SET THEORY

- Texts:
  - **Introduction to Set Theory**, Karel Hrbacek and Thomas Jech, 3rd Edition, Marcel Dekker.
  - **Set Theory**, Charles C. Pinter, reprinted in Korea by KyungMoon.
- References:
  - **Naive Set Theory**, Paul R. Halmos, UTM, Springer.
  - **Elements of Set Theory**, Herbert B. Enderton, Academic Press.
  - **The Joy of Sets**, Keith Devlin, UTM, Springer.
  - **Set Theory**, You-Feng Lin and Shwu-Yeng Lin, reprinted in Korea by Kyung-Moon.

0.1 A Brief History of Mathematical Logic

- Cantor's Set Theory

- Russell's Paradox

- Hilbert's Formalism and Gödel's Work

- ZFC (Zermelo-Fraenkel + Choice) Axioms for Set Theory

Big sets like $\{x|x = x\}$, $\{x|\ x \notin x\}$ are called *(proper) classes*. The assumption of the existence of those proper classes are the main causes of paradoxes such as Russell's. Hence in formal set theory (see Appendix where we summarize the essence of axiomatic method), those classes do not exist. **But this course mainly focuses on elementary treatments of set theory rather than full axiomatic methods, although as the lectures proceed some ideas of the axiomatic methods will occasionally be examined.**[1]

[1] In the chapters of this note, those reviews will be stated after **ZFC:** mark.

0.2 A Review of Mathematical Logic

When we prove a theorem, we use common mathematical reasonings. Indeed mathematical statements or reasonings are often expressed as a sequence of formal symbols for briefness and notational simplicity. Here we point out some of typical logic we use.

Let p, q, r be mathematical statements. By combining symbols $\neg$(not), $\wedge$(and), $\vee$(or), $\rightarrow$, $\leftrightarrow$ to p, q, r properly, we can make further statements (called Boolean combinations of p, q, r) such as

$\neg p$ (not p);

$p \rightarrow q$ (p implies q, equivalently say, if p then q; p only if q; not q implies not p);

$p \vee \neg p$ (p or not p);

$\neg(p \wedge q) \leftrightarrow \neg p \vee \neg q$ (not (p and q) iff(=if and only if) (not p) or (not q));

and so on. As we know some of such Boolean combinations are *tautologies*, i.e. the values of their truth tables are all true. For example, $p \vee \neg p$, $p \rightarrow p \vee q$, $p \wedge r \rightarrow r$, $(p \rightarrow q) \leftrightarrow (\neg q \rightarrow \neg p)$, $\neg(p \wedge q) \leftrightarrow \neg p \vee \neg q$, $\neg(p \vee q) \leftrightarrow \neg p \wedge \neg q$ are all tautologies. We indeed freely use all those tautologies when we prove theorems. Let us see the following.

Theorem 0.1. *$x \in A \rightarrow x \in B$, (i.e. $A \subseteq B$) iff $x \in A \leftrightarrow (x \in A \wedge x \in B)$ (i.e. $A = A \cap B$).*

Proof. ($\Rightarrow$) Assume left. To prove right, first assume $x \in A$. We want to prove $x \in A$ and $x \in B$. By the first assumption, $x \in A$. By the assumption of left, we have $x \in B$ too.

Now assume $x \in A \wedge x \in B$. Then we have $x \in A$. Hence right is proved.

($\Leftarrow$) Assume right. Hence if $x \in A$, then $x \in A$ and B holds. In particular, $x \in B$. $\square$

Now let us go over more complicated logic involving quantifiers $\forall$ (for all), and $\exists$ (there exists; for some). Let $P(x), Q(x, y)$ be certain properties on x and x, y respectively. Note that the followings are *logically valid* (i.e. always true no matter what the properties P, Q exactly be, hence of course all tautologies are logically valid):

$\neg\forall x P x (= \neg(\forall x P x)) \leftrightarrow \exists x \neg P x (= \exists x(\neg P x))$ (not everything satisfies P iff there is something not satisfying P);

$\forall x \neg P x (= \forall x(\neg P x)) \leftrightarrow \neg\exists x P x (= \neg(\exists x P x))$ (everything does not satisfy P iff there does not exist one satisfying P);

$\exists x \forall y Q x y \leftrightarrow \neg\forall x \exists y \neg Q x y$ (for some (fixed) x, Qxy holds for every y iff it is not the case that for each x there corresponds y such that Qxy fails to hold). For example, if Qxy means x cuts y's hair, then saying there is the one who cuts everyone's hair is the same amount of saying that it is not the case that every person can find someone whose hair the person does not cut;

$\exists x \forall y Q x y \rightarrow \forall y \exists x Q x y$ (if there is x such that for every y, Qxy holds then for any y we can find x such that Qxy holds).

But the converse $\forall y \exists x Q x y \rightarrow \exists x \forall y Q x y$ is *not* logically valid, since for example if Qxy means x is a biological father of y, then even if everyone has a father, there is no one who is a biological father of everybody;

$(Px \wedge \exists y Q x y) \leftrightarrow \exists y(Px \wedge Q x y)$ is logically valid.

Here the same holds if $\wedge$ or $\exists$ (or both) is (are) replaced by $\vee$ or $\forall$ (or both), respectively;

$(\forall x P x \wedge \forall x Q x) \leftrightarrow \forall x(Px \wedge Qx)$ is logically valid.

Does the same hold if $\wedge$ or $\exists$ (or both) is (are) replaced by $\vee$ or $\forall$ (or both), respectively?

There are many other logically valid sentences. Logically valid sentences are also freely used in proving theorems.

Definition 0.2. (1) By an *indexed family* $\{A_i|\ i \in I\}$ *of sets*, we mean a collection of sets A_i indexed by $i \in I$.
(2) $\bigcup\{A_i|\ i \in I\} := \{x|\ \exists i \in I\ \ni x \in A_i\}$.[2]
(3) $\bigcap\{A_i|\ i \in I\} := \{x|\ \forall i \in I,\ x \in A_i\}$.[3]

Theorem 0.3. *Let* $\{A_i|\ i \in I\}$ *be an indexed family of sets.*
(1) *If* $A_i \subseteq B$ *for all* $i \in I$, *then* $\bigcup\{A_i|\ i \in I\} \subseteq B$.
(2) $(\bigcup\{A_i|\ i \in I\})^c = \bigcap\{A_i^c|\ i \in I\}$.

Proof. (1) Suppose that $A_i \subseteq B$ for all $i \in I$. Now let $x \in \bigcup\{A_i|\ i \in I\}$. We want to show $x \in B$. By the definition, $x \in A_{i_0}$ for some $i_0 \in I$. Therefore by supposition, as desired $x \in B$.

(2) $x \in (\bigcup\{A_i|\ i \in I\})^c$ iff $x \notin \bigcup\{A_i|\ i \in I\}$ iff $\forall i \in I, x \notin A_i$ iff $\forall i \in I, x \in A_i^c$ iff $x \in \bigcap\{A_i^c|\ i \in I\}$. □

Throughout the course students will be trained to be capable to do these kinds of logical reasonings fairly freely and comfortably.

[2]More precisely $\{x|\ \exists i(i \in I \wedge x \in A_i)\}$.
[3]More precisely $\{x|\ \forall i(i \in I \to x \in A_i)\}$.

1. **Sets**

Definition 1.1. $A = B$ if $\forall x(x \in A \leftrightarrow x \in B)$.

$A \subseteq B$ (A is a *subset* of B) if $\forall x(x \in A \to x \in B)$. We say A is a *proper subset* of B if $A \subseteq B$ and $A \neq B$ ($A \subsetneq B$).

$\emptyset$, called the empty set, is the set having no elements.

We can write a set in the form of $\{x|\ P(x)\}$.

E.g. $\emptyset = \{x|\ x \neq x\}$, $\{x, y\} = \{z|\ z = x \text{ or } z = y\}$.

