1. Axioms of Set Theory

Axioms of Zermelo-Fraenkel

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

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

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

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

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

1.6. *Axiom of Infinity.* *There exists an infinite set.*

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

1.8. *Axiom of Regularity.* *Every nonempty set has an $\in$-minimal element.*

1.9. *Axiom of Choice.* *Every family of nonempty sets has a choice function.*

The theory with axioms 1.1–1.8 is the Zermelo-Fraenkel axiomatic set theory ZF; ZFC denotes the theory ZF with the Axiom of Choice.

Why Axiomatic Set Theory?

Intuitively, a set is a collection of all elements that satisfy a certain given property. In other words, we might be tempted to postulate the following rule of formation for sets.

1.10. *Axiom Schema of Comprehension* (*false*). *If P is a property, then there exists a set $Y = \{x : P(x)\}$.*

This principle, however, is false:

1.11. Russell's Paradox. Consider the set S whose elements are all those (and only those) sets that are not members of themselves: $S = \{X : X \notin X\}$. Question: Does S belong to S? If S belongs to S, then S is not a member of itself, and so $S \notin S$. On the other hand, if $S \notin S$, then S belongs to S. In either case, we have a contradiction.

Thus we must conclude that

$$\{X : X \notin X\}$$

is not a set, and we must revise the intuitive notion of a set.

The safe way to eliminate paradoxes of this type is to abandon the Schema of Comprehension and keep its weak version, the *Schema of Separation*:

If P is a property, then for any X there exists a set $Y = \{x \in X : P(x)\}$.

Once we give up the full Comprehension Schema, Russell's Paradox is no longer a threat; moreover, it provides this useful information: The set of all sets does not exist. (Otherwise, apply the Separation Schema to the property $x \notin x$.)

In other words, it is the concept of the set of all sets that is paradoxical, not the idea of comprehension itself.

Replacing full Comprehension by Separation presents us with a new problem. The Separation Axioms are too weak to develop set theory with its usual operations and constructions. Notably, these axioms are not sufficient to prove that, e.g., the union $X \cup Y$ of two sets exists, or to define the notion of a real number.

Thus we have to add further construction principles that postulate the existence of sets obtained from other sets by means of certain operations.

The axioms of ZFC are generally accepted as a correct formalization of those principles that mathematicians apply when dealing with sets.

Language of Set Theory, Formulas

The Axiom Schema of Separation as formulated above uses the vague notion of a *property*. To give the axioms a precise form, we develop axiomatic set theory in the framework of the first order predicate calculus. Apart from the equality predicate $=$, the language of set theory consists of the binary predicate $\in$, the *membership relation*.

The *formulas* of set theory are built up from the *atomic formulas*

$$x \in y, \qquad x = y$$

by means of *connectives*

$$\varphi \wedge \psi, \qquad \varphi \vee \psi, \qquad \neg\varphi, \qquad \varphi \to \psi, \qquad \varphi \leftrightarrow \psi$$

(conjunction, disjunction, negation, implication, equivalence), and *quantifiers*

$$\forall x\, \varphi, \qquad \exists x\, \varphi.$$

In practice, we shall use in formulas other symbols, namely defined predicates, operations, and constants, and even use formulas informally; but it will be tacitly understood that each such formula can be written in a form that only involves $\in$ and $=$ as nonlogical symbols.

Concerning formulas with free variables, we adopt the notational convention that all free variables of a formula

$$\varphi(u_1, \ldots, u_n)$$

are among $u_1, \ldots, u_n$ (possibly some u_i are not free, or even do not occur, in φ). A formula without free variables is called a *sentence*.

Classes

Although we work in ZFC which, unlike alternative axiomatic set theories, has only one type of object, namely sets, we introduce the informal notion of a *class*. We do this for practical reasons: It is easier to manipulate classes than formulas.

If $\varphi(x, p_1, \ldots, p_n)$ is a formula, we call

$$C = \{x : \varphi(x, p_1, \ldots, p_n)\}$$

a *class*. Members of the class C are all those sets x that satisfy $\varphi(x, p_1, \ldots, p_n)$:

$$x \in C \quad \text{if and only if} \quad \varphi(x, p_1, \ldots, p_n).$$

We say that C is *definable from* $p_1, \ldots, p_n$; if $\varphi(x)$ has no parameters p_i then the class C is *definable*.

