for all $C \in \mathcal{X}$, $f(C)$ is a strict upper bound of C.[5]

Let

(3.9) $\mathcal{F} = \{C \in \mathcal{P}(X) : C$ is a well-ordered chain with $f(C \downarrow c) = c$ for all $c \in C\}$. ac-7

C-ac-1

Claim 3.5.1 (1) $\mathcal{F} \neq \emptyset$.

(2) *For any C, $C' \in \mathcal{F}$, one of them is an initial segment of the other with respect to $\sqsubseteq$.*

$\vdash$ (1): This is clear since $\emptyset \in \mathcal{F}$.

(2): Suppose that C, $C' \in \mathcal{F}$ are such that neither of them is an initial segment of the other. In particular, we have $C \neq C'$. Without loss of generality, we may assume $C \setminus C' \neq \emptyset$. Let c_0 be the $\sqsubseteq$-minimal element $C \setminus C'$. Then we have $C \downarrow c_0 \subseteq C' \downarrow c_0$.

Suppose that $C \downarrow c_0 = C' \downarrow c_0$. If $C' \downarrow c_0 = C'$ then C' is an initial segment of C. This is a contradiction to our assumption. Thus $C' \setminus (C' \downarrow c_0) \neq \emptyset$. Let c_0' be the $\sqsubseteq$-minimal

---

[3] This just means that $\mathcal{A}$ is a set such that, for all $x \in \mathcal{A}$, x is a well-ordered set.

[4] Exercise: show that $\{X : \langle X, \sqsubseteq\rangle \in \mathcal{A}$ for some $\sqsubseteq\}$ is a set in Z.

[5] For each chain $C \in \mathcal{X}$, let $B_C = \{x \in X : x$ is a strict upper bound of $C\}$. By the argument above we have $B_C \neq \emptyset$ for all $C \in \mathcal{X}$. Let $\mathcal{Y} = \{B_C \in \mathcal{P}(X) : C \in \mathcal{X}\}$. By AC there is a choice function $g : \mathcal{Y} \to X$ of $\mathcal{Y}$. The mapping $f : \mathcal{X} \to X$ defined by $f(C) = g(B_C)$ is then as desired.

Note that f is also defined for the initial segment $\emptyset \subseteq X$.

element of this set. Then we have $c_0 = f(C \downarrow c_0) = f(C' \downarrow c_0) = f(C' \downarrow c'_0) = c'_0$. In particular $c_0 \in C'$. This is a contradiction to the choice of c_0.

Thus we have $C \downarrow c_0 \subsetneq C' \downarrow c_0$. Let c_1 be the minimal element of $(C' \downarrow c_0) \setminus (C \downarrow c_0)$. Then again we obtain contradictions in all cases arguing similarly to above (Exercise).

$\dashv$ (Claim 3.5.1)

Let $C^* = \bigcup \mathcal{F}$. By Claim 3.5.1 and Lemma 3.5, (2), $C^* \in \mathcal{F}$. Let $x^* = f(C^*)$ and let $C^{**} = C^* \cup \{x\}$. By Lemma 3.5, (1), we have $C^{**} \in \mathcal{F}$. Since $C^* \subsetneq C^{**}$ this is a contradiction to the definition of C^*.

$(\mathsf{ZL}) \Rightarrow (\mathsf{WO})$: Assume ZL. and Suppose that X is an arbitrary set. We have to show that there is a well-ordering on X. Let

(3.10) $\quad \mathcal{F} = \{\mathcal{Y} : \mathcal{Y} = \langle Y_{\mathcal{Y}}, \sqsubseteq_{\mathcal{Y}} \rangle$ is a well ordered set with $Y_{\mathcal{Y}} \subseteq X\}$

For $\mathcal{Y}$, $\mathcal{Y}' \in \mathcal{F}$ with $\mathcal{Y} = \langle Y_{\mathcal{Y}}, \sqsubseteq_{\mathcal{Y}} \rangle$ and $\mathcal{Y}' = \langle Y_{\mathcal{Y}'}, \sqsubseteq_{\mathcal{Y}'} \rangle$ we define