$A \cup B := \{x|\ x \in A \text{ or } x \in B\}$. $A \cap B := \{x|\ x \in A \text{ and } x \in B\}$. $\mathcal{P}(A) := \{B|\ B \subseteq A\}$.

$A - B = \{x|\ x \in A \wedge x \notin B\}$. $A \triangle B := (A - B) \cup (B - A)$. $A^c := \{x \in U|\ x \notin A\}$.

We say sets A, B are *disjoint* if $A \cap B = \emptyset$.

For a set $\mathcal{S}$,

$\bigcup \mathcal{S} = \bigcup\{X|\ X \in \mathcal{S}\} = \bigcup_{X \in \mathcal{S}} X := \{x|\ x \in A \text{ for some } A \in \mathcal{S}\} = \{x|\ \exists A \in \mathcal{S} \text{ s.t. } x \in A\}$.

For a nonempty set $\mathcal{S}$,

$\bigcap \mathcal{S} = \bigcap\{X|\ X \in \mathcal{S}\} = \bigcap_{X \in \mathcal{S}} X := \{x|\ x \in A \text{ for all } A \in \mathcal{S}\} = \{x|\ \forall A \in \mathcal{S},\ x \in A\}$.

• Laws of Set Operations

$A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \triangle A = \emptyset$,

$A \subseteq B \Rightarrow \bigcup A \subseteq \bigcup B$ and $\bigcap B \subseteq \bigcap A$ (when $A \neq \emptyset$).

Commutativity $A \cup B = B \cup A$, $A \cap B = B \cap A$, $A \triangle B = B \triangle A$.

Associativity $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$, $A \triangle (B \triangle C) = (A \triangle B) \triangle C$.

Distributivity $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

$A \cup \bigcap \mathcal{S} = \bigcap\{A \cup X|\ X \in \mathcal{S}\}$, $A \cap \bigcup \mathcal{S} = \bigcup\{A \cap X|\ X \in \mathcal{S}\}$.

DeMorgan's laws $U - (A \cup B) = (U - A) \cap (U - B)$ $((A \cup B)^c = A^c \cap B^c)$,

$U - (A \cap B) = (U - A) \cup (U - B)$ $((A \cap B)^c = A^c \cup B^c)$.

$U - (\bigcup \mathcal{S}) = \bigcap\{U - X|\ X \in \mathcal{S}\}$ $((\bigcup \mathcal{S})^c = \bigcap\{X^c|\ X \in \mathcal{S}\})$,

$U - (\bigcap \mathcal{S}) = \bigcup\{U - X|\ X \in \mathcal{S}\}$ $((\bigcap \mathcal{S})^c = \bigcup\{X^c|\ X \in \mathcal{S}\})$.

2. **Relations, Functions, and Orderings**

•Relations

Definition 2.1. $(a, b) := \{\{a\}, \{a, b\}\}$, the ordered pair of a, b.

By iteration, we can define ordered tuples: $(a, b, c) := ((a, b), c)$; $(a, b, c, d) := ((a, b, c), d)$; ...

$A \times B := \{(a, b)|\ a \in A,\ b \in B\}$.

$A \times B \times C := (A \times B) \times C = \{(a, b, c)|\ a \in A,\ b \in B,\ c \in C\}$.

...

$A^2 := A \times A,\ \ A^3 := A \times A \times A$, ...

Theorem 2.2. $(a, b) = (a', b')$ *iff* $a = a'$ *and* $b = b'$.

Definition 2.3. A *(binary) relation* R is some set of ordered pairs. We may write xRy for $(x, y) \in R$.

We say R is a *unary relation in* A, if $R \subseteq A$.

R is a *binary relation in* A, if $R \subseteq A^2$.

R is a *ternary relation in* A, if $R \subseteq A^3$.

...

R is an *n-ary relation in* A, if $R \subseteq A^n$.

Definition 2.4. Let R be a relation.

(1) $\operatorname{dom} R := \{x|\ \exists y (x, y) \in R\}$.
(2) $\operatorname{ran} R := \{y|\ \exists x\ \ni xRy\}$.
(3) $\operatorname{field}\ R := \operatorname{dom} R \cup \operatorname{ran} R$.
(4) $R^{-1} := \{(x, y)|\ (y, x) \in R\}$.
(5) $R[A] := \{y \in \operatorname{ran} R|\ \exists x \in A\ \ni xRy\}$, the *image* of A under R. It can be seen that $R^{-1}[B] = \{x \in \operatorname{dom} R|\ \exists y \in B$ s.t. $xRy\}$, the *inverse image* of B under R.
(6) $R\lceil A := \{(x, y) \in R|\ x \in A\}$, the *restriction* of R to A.
(7) Let S be a relation. Then $S \circ R := \{(x, z)|\ \exists y\ (xRy \wedge ySz)\}$.

Exercise 2.5. *Let* R, S, T *be relations.*

$\operatorname{dom} R = \operatorname{ran} R^{-1}$, $\operatorname{ran} R = \operatorname{dom} R^{-1}$, $(R^{-1})^{-1} = R$,

$R[A \cup B] = R[A] \cup R[B]$, $R[A \cap B] \subseteq R[A] \cap R[B]$.

$R \circ (S \circ T) = (R \circ S) \circ T$, $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$.

– – – – – – – – – – – –

•Functions

Definition 2.6. A relation F is said to be a *function* (or *mapping)* if for each $a \in \operatorname{dom} F$, there is a unique b such that aFb holds.

If F is a function, then $F(a) :=$ the unique b such that of $(a, b) \in F$, the *value* of F at a.

We write $f : A \to B$ if f is a function, $A = \operatorname{dom} f$, and $\operatorname{ran} f \subseteq B$, and say f is a function from (or on) A (in)to B.

Lemma 2.7. *Let F, G be functions. Then $F = G$ iff $\operatorname{dom} F = \operatorname{dom} G$ and $\forall x \in \operatorname{dom} F$, $F(x) = G(x)$.*

Definition 2.8. $F : A \to B$ is given.

(1) We say F is 1-1 (or *injective*, an *injection*) if for $a \neq b \in A$, we have $F(a) \neq F(b)$.
(2) F is *onto* (or *surjective*, a *surjection*) if $B = \operatorname{ran} F$.
(3) F is *bijective* (or, a *bijection*) if F is 1-1 and onto.

Theorem 2.9. *Let f, g be functions.*

(1) *$g \circ f$ is a function.*
(2) $\operatorname{dom}(g \circ f) = \{x \in \operatorname{dom} f |\ f(x) \in \operatorname{dom} g\}$.
(3) $(g \circ f)(x) = g(f(x))$ *for all* $x \in \operatorname{dom}(g \circ f)$.

Theorem 2.10. *A function f is invertible (i.e. the relation f^{-1} is a function too) iff f is injective.*

$f : A \to B$ is invertible and $f^{-1} : B \to A$ iff $f : A \to B$ is bijective.

Notation 2.11. ${}^A B := \{f |\ f : A \to B\}$.

Note that for all A, ${}^{\emptyset}A = \{\emptyset\}$ and if $A \neq \emptyset$, then ${}^A\emptyset = \emptyset$.

Now assume $f : I \to \mathcal{S} = \{X |\ X \in \mathcal{S}\}$ is onto. Then $\mathcal{S} = \operatorname{ran} f = \{f(i) |\ i \in I\}$. We may write $f(i) = X_i$, and write $\mathcal{S} = \{X_i |\ i \in I\} = \{X_i\}_{i \in I}$. We call $\mathcal{S}$, an *indexed family* of sets.