Two classes are considered equal if they have the same elements: If

$$C = \{x : \varphi(x, p_1, \ldots, p_n)\}, \qquad D = \{x : \psi(x, q_1, \ldots, q_m)\},$$

then $C = D$ if and only if for all x

$$\varphi(x, p_1, \ldots, p_n) \leftrightarrow \psi(x, q_1, \ldots, q_m).$$

The *universal class*, or *universe*, is the class of all sets:

$$V = \{x : x = x\}.$$

We define *inclusion* of classes (C is a *subclass* of D)

$$C \subset D \text{ if and only if for all } x,\ x \in C \text{ implies } x \in D,$$

and the following operations on classes:

$$\begin{aligned} C \cap D &= \{x : x \in C \text{ and } x \in D\}, \\ C \cup D &= \{x : x \in C \text{ or } x \in D\}, \\ C - D &= \{x : x \in C \text{ and } x \notin D\}, \\ \bigcup C &= \{x : x \in S \text{ for some } S \in C\} = \bigcup\{S : S \in C\}. \end{aligned}$$

Every set can be considered a class. If S is a set, consider the formula $x \in S$ and the class

$$\{x : x \in S\}.$$

That the set S is uniquely determined by its elements follows from the Axiom of Extensionality.

A class that is not a set is a *proper class*.

Extensionality

If X and Y have the same elements, then $X = Y$:

$$\forall u\,(u \in X \leftrightarrow u \in Y) \rightarrow X = Y.$$

The converse, namely, if $X = Y$ then $u \in X \leftrightarrow u \in Y$, is an axiom of predicate calculus. Thus we have

$$X = Y \quad \text{if and only if} \quad \forall u\,(u \in X \leftrightarrow u \in Y).$$

The axiom expresses the basic idea of a set: A set is determined by its elements.

Pairing

For any a and b there exists a set $\{a, b\}$ that contains exactly a and b:

$$\forall a\,\forall b\,\exists c\,\forall x\,(x \in c \leftrightarrow x = a \lor x = b).$$

By Extensionality, the set c is unique, and we can define the *pair*

$$\{a, b\} = \text{the unique } c \text{ such that } \forall x\, (x \in c \leftrightarrow x = a \vee x = b).$$

The *singleton* $\{a\}$ is the set

$$\{a\} = \{a, a\}.$$

Since $\{a, b\} = \{b, a\}$, we further define an *ordered pair*

$$(a, b)$$

so as to satisfy the following condition:

$$(a, b) = (c, d) \text{ if and only if } a = c \text{ and } b = d. \tag{1.1}$$

For the formal definition of an ordered pair, we take

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

We leave the verification of (1.1) to the reader (Exercise 1.1).

We further define ordered triples, quadruples, etc., as follows:

$$\begin{aligned}(a, b, c) &= ((a, b), c),\\ (a, b, c, d) &= ((a, b, c), d),\\ &\vdots\\ (a_1, \dots, a_{n+1}) &= ((a_1, \dots, a_n), a_{n+1}).\end{aligned}$$

It follows that two ordered n-tuples $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$ are equal if and only if $a_1 = b_1, \dots, a_n = b_n$.

Separation Schema

Let $\varphi(u, p)$ be a formula. For any X and p, there exists a set $Y = \{u \in X : \varphi(u, p)\}$:

$$\forall X\, \forall p\, \exists Y\, \forall u\, (u \in Y \leftrightarrow u \in X \wedge \varphi(u, p)). \tag{1.2}$$

For each formula $\varphi(u, p)$, the formula (1.2) is an Axiom (of Separation). The set Y in (1.2) is unique by Extensionality.