(3.11) $\quad \mathcal{Y} \preccurlyeq \mathcal{Y}' \;:\Leftrightarrow\; \mathcal{Y}$ is an initial segment of $\mathcal{Y}'$

Then it is easy to show that $\preccurlyeq$ is a partial ordering on $\mathcal{F}$ (Exercise). By Lemma 3.5, (2), any chain $\mathcal{C} \subseteq \mathcal{F}$ has its upper bound $\langle Y^{\mathcal{C}}, \sqsubseteq^{\mathcal{C}} \rangle$ where

(3.12) $\quad Y^{\mathcal{C}} = \bigcup\{Y : \langle Y, \sqsubseteq \rangle \in \mathcal{C}$ for some $\sqsubseteq\}$ and
$\sqsubseteq^{\mathcal{C}} = \bigcup\{\sqsubseteq : \langle Y, \sqsubseteq \rangle \in \mathcal{C}$ for some $Y\}$.

By ZL, there is a maximal element $\langle Y^*, \sqsubseteq^* \rangle \in \mathcal{F}$. We are done by showing that Y^*. Suppose that this is not the case and let $x^* \in X \setminus Y^*$. Let $Y^{**} = Y^* \cup \{x^*\}$ and $\sqsubseteq^{**} = \sqsubseteq^* \cup (Y^{**} \times \{x^*\})$. Then we have $\langle Y^{**}, \sqsubseteq^{**} \rangle \in \mathcal{F}$ by Lemma 3.5, (1) and $\langle Y^*, \sqsubseteq^* \rangle \preccurlyeq \langle Y^{**}, \sqsubseteq^{**} \rangle$. This is a contradiction to the maximality of $\langle Y^*, \sqsubseteq^* \rangle$.

$(\mathsf{WO}) \Rightarrow (\mathsf{AC})$: Assume WO and let X be such that $\emptyset \notin X$. Let $\sqsubseteq$ be a well-ordering on $\bigcup X$ and define $f : X \to \bigcup X$ by

(3.13) $\quad f := \{\langle x, y \rangle \in X \times \bigcup X : y$ is the minimal element of X with respect to $\sqsubseteq\}$.

Then f is a choice function on X. $\square$ (Theorem 3.4)

The next Theorem 3.6 is an application of Theorem 3.4. We show later that the assertion of the Theorem 3.6 is equivalent to AC over ZF.

Theorem 3.6 (ZC) *For any X, Y, there is a one-to-one function from one of X and Y to the other.* injection

Proof. Let $\mathcal{F} = \{f : f$ is a one-to-one function from a subset of X to $Y\}$. $\langle \mathcal{F}, \subseteq \rangle$ is a partially ordered set satisfying the condition in ZL (Exercise).

By ZL (which is equivalent to AC by Theorem 3.4), there is a maximal element f^* in $\mathcal{F}$ with respect to $\subseteq$.

We show that either $\mathrm{dom}(f^*) = X$ or $\mathrm{range}(f^*) = Y$. If $\mathrm{dom}(f^*)$ then we are done and if $\mathrm{range}(f^*)$ then $(f^*)^{-1}$ is a one-to-one function from Y to X.

Suppose that $\mathrm{dom}(f^*) \neq X$ and $\mathrm{range}(f^*) \neq Y$. Then, letting $x^* \in X \setminus \mathrm{dom}(f^*)$ and $y^* \in Y \setminus \mathrm{range}(f^*)$, $f^{**} = f^* \cup \{\langle x^*, y^*\rangle\}$ is an element of $\mathcal{F}$ extending f^*. This is a contradiction to the maximality of f^*. ❑ (Theorem 3.6)

For some other theorems in set theory, the first proofs with AC could be replaced later with proofs without AC. The following theorems are also such examples though, when Cantor proved assertions corresponding to these theorems, the Axiom of Choice was not yet singled out as a separate set-theoretic principle.