For $\mathcal{S} = \{X_i |\ i \in I\}$,

$$\bigcup_{i \in I} X_i \text{ denotes } \bigcup \mathcal{S}, \quad \bigcap_{i \in I} X_i \text{ (with nonempty } I\text{) denotes } \bigcap \mathcal{S}, \text{ and}$$

$$\prod_{i \in I} X_i := \{f |\ f : I \to \bigcup \mathcal{S} \text{ s.t. } \forall i \in I,\ f(i) \in X_i\}.$$

Exercise 2.12. *f is a function.*

(1) *Let $f : A \to B$. If f is bijective, then $f^{-1} \circ f = \mathrm{Id}_A$, and $f \circ f^{-1} = \mathrm{Id}_B$. Conversely if $g \circ f = \mathrm{Id}_A$ and $f \circ g = \mathrm{Id}_B$ for some $g : B \to A$, then f is bijective and $g = f^{-1}$.*
(2) *Let $f : A \to B$, $h : B \to C$ be given. If both f, h are 1-1 (onto resp.), then so is $h \circ f$.*
(3) $f[\bigcup_{i \in I} A_i] = \bigcup_{i \in I} f[A_i]$.
$f^{-1}[\bigcup_{i \in I} A_i] = \bigcup_{i \in I} f^{-1}[A_i]$.
$f[\bigcap_{i \in I} A_i] \subseteq \bigcap_{i \in I} f[A_i]$.
$f^{-1}[\bigcap_{i \in I} A_i] = \bigcap_{i \in I} f^{-1}[A_i]$.

– – – – – – – – – – – –

•Equivalence Relations

Definition 2.13. Let R be a relation in A (i.e. $R \subseteq A^2$).

(1) R is *reflexive* if for all $a \in A$, aRa.
(2) R is *symmetric* if $\forall a, b \in A$, $aRb \to bRa$.
(3) R is *transitive* if $\forall a, b, c \in A$, $aRb \wedge bRc \to aRc$.

R is said to be an *equivalence relation on* A, if R is reflexive, symmetric, and transitive.

Definition 2.14. Let E be an equivalence relation on A. For $a, b \in A$ with aEb, we say a and b are *equivalent* (or, a is *equivalent* to b) modulo E.

$[a]_E = a/E := \{x|\ xEa\}$ is called the *equivalence class* of a modulo E (or, the E-equivalence class of a). $A/E := \{[a]_E|\ a \in A\}$.

Lemma 2.15. *E is an equivalence relation on A. Then for $x, y \in A$, xEy iff $[x]_E = [y]_E$.*

Definition 2.16. By a *partition* of a set A, we mean a family $\mathcal{S}$ of nonempty subsets of A such that

(a) for $C \neq D \in \mathcal{S}$, C and D are disjoint,
(b) $A = \bigcup \mathcal{S}$.

Theorem 2.17. *Let F be an equivalence relation on A, and let $\mathcal{S}$ be a partition of A.*

(1) *$\{[a]_F|\ a \in A\}$ partitions A.*
(2) *If we define*

$$E_{\mathcal{S}} := \{(x, y) \in A^2|\ x, y \in C \text{ for some } C \in \mathcal{S}\},$$

then it is an equivalence relation on A such that $\mathcal{S} = A/E_{\mathcal{S}}$. Moreover, if $\mathcal{S} = A/F$, then $E_{\mathcal{S}} = F$.

Definition 2.18. (1) Let E, F be equivalence relations on A. We say E *refines* (*is an refinement of*, or *is finer than*) F (equivalently say F is *coarser than* E), if $E \subseteq F$.

(2) Let both $E \subseteq F$ be equivalence relations on A. The quotient of F by E (written as F/E) is

$$\{([x]_E, [y]_E)|\ xFy\}.$$

Theorem 2.19. *Let $E \subseteq F$ be equivalence relations on A. Then F/E is an equivalence relation on A/E.*

Theorem 2.20. *Let $f : A \to B$. Then f induces an equivalence relation $\sim$ on A such that $a \sim b$ iff $f(a) = f(b)$. Define $\varphi : A \to A/\sim$ by mapping $x \in A$ to $[x]_\sim$. Then there is a unique $\hat{f} : A/\sim \to B$ such that $f = \hat{f} \circ \varphi$. Moreover $\hat{f}$ is 1-1, and if f surjective, so is $\hat{f}$.*

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle\varphi}\big\downarrow & \nearrow \hat{f} & \\ A/\sim & & \end{array}$$

Corollary 2.21. *Let $E \subseteq F$ be equivalence relations on A. Then the canonical map $f : A/E \to A/F$ sending $[x]_E$ to $[x]_F$ is well-defined and onto. Moreover F/E is an equivalence relation on A/E induced by f. Hence $\hat{f} : (A/E)/(F/E) \to A/F$ is a bijection.*

$$\begin{array}{ccc} A/E & \xrightarrow{f} & A/F \\ {\scriptstyle\varphi}\big\downarrow & \nearrow \hat{f} & \\ (A/E)/(F/E) & & \end{array}$$

•Orderings

Definition 2.22. Let R be a relation in A.

(1) R is said to be *antisymmetric* if $\forall a, b \in A,\ aRb \wedge bRa \to a = b$.
(2) R is *asymmetric* if $\forall a, b \in A,\ aRb \to \neg bRa$.
(3) We say R is a (*partial*) *ordering* (or an *order relation*) of A if R is reflexive, antisymmetric, and transitive. The pair (A, R) is called an *ordered* set (or, a poset). If R is asymmetric and transitive, then we call R a *strict ordering* of A.
(4) Let $(A, \leq)$ be a poset. Clearly $\leq$ partially orders any subset of A (i.e. for $B(\subseteq A)$, $(B, \leq_B)$ is a poset where $\leq_B := \{(a, b) \in \leq \mid a, b \in B\}$.) We say $a, b \in A$ are *comparable* if $a \leq b$ or $b \leq a$. Otherwise a, b are *incomparable*. $C(\subseteq A)$ is called a *chain* in A if any two elements of C are comparable. We say $\leq$ is a *linear* (or *total*) ordering of A if any two elements of A are comparable, i.e. A itself is a chain in A.

As the reader may notice, an order relation is an abstraction of the notion 'less than or equal to', while a strict relation is an abstraction of the notion 'strictly less than'.

Theorem 2.23. *Let both $\leq, \prec$ be relations in A.*

(1) *$\prec$ is a strict ordering iff $\prec$ is irreflexive (i.e. $\forall x \in A,\ x \not\prec x$) and transitive.*
(2) (a) *If $\leq$ is an ordering, then the relation $<$ defined by $x \leq y \wedge x \neq y$ is a strict ordering.*
    (b) *Similarly, if $\prec$ is a strict ordering, then the relation $\preceq$ defined by $x \prec y \vee x = y$ is an ordering.*
(3) *$(A, \leq)$ is linearly ordered iff $\leq$ is reflexive, transitive, and $\forall x \neq y \in A$, exactly one of $x \leq y$ or $y \leq x$ holds.*

Definition 2.24. Let A be ordered by $\leq$, and $B \subseteq A$.

(1) $b \in B$ is the *least* element of B if $b \leq x$ for all $x \in B$. b is a *minimal* element of B if $\nexists x \in B$ s.t. $x < b$.
(2) $b \in B$ is the *greatest* element of B if $x \leq b$ for all $x \in B$. b is a *maximal* element of B if $\nexists x \in B$ s.t. $b < x$.
(3) $a \in A$ is a *lower bound* of B if $a \leq x$ for all $x \in B$. We say B is *bounded below* (in A) if there is a lower bound of B.
    $a \in A$ is an *upper bound* of B if $x \leq a$ for all $x \in B$. We say B is *bounded above* (in A) if there is an upper bound of B.
(4) $a \in A$ is the *infimum* (or *g.l.b.*) of B (in A) if a is the greatest element of the set of all lower bounds of B. We write $\inf_A B = \mathrm{glb}_A B = a$.
    $a \in A$ is the *supremum* of B in A if a is the least element of the set of all upper bounds of B. We write $\sup_A B = \mathrm{lub}_A B = a$.
(5) Assume $\leq$ linearly orders A. For $a, b \in A$, we say b is a *successor* of a (a is a *predecessor* of b) if $a < b$ and there is no c such that $a < c < b$.