Note that a more general version of Separation Axioms can be proved using ordered n-tuples: Let $\psi(u, p_1, \dots, p_n)$ be a formula. Then

$$\forall X\, \forall p_1 \dots \forall p_n\, \exists Y\, \forall u\, (u \in Y \leftrightarrow u \in X \wedge \psi(u, p_1, \dots, p_n)). \tag{1.3}$$

Simply let $\varphi(u, p)$ be the formula

$$\exists p_1, \dots \exists p_n \, (p = (p_1, \dots, p_n) \text{ and } \psi(u, p_1, \dots, p_n))$$

and then, given X and $p_1, \dots, p_n$, let

$$Y = \{u \in X : \varphi(u, (p_1, \dots, p_n))\}.$$

We can give the Separation Axioms the following form: Consider the class $C = \{u : \varphi(u, p_1, \dots, p_n)\}$; then by (1.3)

$$\forall X \, \exists Y \, (C \cap X = Y).$$

Thus the intersection of a class C with any set is a set; or, we can say even more informally

a subclass of a set is a set.

One consequence of the Separation Axioms is that the intersection and the difference of two sets is a set, and so we can define the operations

$$X \cap Y = \{u \in X : u \in Y\} \quad \text{and} \quad X - Y = \{u \in X : u \notin Y\}.$$

Similarly, it follows that the empty class

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

is a set—the *empty set*; this, of course, only under the assumption that at least one set X exists (because $\emptyset \subset X$):

$$\exists X \, (X = X). \tag{1.4}$$

We have not included (1.4) among the axioms, because it follows from the Axiom of Infinity.

Two sets X, Y are called *disjoint* if $X \cap Y = \emptyset$.

If C is a nonempty class of sets, we let

$$\bigcap C = \bigcap \{X : X \in C\} = \{u : u \in X \text{ for every } X \in C\}.$$

Note that $\bigcap C$ is a set (it is a subset of any $X \in C$). Also, $X \cap Y = \bigcap \{X, Y\}$.

Another consequence of the Separation Axioms is that the universal class V is a proper class; otherwise,

$$S = \{x \in V : x \notin x\}$$

would be a set.

Union

For any X there exists a set $Y = \bigcup X$:

$$\forall X\, \exists Y\, \forall u\, (u \in Y \leftrightarrow \exists z\, (z \in X \wedge u \in z)). \tag{1.5}$$

Let us introduce the abbreviations

$$(\exists z \in X)\, \varphi \quad \text{for} \quad \exists z\, (z \in X \wedge \varphi),$$

and

$$(\forall z \in X)\, \varphi \quad \text{for} \quad \forall z\, (z \in X \rightarrow \varphi).$$

By (1.5), for every X there is a unique set

$$Y = \{u : (\exists z \in X)\, u \in z\} = \bigcup\{z : z \in X\} = \bigcup X,$$

the *union* of X.

Now we can define

$$X \cup Y = \bigcup\{X, Y\}, \qquad X \cup Y \cup Z = (X \cup Y) \cup Z, \qquad \text{etc.,}$$

and also

$$\{a, b, c\} = \{a, b\} \cup \{c\},$$

and in general

$$\{a_1, \ldots, a_n\} = \{a_1\} \cup \ldots \cup \{a_n\}.$$

We also let

$$X \triangle Y = (X - Y) \cup (Y - X),$$

the *symmetric difference* of X and Y.

Power Set

For any X there exists a set $Y = P(X)$:

$$\forall X\, \exists Y\, \forall u\, (u \in Y \leftrightarrow u \subset X).$$

A set U is a *subset* of X, $U \subset X$, if

$$\forall z\, (z \in U \rightarrow z \in X).$$

If $U \subset X$ and $U \neq X$, then U is a *proper subset* of X.

The set of all subsets of X,

$$P(X) = \{u : u \subset X\},$$

is called the *power set* of X.

Using the Power Set Axiom we can define other basic notions of set theory.

The *product* of X and Y is the set of all pairs (x, y) such that $x \in X$ and $y \in Y$:

$$X \times Y = \{(x, y) : x \in X \text{ and } y \in Y\}. \tag{1.6}$$

The notation $\{(x, y) : \dots\}$ in (1.6) is justified because

$$\{(x, y) : \varphi(x, y)\} = \{u : \exists x \, \exists y \, (u = (x, y) \text{ and } \varphi(x, y))\}.$$

The product $X \times Y$ is a set because

$$X \times Y \subset PP(X \cup Y).$$

Further, we define

$$X \times Y \times Z = (X \times Y) \times Z,$$

and in general

$$X_1 \times \dots \times X_{n+1} = (X_1 \times \dots \times X_n) \times X_{n+1}.$$