Theorem 3.7 (Z) *Suppose that $f : X \to X$ is one-to-one. If $f''X \subseteq Y \subseteq X$, then there is a bijection from X to Y.* T-cantor-0

Proof. The following proof is taken from [3]. Let

$$Q = Y \setminus f''X \tag{3.14}$$ ac-7-0

and let

$$\mathcal{F} = \{Z \in \mathcal{P}(X) \,:\, Q \subseteq Z \text{ and } f''Z \subseteq Z\}\ . \tag{3.15}$$ ac-8

$\mathcal{F} \neq \emptyset$ since we have, for example, $X \in \mathcal{F}$. Let $Z_0 = \bigcap \mathcal{F}$. Then clearly we have:

C-cantor-0-1

Claim 3.7.1 (1) *$Q \subseteq Z_0$;*

(2) *$f''Z_0 \subseteq Z_0$.*

$\vdash$ Exercise. $\dashv$ (Claim 3.7.1)

Hence Z_0 is the minimal element of $\mathcal{F}$ (with respect to $\subseteq$).

Note that by (1) and (2) of Claim 3.7.1, we have

$$f''Q \subseteq f''Z_0 \text{ and} \tag{3.16}$$ ac-9

$$f''(f''Z_0) \subseteq f''Z_0. \tag{3.17}$$ ac-10

C-cantor-0-2

Claim 3.7.2 *$Z_0 = Q \,\dot\cup\, f''Z_0$.*

$\vdash$ Q and $f''Z_0$ are disjoint since $f''Z_0 \subseteq f''X$ and, Q and $f''X$ are disjoint by (3.14). $Q \subseteq Q \,\dot\cup\, f''Z_0$ and $f''(Q \,\dot\cup\, f''Z_0) \subseteq f''Z_0$ by (3.16) and (3.17). Hence, by the minimality of Z_0, $Z_0 \subseteq Q \,\dot\cup\, f''Z_0$.

On the other hand, we also have $Q \,\dot\cup\, f''Z_0 \subseteq Z_0$ by Claim 3.7.1, (1) and (2).

$\dashv$ (Claim 3.7.2)

Now let $Y = f''X \,\dot\cup\, Q$. Y can be also represented as

$$Y = (f''X \setminus f''Z_0) \,\dot\cup\, (Q \,\dot\cup\, f''Z_0). \tag{3.18}$$ ac-11

Let $g : Y \rightarrow f''X$ be defined by

$$(3.19) \quad g(x) = \begin{cases} x, & \text{if } x \in g''X \setminus f''Z_0 \\ f(x), & \text{otherwise.} \end{cases}$$

Since $f \restriction Q \,\dot{\cup}\, f''Z_0 : Q \,\dot{\cup}\, f''Z_0 \rightarrow f''Z_0$ Claim 3.7.2 and this mapping is one-to-one. Hence g is also one-to-one.

Let $h = g^{-1} \circ f$. Then $h : X \rightarrow Y$ is a bijection. ❏ (Theorem 3.7)

Theorem 3.8 (Cantor-Bernstein-Schroeder Theorem) (Z) *Suppose the $f : X \rightarrow Y$ and $g : Y \rightarrow X$ are one-to-one functions. Then there is a bijection from X to Y.* cantor-bernstein-schroeder

Proof. Let $h = g \circ f$. Then h is one-to-one and $h''X = g''(f''X) \subseteq g''Y \subseteq X$. Hence there is a bijection $h_1 : X \rightarrow g''Y$ by Theorem 3.7. $h_2 = g^{-1} \circ h_1$ is then a bijection from X to Y. ❏ (Theorem 3.8)

4 First-order predicate logic and strict axiomatization of set theory

logic

References

the-refs

[1] S. Fuchino, 数理論理学, lecture notes of “数理論理学” and “数理論理学特論” given in Spring Semester 2012.
http://kurt.scitec.kobe-u.ac.jp/~fuchino/kobe/predicate-logic-ss11.pdf

[2] Jonathan W. Lewin, A simple proof of Zorn’s Lemma, The American Mathematical Monthly, 98(4), (1991), 353–354.

[3] Ernst Zermelo, Untersuchungen über die Grundlagen der Mengenlehre. I, Mathematische Annalen 65 (1908), 261–281.