Exercise 2.25. *Let $(A, \leq)$ be an ordered set, and $B \subseteq A$. Then $b \in A$ is the least (greatest, resp.) element of B iff $b = \inf B$ ($= \sup B$, resp.) and $b \in B$.*

Definition 2.26. Let $(A, \leq_A), (B, \leq_B)$ be posets. We say A and B are *order isomorphic* (write $A \cong B$) if there is a bijective $f : A \to B$, called an *order isomorphism*, such that for all $a, b \in A$, $a <_A b$ iff $f(a) <_B f(b)$.

Definition 2.27. An ordering of a set A is called a *well-ordering* if for any nonempty subset B of A has a least element. Every well-ordering is a linear ordering. If $\leq$ well-orders a set A, then clearly it well-orders any subset of A, too.

Let $(W, \leq)$ be well-ordered. We say $S(\subseteq W)$ is an *initial segment* (of $a \in W$) if $S = \{x \in W |\ x < a\}$. We write $S = W[a]$ or $= \text{seg } a$.

Lemma 2.28. *Any well-ordered set A is* not *order-isomorphic with a subset of an initial segment of A.*

Theorem 2.29. (Well-Ordering Isomorphism Theorem) *Let $(A, \leq_A), (B, \leq_B)$ be well-ordered. Then exactly one of the following holds.*

(1) $A \cong B$.
(2) $A \cong B[b]$ *for unique* $b \in B$.
(3) $B \cong A[a]$ *for unique* $a \in A$.

In each case, the isomorphism is unique.

Corollary 2.30. *Let $(A, \leq_A)$ be well-ordered. Then for $B \subseteq A$, it is order-isomorphic with A or with an initial segment of A.*

Definition 2.31. Let $(A, <_A), (B, <_B)$ be linear orderings.

Then $A \otimes_l B$, called the *lexicographic product* of A and B, is the set $A \times B$ ordered by $<_l$, called the *lexicographic* ordering of A and B, as follows:

$$(x_1, y_1) <_l (x_2, y_2) \text{ iff } x_1 <_A x_2 \text{ or } (x_1 = x_2 \text{ and } y_1 <_B y_2).$$

Similarly, $A \otimes_{al} B$, called the *antilexicographic product* of A and B, is again the set $A \times B$ ordered by $<_{al}$, called the *antilexicographic* ordering of A and B, as follows:

$$(x_1, y_1) <_{al} (x_2, y_2) \text{ iff } y_1 <_B y_2 \text{ or } (y_1 = y_2 \text{ and } x_1 <_A x_2).$$

If A, B are disjoint, then $A \oplus B$, called the *sum* of A and B, is the set $A \cup B$ ordered by $<$ as follows:

$$\begin{aligned} x < y \text{ iff } \quad & x, y \in A \text{ and } x <_A y \text{ or} \\ & x, y \in B \text{ and } x <_B y \text{ or} \\ & x \in A \text{ and } y \in B. \end{aligned}$$

Lemma 2.32. *Assume that $(A, <_A), (B, <_B)$ are linear orderings. Then so are $A \otimes_l B$, $A \otimes_{al} B$, and $A \oplus B$ (when $A \cap B = \emptyset$). Moreover $A \otimes_l B \cong B \otimes_{al} A$.*

Similarly, if $(A, <_A), (B, <_B)$ are well-orderings, then so are the products and the sum.

3. **Natural Numbers**

ZFC: The hidden intension of this chapter is to figure out how to formalize ω, the set of natural numbers, and its arithmetic system in ZFC. To show the existence of ω, we need Axiom of Infinity, and to define arithmetic, we need the Recursion Theorem which we shall prove later.

Definition 3.1. For a set a, its *successor* $S(a) = a^+ = a + 1 := a \cup \{a\}$. Note that $a \in a^+$ and $a \subseteq a^+$. $0 := \emptyset,\ 1 := 0^+,\ 2 := 1^+,\ 3 := 2^+, ...$

A set A is said to be *inductive*, if $\emptyset \in A$, and is closed under successor (i.e. $x \in A \to x^+ \in A$).

$\mathbb{N} = \omega := \{x|\ \forall I(I \text{ is an inductive set } \to x \in I)\}$. A *natural number* is an element in ω.

Theorem 3.2. *ω itself is inductive.*

Induction Principle Any inductive subset of ω is equal to ω. Namely the following holds.

Let $P(x)$ be a property.[4] Assume that

(a) $P(0)$ holds, and

(b) for all $n \in \omega$, $P(n)$ implies $P(n^+)$.

Then P holds for all $n \in \omega$.

– – – – – – – – – – – –

Definition 3.3. A set A is said to be *transitive* if every member of a member of A is a member of A, i.e. $x \in a \in A \to x \in A$.

Theorem 3.4. *Every natural number is transitive. ω is transitive as well.*

Definition 3.5. Define the relation $<$ on ω by $m < n$ iff $m \in n$.

Then $m \leq n$ iff ($m < n$ or $m = n$) iff $m \in n^+$ iff $m < n^+$.

Theorem 3.6. *$< (= \in)$ is a well-ordering of ω.*

Corollary 3.7. *Any natural number is well-ordered. Any subset of ω is order isomorphic with either a natural number or ω. Any $X(\subseteq n \in \omega)$ is isomorphic with some $m \leq n$.*

Corollary 3.8. (1) *$\leq$ is a linear ordering on ω.*

(2) *For any $m, n \in \omega$, exactly one of $m < n,\ m = n,\ n < m$ holds.*

(3) *For any $m, n \in \omega$, we have $m < n$ iff $m \in n$ iff $m \subsetneq n$, and $m \leq n$ iff $m \subseteq n$.*

[4] **ZFC:** Indeed a formula $\varphi(x)$.

Induction Principle (2nd Version) Let $P(x)$ be a property. Assume that for every $n \in \omega$,

$$\text{if } P(k) \text{ holds for all } k < n, \text{ then } P(n) \text{ holds.}$$

Then $P(n)$ holds for all $n \in \omega$.

– – – – – – – – – – –

The Recursion Theorem Let A be a set and let $c \in A$. Suppose that $f : A \to A$. Then there is a unique function $h : \omega \to A$ such that $h(0) = c$ and $h(k^+) = f(h(k))$ for all $k \in \omega$.[5]

Proof. Let $\mathcal{A} = \{R|\ R \subseteq \omega \times A\ \wedge (0, c) \in R \wedge\ \forall k \in \omega((k, x) \in R \to (k^+, f(x)) \in R)\}$. Since $\omega \times A \in \mathcal{A} \neq \emptyset$, $\bigcap \mathcal{A}$ exists. We let $h = \bigcap \mathcal{A}$.

Claim 1) h is the desired function:

Claim 2) If $h' : \omega \to A$ satisfies the two conditions then $h = h'$: □

• **Arithmetic of Natural Numbers** We will use the Recursion Theorem to define $+$ and $\times$ on ω.

Definition 3.9. (1) Fix $k \in \omega$. We define $+_k : \omega \to \omega$ as follows.
$+_k(0) = k$,
$+_k(n^+) = (+_k(n))^+$ for each $n \in \omega$.
By the Recursion Theorem, there is a unique such $+_k$. Now define
$+ : \omega^2 \to \omega$ by $k + m = +_k(m)$.

(2) Fix $k \in \omega$. We define $\times_k : \omega \to \omega$ as follows.
$\times_k(0) = 0$,
$\times_k(n^+) = (\times_k(n)) + k$ for each $n \in \omega$.
By the Recursion Theorem, there is a unique such $\times_k$. Now define
$\times : \omega^2 \to \omega$ by $k \times m = \times_k(m)$.

Theorem 3.10. $+, \times$ *satisfy the usual arithmetic laws such as associativity, commutativity, distributive law, cancellation laws, and 0 is an identity for* $+$*, and 1 is an identity for* $\times$*.*

From ω, we can consecutively build $\mathbb{Z}$ = the set of integers; $\mathbb{Q}$ = the set of rational numbers; $\mathbb{R}$ = the set of real numbers; $\mathbb{C}$ = the set of complex numbers. In particular, when we construct $\mathbb{R}$ from $\mathbb{Q}$, well-known notions such as *Dedekind cuts* or *Cauchy sequences* are used.