Thus

$$X_1 \times \dots \times X_n = \{(x_1, \dots, x_n) : x_1 \in X_1 \wedge \dots \wedge x_n \in X_n\}.$$

We also let

$$X^n = \underbrace{X \times \dots \times X}_{n \text{ times}}.$$

An *n-ary relation* R is a set of n-tuples. R is a relation *on* X if $R \subset X^n$. It is customary to write $R(x_1, \dots, x_n)$ instead of

$$(x_1, \dots, x_n) \in R,$$

and in case that R is binary, then we also use

$$x \, R \, y$$

for $(x, y) \in R$.

If R is a binary relation, then the *domain* of R is the set

$$\operatorname{dom}(R) = \{u : \exists v \, (u, v) \in R\},$$

and the *range* of R is the set

$$\operatorname{ran}(R) = \{v : \exists u \, (u, v) \in R\}.$$

Note that $\operatorname{dom}(R)$ and $\operatorname{ran}(R)$ are sets because

$$\operatorname{dom}(R) \subset \bigcup \bigcup R, \qquad \operatorname{ran}(R) \subset \bigcup \bigcup R.$$

The *field* of a relation R is the set $\operatorname{field}(R) = \operatorname{dom}(R) \cup \operatorname{ran}(R)$.

In general, we call a class R an n-ary relation if all its elements are n-tuples; in other words, if

$$R \subset V^n = \text{the class of all } n\text{-tuples},$$

where C^n (and $C \times D$) is defined in the obvious way.

A binary relation f is a *function* if $(x, y) \in f$ and $(x, z) \in f$ implies $y = z$. The unique y such that $(x, y) \in f$ is the *value* of f at x; we use the standard notation

$$y = f(x)$$

or its variations $f : x \mapsto y$, $y = f_x$, etc. for $(x, y) \in f$.

f is a function *on* X if $X = \mathrm{dom}(f)$. If $\mathrm{dom}(f) = X^n$, then f is an *n-ary function* on X.

f is a function *from* X to Y,

$$f : X \to Y,$$

if $\mathrm{dom}(f) = X$ and $\mathrm{ran}(f) \subset Y$. The set of all functions from X to Y is denoted by Y^X. Note that Y^X is a set:

$$Y^X \subset P(X \times Y).$$

If $Y = \mathrm{ran}(f)$, then f is a function *onto* Y. A function f is *one-to-one* if

$$f(x) = f(y) \quad \text{implies} \quad x = y.$$

An *n-ary operation on* X is a function $f : X^n \to X$.

The *restriction* of a function f to a set X (usually a subset of $\mathrm{dom}(f)$) is the function

$$f \restriction X = \{(x, y) \in f : x \in X\}.$$

A function g is an *extension* of a function f if $g \supset f$, i.e., $\mathrm{dom}(f) \subset \mathrm{dom}(g)$ and $g(x) = f(x)$ for all $x \in \mathrm{dom}(f)$.

If f and g are functions such that $\mathrm{ran}(g) \subset \mathrm{dom}(f)$, then the *composition* of f and g is the function $f \circ g$ with domain $\mathrm{dom}(f \circ g) = \mathrm{dom}(g)$ such that $(f \circ g)(x) = f(g(x))$ for all $x \in \mathrm{dom}(g)$.

We denote the *image* of X by f either $f``X$ or $f(X)$:

$$f``X = f(X) = \{y : (\exists x \in X)\, y = f(x)\},$$

and the *inverse image* by

$$f_{-1}(X) = \{x : f(x) \in X\}.$$

If f is one-to-one, then f^{-1} denotes the *inverse* of f:

$$f^{-1}(x) = y \quad \text{if and only if} \quad x = f(y).$$

The previous definitions can also be applied to classes instead of sets. A class F is a function if it is a relation such that $(x, y) \in F$ and $(x, z) \in F$

implies $y = z$. For example, $F``C$ or $F(C)$ denotes the image of the class C by the function F.

It should be noted that a function is often called a *mapping* or a *correspondence* (and similarly, a set is called a *family* or a *collection*).