[5] **ZFC:** Formally,

$$\forall A \forall c \forall f[c \in A \wedge f \text{ is a function } \to \exists! h[h : \omega \to A \wedge h(0) = c \wedge \forall k(k \in \omega \to h(k^+) = f(h(k)))]]$$

is provable.

4. **Axiom of Choice**

ZFC: One of historically important issues is whether Axiom of Choice (AC) is provable from other axioms (ZF). Kurt Gödel and Paul Cohen proved that AC is independent from other axioms. It means even if we deny AC, we can still do consistent mathematics. Indeed there are a few of mathematicians who do not accept AC mainly because AC supplies non constructible existence proofs mostly in the form of Zorn's Lemma, and also Banach-Tarski phenomena. However the absolute majority of contemporary mathematicians accept AC since it is so natural and again a great deal of abstract existence theorems in basic algebra and analysis are relying on AC.

Now in this chapter, formally we only assume ZF. We then prove the equivalence of several statements to AC. From chapter 5, we freely assume AC (so formally then we work under ZFC).[6]

Definition 4.1. Axiom of Choice (Version 1): Let $\mathcal{A}$ be a family of mutually disjoint nonempty sets. Then there is a set C containing exactly one element from each set in $\mathcal{A}$.

Axiom of Choice (Version 2): Let $\{A_i\}_{i\in I}$ be an indexed family of nonempty sets. If $I \neq \emptyset$, then $\prod_{i\in I} A_i$ is nonempty.

Axiom of Choice (Version 3): For any nonempty set A, there is a function $F : \mathcal{P}(A) \to A$ such that $F(X) \in X$ for any nonempty subset X of A. (Such a function is called a *choice function* for $\mathcal{P}(A)$.)

Theorem 4.2. *AC 1st, 2nd, and 3rd versions are all equivalent.*

Theorem 4.3. *The following are equivalent.*
(1) **AC**.
(2) **(Well-Ordering Theorem)** *Every set can be well ordered.*
(3) **(Zorn's Lemma)** *Given a poset $(A, \leq)$, if every chain B of A has an upper bound, then A has a maximal element.*
$(3)'$ **(Zorn's Lemma)$'$** *Given a poset $(A, \leq)$, if every chain B of A has the supremum, then A has a maximal element.*
(4) **(Hausdorff's Maximal Principle)** *Every poset has a maximal chain.*

Proof. (1)⇒$(3)'$ Assume (1). Let $(A, \leq)$ be a poset such that

$$\text{each chain has the supremum.} \tag{*}$$

In particular there is the least element $p = \sup \emptyset$ in A.

To lead a contradiction, suppose that A has no maximal element. Then for each $x \in A$, $S(x) := \{y \in A |\ x < y\}$ is a nonempty subset of A. Let F be a choice function for $\mathcal{P}(A)$ so that $F(S(x)) \in S(x)$. Define $f : A \to A$ by $f(x) = F(S(x))$. Then since $f(x) \in S(x)$,

$$x < f(x), \text{ for each } x \in A. \tag{**}$$

Now let $\mathcal{H} = \{B \subseteq A |\quad$ (i) $p \in B$, (ii) $x \in B \to f(x) \in B$,
(iii) if $C(\subseteq B)$ is a chain then $\sup_A C \in B\}$.

[6]In some set theory books, authors are sensitive to discern between proofs using AC and not.

Due to (*), $A \in \mathcal{H} \neq \emptyset$. Hence we let $P := \bigcap \mathcal{H}$. It is easy to check that $P \in \mathcal{H}$ and $p \in P \subseteq A$. We shall show P is a chain. For this end we claim the following.

Claim 1) Let $P' := \{a \in P |\ \forall x \in P,\ x < a \to f(x) \leq a\}$. Let $a \in P'$. Then

$$B_a := \{x \in P |\ x \leq a \text{ or } f(a) \leq x\}$$

is equal to P. Indeed, a is comparable with any element in P, and P' is a chain: It is not hard to show that $B_a (\subseteq P)$ satisfies (i)(ii)(iii) (Exercise). Hence $P \subseteq B_a \in \mathcal{H}$, and $P = B_a$. Now given $b \in P = B_a$, either $b \leq a$ or $a < f(a) \leq b$ (by (**)), so a, b are comparable, and P' is a chain. We have proved Claim 1.

Claim 2) P' satisfies (i)(ii)(iii). Hence $P' = P$:

Firstly, that $p \in P'$ is vacuously true.

Secondly, suppose that $a \in P'$. We want to prove $f(a) \in P'$, i.e. given $x (\in P) < f(a)$, show $f(x) \leq f(a)$. Now due to Claim 1, we have $P = B_a$, and thus $x \leq a$. Then since $a \in P'$, either $f(x) \leq a (< f(a))$ or $(x = a$ and $f(x) = f(a))$. Hence always $f(x) \leq f(a)$.

Thirdly, suppose that $C (\subseteq P' \subseteq P)$ is a chain. Let $c = \sup_A C$. Then since $P (\in \mathcal{H})$ satisfies (iii), $c \in P$. It remains to show $c \in P'$. Let $x (\in P) < c$. We want to see $f(x) \leq c$. Now there is $y_0 \in C$ s.t. $x < y_0$, since otherwise $x \not< y$ for each $y \in C$, so due to Claim 1 (y is comparable with x), we have $y \leq x$ and $c = \sup_A C \leq x$, a contradiction. Then since $y_0 \in P'$, $f(x) \leq y_0 \leq c$. Therefore Claim 2 is proved.

Now by Claim 1, 2, we conclude that P is a chain. Then due to (*), there is $m := \sup_A P$. Since $P \in \mathcal{H}$, by (iii)(ii) we have $m \in P$ and $f(m) \in P$. But by (**), $m < f(m)$ and $m \neq \sup_A P$, a desired contradiction is obtained. Hence A must have a maximal element. We have proved (1)⇒(3)$'$.

(3)$'$ ⇒(4)

(4)⇒(3)

(3)⇒(2) Let A be a set. Let

$$\mathcal{A} = \{(B, \leq) |\ B \subseteq A,\ \leq \subseteq B^2,\ \text{and}\ \leq\ \text{well-orders}\ B\}.$$

Given $(B_1, \leq_1), (B_2, \leq_2) \in \mathcal{A}$, define

$$(B, \leq_1) \preccurlyeq (B_2, \leq_2) \text{ iff } \leq_1 \subseteq \leq_2 (\therefore B_1 \subseteq B_2), \text{ and } \forall x \in B_1, \forall y \in B_2 - B_1, x <_2 y.$$

Then $(\mathcal{A}, \preccurlyeq)$ forms an ordered set (Exercise).

Claim) Assume a chain $\mathcal{C} = \{(B_i, \leq_i)\}_{i \in I} \subseteq \mathcal{A}$ is given. Let $C := \bigcup \{B_i |\ i \in I\}$ and let $\leq_C := \bigcup \{\leq_i |\ i \in I\}$. Then $(C, \leq_C) \in \mathcal{A}$, and is an an upper bound of $\mathcal{C}$ (i.e. $(B_i, \leq_i) \preccurlyeq (C, \leq_C)$ for all $(B_i, \leq_i) \in \mathcal{C}$): Firstly, it is easy to see that $\leq_C$ is a relation in C (Exercise).