An *equivalence relation* on a set X is a binary relation $\equiv$ which is *reflexive*, *symmetric*, and *transitive*: For all $x, y, z \in X$,

$$x \equiv x,$$
$$x \equiv y \text{ implies } y \equiv x,$$
$$\text{if } x \equiv y \text{ and } y \equiv z \text{ then } x \equiv z.$$

A family of sets is *disjoint* if any two of its members are disjoint. A *partition* of a set X is a disjoint family P of nonempty sets such that

$$X = \bigcup\{Y : Y \in P\}.$$

Let $\equiv$ be an equivalence relation on X. For every $x \in X$, let

$$[x] = \{y \in X : y \equiv x\}$$

(the *equivalence class* of x). The set

$$X/{\equiv} = \{[x] : x \in X\}$$

is a partition of X (the *quotient* of X by $\equiv$). Conversely, each partition P of X defines an equivalence relation on X:

$$x \equiv y \quad \text{if and only if} \quad (\exists Y \in P)(x \in Y \text{ and } y \in Y).$$

If an equivalence relation is a class, then its equivalence classes may be proper classes. In Chapter 6 we shall introduce a trick that enables us to handle equivalence classes as if they were sets.

Infinity

There exists an infinite set.

To give a precise formulation of the Axiom of Infinity, we have to define first the notion of finiteness. The most obvious definition of finiteness uses the notion of a natural number, which is as yet undefined. We shall define natural numbers (as finite ordinals) in Chapter 2 and give only a quick treatment of natural numbers and finiteness in the exercises below.

In principle, it is possible to give a definition of finiteness that does not mention numbers, but such definitions necessarily look artificial.

We therefore formulate the Axiom of Infinity differently:

$$\exists S\,(\emptyset \in S \wedge (\forall x \in S) x \cup \{x\} \in S).$$

We call a set S with the above property *inductive*. Thus we have:

Axiom of Infinity. *There exists an inductive set.*

The axiom provides for the existence of infinite sets. In Chapter 2 we show that an inductive set is infinite (and that an inductive set exists if there exists an infinite set).

We shall introduce natural numbers and finite sets in Chapter 2, as a part of the introduction of ordinal numbers. In Exercises 1.3–1.9 we show an alternative approach.

Replacement Schema

If a class F is a function, then for every set X, $F(X)$ is a set.

For each formula $\varphi(x, y, p)$, the formula (1.7) is an Axiom (of Replacement):

$$\text{(1.7)}\quad \forall x\, \forall y\, \forall z\, (\varphi(x, y, p) \wedge \varphi(x, z, p) \rightarrow y = z) \\ \rightarrow \forall X\, \exists Y\, \forall y\, (y \in Y \leftrightarrow (\exists x \in X)\, \varphi(x, y, p)).$$

As in the case of Separation Axioms, we can prove the version of Replacement Axioms with several parameters: Replace p by $p_1, \ldots, p_n$.

If $F = \{(x, y) : \varphi(x, y, p)\}$, then the premise of (1.7) says that F is a function, and we get the formulation above. We can also formulate the axioms in the following ways:

If a class F is a function and $\mathrm{dom}(F)$ *is a set, then* $\mathrm{ran}(F)$ *is a set.*
If a class F is a function, then $\forall X\, \exists f\, (F{\restriction}X = f)$.

The remaining two axioms, Choice and Regularity, will by introduced in Chapters 5 and 6.

Exercises

1.1. Verify (1.1).

1.2. There is no set X such that $P(X) \subset X$.

Let
$$\boldsymbol{N} = \bigcap\{X : X \text{ is inductive}\}.$$
$\boldsymbol{N}$ is the smallest inductive set. Let us use the following notation:
$$0 = \emptyset, \qquad 1 = \{0\}, \qquad 2 = \{0, 1\}, \qquad 3 = \{0, 1, 2\}, \qquad \ldots.$$
If $n \in \boldsymbol{N}$, let $n + 1 = n \cup \{n\}$. Let us define $<$ (on $\boldsymbol{N}$) by $n < m$ if and only if $n \in m$.

A set T is *transitive* if $x \in T$ implies $x \subset T$.