Secondly, we shall show that $\leq_C$ well-orders C. It can be seen that $\leq_C$ is reflexive and antisymmetric (Exercise). To show transitivity, assume $x \leq_C y$ and $y \leq_C z$. Then $x \leq_i y$ and $y \leq_j z$ for some $i, j \in I$. Since $\mathcal{C}$ is a chain we have say $\leq_i \subseteq \leq_j$. Then $x \leq_j y$ and by the transitivity of $\leq_j$, we have $x \leq_j z$ and $x \leq_C z$. It remains to show $\leq_C$ well-orders C. Let D be a nonempty subset of $C = \bigcup_{i \in I} B_i$. For some i_0, $D' := D \cap B_{i_0} \neq \emptyset$. Hence D' has the least element b in $(B_{i_0}, \leq_{i_0})$. We shall prove in fact b is the desired least element of D in $(C, \leq_C)$. Let $x \in D$. Hence $x \in B_j$ for some $j \in I$. Now if $x \in B_{i_0}$, then $(b, x) \in \leq_{i_0} \subseteq \leq_C$ and we are done. If $x \in B_j - B_{i_0}$, then since $\mathcal{C}$ is a chain, we must have $B_{i_0} \subsetneq B_j$ and

$(B_{i_0}, \leq_{i_0}) \prec (B_j, \leq_j)$. Hence by the definition of $\preccurlyeq$, we have $b <_j x$ and $b <_C x$. We are done.

Thirdly, it can be seen $(C, \leq_C)$ is an upper bound of $\mathcal{C}$ (Exercise). We have proved CLAIM.

By CLAIM and Zorn's Lemma, there is a maximal element $(D, \leq_D) \in \mathcal{A}$. It remains to show $A = D$. Note that $D \subseteq A$. So if $D \neq A$, then there is $d \in A - D$. Now let $\leq'_D := \leq_D \cup \{(y, d) | y \in D\}$. Then it is easy to see that $(D, \leq_D) \prec (D \cup \{d\}, \leq'_D) \in \mathcal{A}$ (Exercise). It contradicts the maximality of $(D, \leq_D)$. Therefore $A = D$, and the proof is done.

(2)⇒(1) □

•Applications of Zorn's Lemma

Typical steps of the existence proofs using Zorn's Lemma (e.g. above proof (3)⇒(2)).

1st, consider a family $\mathcal{F}$ of all the candidate sets. $\mathcal{F}$ is suitably ordered (often ordered by inclusion).

2nd, show that every chain $\mathcal{C}$ in $\mathcal{F}$ has an upper bound (often showing $\bigcup \mathcal{C} \in \mathcal{F}$).

3rd, thus by Zorn's Lemma, there is a maximal element M in $\mathcal{F}$. To show M is the desired object, often, arguing that if not, then the maximality of M fails.

Theorem 4.4. *Every vector space V has a basis.*

Proof. Let $\mathcal{F}$ be the family of linearly independent subsets of V. $\mathcal{F}$ is ordered by inclusion. Let $\mathcal{C}$ be a chain in $\mathcal{F}$. Then $\bigcup \mathcal{C} (\subseteq V)$ is also linearly independent. Hence $\bigcup \mathcal{C} \in \mathcal{F}$ and indeed $\bigcup \mathcal{C} = \sup \mathcal{C}$. Hence by Zorn's Lemma, $\mathcal{F}$ has a maximal element, say $B \in \mathcal{F}$. We shall show B is a basis for V. If not, then there is $v_0 (\neq 0) \in V$ such that

$$v_0 \text{ is not a linear combination of any finite subset of } B. \qquad (*)$$

In particular, $v_0 \notin B$, and $B' := B \cup \{v_0\}$ is a proper superset of B.

CLAIM) B' is linearly independent as well: Assume $\alpha_0 \cdot v_0 + \alpha_1 \cdot v_1 ... + \alpha_n \cdot v_n = 0$ where $v_1, ..., v_n \in B$ and scalars α_i. We want to show $\alpha_0 = ... = \alpha_n = 0$. Note that $\alpha_0 = 0$ since otherwise v_0 is a linear combination of $\{v_1, ..., v_n\}$ contradicting (*). Now then since B is linearly independent, $\alpha_1 = ... = \alpha_n = 0$ and we are done.

The claim contradicts then the maximality of B. Hence B must be a basis for V. □

Theorem 4.5. *Every ring with identity has a proper maximal ideal.*

Theorem 4.6. *Let G be a group and let $A \subseteq G$ with $1_G \in A$. Among the subgroups of G contained in A, there is a maximal one.*

Proof. Let $\mathcal{F} = \{H \subseteq A |\ H \text{ is a subgroup of } G\}$. $\mathcal{F}$ is ordered by inclusion. Let $\mathcal{C} (\subseteq \mathcal{F})$ be a chain. We want to show $\mathcal{C}$ has an upper bound. If $\mathcal{C} = \emptyset$, then $\{1_G\} (\in \mathcal{F})$ is an upper bound. Hence we can assume $\mathcal{C} \neq \emptyset$. Now let $C := \bigcup \mathcal{C}$. Then clearly $1_G \in C \subseteq A$. We claim that C is a subgroup of G (and thus $C \in \mathcal{F}$): If $x, y \in C$, then $x \in H_1 \in \mathcal{C}$, $y \in H_2 \in \mathcal{C}$. Since $\mathcal{C}$ is a chain, say $H_1 \subseteq H_2$, and since H_2 is a group, we have $x \cdot y \in H_2 \subseteq C$. Moreover $x^{-1} \in H_1 \subseteq C$. Hence C is a subgroup.

Therefore either case, $\mathcal{C}$ has an upper bound. Hence by Zorn's Lemma, there is a maximal element in $\mathcal{F}$, which is the desired one. $\square$

5. Finite, Countable, and Uncountable Sets

Definition 5.1. Two sets A, B are *equipotent* (or *equinumerous*) (write $A \approx B$) if there is a bijection $f : A \to B$.

Assumption There are sets called *cardinal numbers* (or *cardinals*) such that for any set A, there is a unique cardinal $|A|$ (or written $\mathrm{Card}(A)$, the cardinality of A) equipotent to A.[7] Moreover for $n \in \omega$, we have $|n| = n$, and $|\omega| = \omega$. We use lowercase Greek letters $\kappa, \lambda, \mu..$ to denote cardinals.

Remark 5.2. (1) $A \approx A$.
(2) $A \approx B$ iff $B \approx A$.
(3) If $A \approx B$ and $B \approx C$, then $A \approx C$.
Hence $A \approx B$ iff $\mathrm{Card}(A) = \mathrm{Card}(B)$, and we equivalently say A, B have the same *cardinality* when A, B are equipotent.

Definition 5.3. Write $A \preccurlyeq B$ (or $|A| \leq |B|$) if there is an injection $f : A \to B$.

Remark 5.4. (1) $A \subseteq B \to A \preccurlyeq B$. In particular, $A \preccurlyeq A$.
(2) $A \preccurlyeq B$ and $A \approx C$ implies $C \preccurlyeq B$.
(3) If $A \preccurlyeq B$ and $B \approx C$, then $A \preccurlyeq C$.
(4) $A \preccurlyeq B$ and $B \preccurlyeq C$ implies $A \preccurlyeq C$.

Theorem 5.5. (AC) *If $f : B \to A$ is onto, then $A \preccurlyeq B$. The converse holds if $A \neq \emptyset$.*

Theorem 5.6. (Cantor-Bernstein) *If $|A| \leq |B|$ and $|B| \leq |A|$, then $A \approx B$ (i.e. $\kappa \leq \lambda$ and $\lambda \leq \kappa$ implies $\kappa = \lambda$).*

Theorem 5.7. (AC) *For any cardinals κ, λ, either $\kappa \leq \lambda$ or $\lambda \leq \kappa$.*

Theorem 5.8. (Cantor)
(1) $|A| < |\mathcal{P}(A)|$.
(2) $|\mathcal{P}(A)| = |{}^A 2|$.
(3) $\omega < \mathrm{Card}(\mathbb{R}) = \mathrm{Card}(]0, 1[) = \mathrm{Card}(\mathcal{P}(\omega))$.

The Continuum Hypothesis There is no cardinal κ such that $\omega < \kappa < |\mathbb{R}|$.

– – – – – – – – – – – –

•Finite Sets, Countable Sets

Definition 5.9. (1) We say a set is *finite* having n elements if it is equinumerous to $n (\in \omega)$. A set is *infinite* if it is not finite.
(2) A set A is called *countable* if $|A| = \omega$. It is called *at most countable* if $|A| \leq \omega$, *uncountable* if not at most countable.