1.3. If X is inductive, then the set $\{x \in X : x \subset X\}$ is inductive. Hence $\boldsymbol{N}$ is transitive, and for each n, $n = \{m \in \boldsymbol{N} : m < n\}$.

1.4. If X is inductive, then the set $\{x \in X : x \text{ is transitive}\}$ is inductive. Hence every $n \in \boldsymbol{N}$ is transitive.

1.5. If X is inductive, then the set $\{x \in X : x \text{ is transitive and } x \notin x\}$ is inductive. Hence $n \notin n$ and $n \neq n+1$ for each $n \in \boldsymbol{N}$.

1.6. If X is inductive, then $\{x \in X : x$ is transitive and every nonempty $z \subset x$ has an $\in$-minimal element$\}$ is inductive (t is $\in$-*minimal* in z if there is no $s \in z$ such that $s \in t$).

1.7. Every nonempty $X \subset \boldsymbol{N}$ has an $\in$-minimal element.
[Pick $n \in X$ and look at $X \cap n$.]

1.8. If X is inductive then so is $\{x \in X : x = \emptyset \text{ or } x = y \cup \{y\} \text{ for some } y\}$. Hence each $n \neq 0$ is $m+1$ for some m.

1.9 (Induction). Let A be a subset of $\boldsymbol{N}$ such that $0 \in A$, and if $n \in A$ then $n+1 \in A$. Then $A = \boldsymbol{N}$.

A set X *has n elements* (where $n \in \boldsymbol{N}$) if there is a one-to-one mapping of n onto X. A set is *finite* if it has n elements for some $n \in \boldsymbol{N}$, and *infinite* if it is not finite.

A set S is *T-finite* if every nonempty $X \subset P(S)$ has a $\subset$-maximal element, i.e., $u \in X$ such that there is no $v \in X$ with $u \subset v$ and $u \neq v$. S is T-*infinite* if it is not T-finite. (T is for Tarski.)

1.10. Each $n \in \boldsymbol{N}$ is T-finite.

1.11. $\boldsymbol{N}$ is T-infinite; the set $\boldsymbol{N} \subset P(\boldsymbol{N})$ has no $\subset$-maximal element.

1.12. Every finite set is T-finite.

1.13. Every infinite set is T-infinite.
[If S is infinite, consider $X = \{u \subset S : u \text{ is finite}\}$.]

1.14. The Separation Axioms follow from the Replacement Schema.
[Given φ, let $F = \{(x, x) : \varphi(x)\}$. Then $\{x \in X : \varphi(x)\} = F(X)$, for every X.]

1.15. Instead of Union, Power Set, and Replacement Axioms consider the following weaker versions:

(1.8) $\forall X\, \exists Y\, \bigcup X \subset Y$, i.e., $\forall X\, \exists Y\, (\forall x \in X)(\forall u \in x)\, u \in Y$,

(1.9) $\forall X\, \exists Y\, P(X) \subset Y$, i.e., $\forall X\, \exists Y\, \forall u\, (u \subset X \rightarrow u \in Y)$,

(1.10) If a class F is a function, then $\forall X\, \exists Y\, F(X) \subset Y$.

Then axioms 1.4, 1.5, and 1.7 can be proved from (1.8), (1.9), and (1.10), using the Separation Schema (1.3).

Historical Notes

Set theory was invented by Georg Cantor. The first attempt to consider infinite sets is attributed to Bolzano (who introduced the term *Menge*). It was however Cantor who realized the significance of one-to-one functions between sets and introduced the notion of cardinality of a set. Cantor originated the theory of cardinal and ordinal numbers as well as the investigations of the topology of the real line. Much of the development in the first four chapters follows Cantor's work. The main reference to Cantor's work is his collected works, Cantor [1932]. Another source of references to the early research in set theory is Hausdorff's book [1914].

Cantor started his investigations in [1874], where he proved that the set of all real numbers is uncountable, while the set of all algebraic reals is countable. In [1878] he gave the first formulation of the celebrated Continuum Hypothesis.

The axioms for set theory (except Replacement and Regularity) are due to Zermelo [1908]. The Replacement Schema is due to Fraenkel [1922a] and Skolem (see [1970], pp. 137–152).

Exercises 1.12 and 1.13: Tarski [1925a].