Due to 3.7, any subset of a finite set is finite, so any set having an infinite subset is infinite. We shall see that any countable set (in particular ω) is infinite.

[7] This is a temporary assumption. We shall see in Chapter 8 that this follows as a theorem.

Theorem 5.10. (1) *Any infinite subset of a countable set is countable. Hence, any subset of a countable set is either finite or countable.*

(2) *A set B is at most countable iff B is either finite or countable. If $B \neq \emptyset$, then B is at most countable iff there is surjective $f : \omega \to B$.*

Lemma 5.11. (Pigeonhole Principle) *If $X \subsetneq n (\in \omega)$, then $X \not\approx n$.*

Corollary 5.12. (1) *Let $m, n \in \omega$. If $m \approx n$, then $m = n$. Hence if $A \approx m$ and $A \approx n$, then $m = n$. Moreover, $m \leq n$ as natural numbers iff $m \leq n$ as finite cardinals.*

(2) *If A is finite, then for any proper subset B of A, $B \not\approx A$.*

Theorem 5.13. (AC) *The following are equivalent.*

(1) *A is infinite.*
(2) *A has a countable subset.*
(3) *There is $B \subsetneq A$ such that $A \approx B$.*

Theorem 5.14. (1) *If A, B are countable, then so is $A \times B$. Hence for $n \neq 0$, $\omega = |\omega^n| = |{}^n\omega|$.*

(2) *Let $\{A_n |\ n \in \omega\}$ be a family of (at most, resp.) countable sets. Then $\bigcup_{n\in\omega} A_n$ is (at most, resp.) countable. Hence if A, B are (at most, resp.) countable, then so is $A \cup B$.*

(3) *Moreover if A is countable, then both $\mathrm{FinSeq}(A) = A^{<\omega} := \bigcup_{n\in\omega} {}^nA$, and $\{B \subset A |\ B$ is finite$\}$ are countable.*

(4) *Both $\mathbb{Z}, \mathbb{Q}$ are countable.*

6. **Cardinal Numbers**

Definition 6.1. Let κ, λ be cardinals.

(1) $\kappa + \lambda := |A \cup B|$ where $|A| = \kappa, |B| = \lambda$ and A, B are disjoint.
(2) $\kappa \cdot \lambda := |A \times B|$ where $|A| = \kappa, |B| = \lambda$.
(3) $\kappa^\lambda := |{}^B A|$ where $|A| = \kappa, |B| = \lambda$.

To check the definitions are independent from the choice of sets A, B, we need the following lemma.

Lemma 6.2. *Let $A \approx A'$ and $B \approx B'$. Then*

(1) *$A \cup B \approx A' \cup B'$ if $A \cap B = A' \cap B' = \emptyset$.*
(2) *$A \times B \approx A' \times B'$.*
(3) *${}^B A \approx {}^{B'} A'$.*

Corollary 6.3. *$\omega + \omega = \omega$, $\omega \cdot \omega = \omega$, $|{}^A 2| = |\mathcal{P}(A)| = 2^{|A|}$.*

Theorem 6.4. (1) *$\kappa + \lambda = \lambda + \kappa$; $\kappa \cdot \lambda = \lambda \cdot \kappa$.*
(2) *$(\kappa + \lambda) + \mu = \kappa + (\lambda + \mu)$; $(\kappa \cdot \lambda) \cdot \mu = \kappa \cdot (\lambda \cdot \mu)$.*
(3) *$\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$.*
(4) *$\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$; $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$; $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$.*

For finite cardinals (i.e. natural numbers) the addition, multiplication, and exponentiation defined in 6.1 are equal to those in 3.9.

Theorem 6.5. *Assume $\kappa_1 \le \kappa_2$ and $\lambda_1 \le \lambda_2$. Then $\kappa_1 + \lambda_1 \le \kappa_2 + \lambda_2$; $\kappa_1 \cdot \lambda_1 \le \kappa_2 \cdot \lambda_2$; and if $\kappa_1 \neq 0$, $\kappa_1^{\lambda_1} \le \kappa_2^{\lambda_2}$.*

Theorem 6.6. (AC) *For infinite cardinal κ, $\kappa \cdot \kappa = \kappa$.*

Corollary 6.7. (1) *For infinite cardinals κ, λ, $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$.*
(2) *For $2 \le \lambda \le \kappa$ with infinite κ, we have $2^\kappa = \lambda^\kappa = \kappa^\kappa$.*

7. **Ordinal Numbers**

Definition 7.1. We say a set α is an *ordinal* (*number*) if

(1) it is transitive and
(2) well-ordered by ϵ (i.e. $\epsilon_\alpha := \{(x,y)|\ x \in \alpha,\ y \in \alpha,\ x \in y\}$ well-orders α).

We use lower-case Greek letters α, β, γ to denote ordinals.

Fact 7.2. *Each natural number and ω are ordinals. Any element of an ordinal is an ordinal.*

Lemma 7.3. *Let α, β be ordinals.*

(1) $\alpha \notin \alpha$.
(2) $\alpha^+ = \alpha \cup \{\alpha\}$ *is an ordinal.*
(3) $\alpha \cap \beta$ *is an ordinal too.*
(4) $\alpha \in \beta$ *iff* $\alpha \subsetneq \beta$.

For ordinals α, β we write $\alpha < \beta$ iff $\alpha \in \beta$. Therefore $\alpha \leq \beta$ iff ($\alpha \in \beta$ or $\alpha = \beta$) iff $\alpha \subseteq \beta$.

Theorem 7.4. *The class* $\mathrm{OR} := \{\alpha|\ \alpha$ *ordinal* $\}$ *is well-ordered by* $<$.[8] *Namely, for ordinals* α, β, γ,

(1) *if* $\alpha < \beta$ *and* $\beta < \gamma$, *then* $\alpha < \gamma$;
(2) *if* $\alpha < \beta$, *then* $\beta \not< \alpha$;
(3) *exactly one of* $\alpha < \beta$, *or* $\alpha = \beta$, *or* $\beta < \alpha$ *holds;*
(4) *if* $\mathcal{A}$ *is a nonempty set of ordinals, then* $\bigcap \mathcal{A}$ *is the* $<$*-least ordinal in* $\mathcal{A}$.

Moreover,

(5) $\alpha < \beta$ *iff* $\alpha^+ \leq \beta$ *(hence,* α^+ *is a successor of* α*);*
(6) *if* $\mathcal{A}$ *is a set of ordinals, then* $\bigcup \mathcal{A}$ *is an ordinal* $\sup \mathcal{A}$.

Note that for an ordinal α,

$$\alpha = \{x|\ x \in \alpha\} = \{\beta \text{ ordinal} \mid \beta < \alpha\} = \mathrm{OR}[\alpha] = \mathrm{seg}[\alpha].$$

Moreover for $\gamma < \alpha$, it follows $\alpha[\gamma] = \gamma$.

Lemma 7.5. *Any two order isomorphic ordinals are equal.*

Theorem 7.6. *Every well-ordered set is order isomorphic with exactly one ordinal number.*

ZFC: If Well-ordering Isomorphism Theorem holds for proper classes such as OR, then above Theorem 7.6 is a direct consequence of it. Indeed the proof of Theorem 7.6 can be completed by mimicking that of W.I.T., but a careful investigation says **Replacement Axioms** should be used at a step. For more details see the textbook.

Transfinite Induction (First Version) Let $P(x)$ be a property. Assume that for all ordinal α,

if $P(\beta)$ holds for all ordinal $\beta < \alpha$, then $P(\alpha)$ holds.

Then $P(\alpha)$ holds for all α.

[8] Formally, for example, for (1)

$$\forall xyz(\mathrm{OR}(x) \wedge \mathrm{OR}(y) \wedge \mathrm{OR}(z) \wedge x \in y \wedge y \in z \to x \in z)$$

is provable, where $\mathrm{OR}(x)$ is a formula saying x is an ordinal.

Definition 7.7. We call α a *successor ordinal* if $\alpha = \beta^+$ for some β. We call it a *limit ordinal* if it is non-zero and not a successor ordinal.

Transfinite Induction (Second Version) Let $P(x)$ be a property.
Assume that
$P(0)$ holds;
if $P(\alpha)$ holds, then $P(\alpha^+)$ holds;
for a limit α, if $P(\beta)$ holds for all $\beta < \alpha$, then $P(\alpha)$ holds.
Then $P(\alpha)$ holds for all α.

Transfinite Recursion Let $H(x, y)$ be a functional property and let A be a set. Then there is a functional property $F(x, y)$ on OR such that
$F(0) = A$,
$F(\alpha^+) = H(F(\alpha))$,
$F(\alpha) = \bigcup\{F(\beta)|\ \beta < \alpha\}$ for a limit α.[9]

• **Ordinal Arithmetic** We will use transfinite recursion to define + and $\times$ for ordinals.

Definition 7.8. (1) Fix an ordinal γ. We define $+_\gamma$ as follows.
$+_\gamma(0) = \gamma$,
$+_\gamma(\alpha^+) = (+_\gamma(\alpha))^+$ for each ordinal α,
$+_\gamma(\alpha) = \sup\{+_\gamma(\beta)|\ \beta < \alpha\}$ for limit α.
By the Transfinite Recursion Theorem, such $+_\gamma$ is defined. Now define + such that, for ordinals α, β,
$\alpha + \beta := +_\alpha(\beta)$.
Similarly define $\cdot$ and exponentiation between ordinals.
(2) $\gamma \cdot 0 := 0,\ \gamma \cdot \alpha^+ := \gamma \cdot \alpha + \gamma,\ \gamma \cdot \alpha := \sup\{\gamma \cdot \beta|\ \beta < \alpha\}$ for limit α.
(3) $\gamma^0 := 1,\ \gamma^{\alpha^+} := \gamma^\alpha \cdot \gamma,\ \gamma^\alpha := \sup\{\gamma^\beta|\ \beta < \alpha\}$ for limit α.
Due to transfinite induction 2nd version, for ordinals α, β, each of $\alpha + \beta,\ \alpha \cdot \beta,\ \alpha^\beta$ is an ordinal.

Theorem 7.9. *For ordinals α, β if $\alpha \cong A$ and $\beta \cong B$ $(A \cap B = \emptyset)$, then*
(1) $\alpha + \beta \cong A \oplus B$,
(2) $\alpha \cdot \beta \cong A \otimes_{al} B$,
(3) $\beta \cdot \alpha \cong A \otimes_l B$.

• **Laws of Ordinal Arithmetic**

$\alpha^+ = \alpha + 1;\ \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma;$ but $\omega + 1 \neq 1 + \omega$.
$\alpha_1 \leq \alpha_2 \Rightarrow \alpha_1 + \beta \leq \alpha_2 + \beta;\ \alpha_1 + \beta < \alpha_2 + \beta \Rightarrow \alpha_1 < \alpha_2;$ but even if $1 < 2$, we have $1 + \omega = \omega = 2 + \omega$.

[9] **ZFC:** More formally, for any formula $H(x, y)$, there corresponds a formula $F(x, y; A) = F_A(x, y)$ such that
$$\forall x \exists! y H(x, y) \to \forall A[\forall \alpha[\mathrm{OR}(\alpha) \to \ [\exists! z F_A(\alpha, z) \wedge F_A(0, A) \wedge F_A(\alpha^+, H(F_A(\alpha))) \wedge (\alpha \text{ limit } \to F_A(\alpha, \bigcup\{F_A(\beta)|\beta < \alpha\}))]]]$$
is provable.

(Left Cancellation) $\beta + \alpha_1 < \beta + \alpha_2$ iff $\alpha_1 < \alpha_2$; $\beta + \alpha_1 = \beta + \alpha_2$ iff $\alpha_1 = \alpha_2$.

If $\alpha \leq \beta$, then there is a unique γ such that $\alpha + \gamma = \beta$.

$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$; $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$; but $(1+1) \cdot \omega = 2 \cdot \omega = \omega < 1 \cdot \omega + 1 \cdot \omega = \omega \cdot 2$.
$\alpha_1 \leq \alpha_2 \Rightarrow \alpha_1 \cdot \beta \leq \alpha_2 \cdot \beta$; $\alpha_1 \cdot \beta < \alpha_2 \cdot \beta \Rightarrow \alpha_1 < \alpha_2$; but again $1 \cdot \omega = \omega = 2 \cdot \omega$.
(Left Cancellation) $(\beta \neq 0)$ $\beta \cdot \alpha_1 < \beta \cdot \alpha_2$ iff $\alpha_1 < \alpha_2$; $\beta \cdot \alpha_1 = \beta \cdot \alpha_2$ iff $\alpha_1 = \alpha_2$.

(Division Theorem) For any α and nonzero β, there are a unique $\gamma < \beta$ and a unique δ such that $\alpha = \beta \cdot \delta + \gamma$.

$(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$; $\alpha^\beta \cdot \alpha^\gamma = \alpha^{\beta + \gamma}$; but $(2 \cdot 2)^\omega = \omega < 2^\omega \cdot 2^\omega = \omega^2$.
$\alpha_1 \leq \alpha_2 \Rightarrow {\alpha_1}^\beta \leq {\alpha_2}^\beta$; ${\alpha_1}^\beta < {\alpha_2}^\beta \Rightarrow \alpha_1 < \alpha_2$; but for $2 < 3$, we have $2^\omega = 3^\omega = \omega$.
$(1 < \beta)$ $\beta^{\alpha_1} < \beta^{\alpha_2}$ iff $\alpha_1 < \alpha_2$; $\beta^{\alpha_1} = \beta^{\alpha_2}$ iff $\alpha_1 = \alpha_2$.

0,1,2,...,$\omega, \omega + 1, \omega + 2, ..., \omega + \omega = \omega \cdot 2, \omega \cdot 2 + 1 ..., \omega \cdot 3, ..., \omega \cdot \omega = \omega^2, ..., \omega^3, ..., \omega^\omega$, $\omega^\omega + 1, ..., \omega^\omega \cdot 2, ..., \omega^\omega \cdot \omega = \omega^{\omega+1}, ..., \omega^{\omega^2}, ..., \omega^{\omega^\omega}, ..., \epsilon, ...$
where $\epsilon = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \omega^{\omega^{\omega^\omega}}, ...\}$.

Theorem 7.10. (The Normal Form Theorem) *Every ordinal $\alpha > 0$ is uniquely expressed as*

$$\alpha = \omega^{\beta_1} \cdot k_1 + ... + \omega^{\beta_n} \cdot k_n$$

where $\beta_1 > ... > \beta_n$, and $0 < k_i \in \omega$ $(i = 1, ...n)$.

8. **Alephs**

Definition 8.1. By an *initial* ordinal we mean an ordinal α not equinumerous to any ordinal $\beta < \alpha$.

Theorem 8.2. (AC) *Any set is equinumerous to a unique initial ordinal.*

Now as promised, we define cardinal numbers as initial ordinals, and for a set A, its cardinality $|A|$ is the unique cardinal as in Theorem 8.2. So now we can remove Assumption in Chapter 5.[10] Note also that **cardinal and ordinal arithmetic are different** even if they use the same notation for addition, multiplication, and exponentiation. For example as ordinal addition $2 + \omega = \omega \neq \omega + 2$, while as cardinal addition, $2 + \omega = \omega = \omega + 2$. Similarly, as ordinals $2^\omega = \omega$ and $\omega < \omega^2$ while as cardinals $\omega < 2^\omega$ and $\omega = \omega^2$. **But up to the context, it should be clear whether what is stated is for ordinal, or cardinal arithmetic.**[11]

Both CD := $\{\kappa|\ \kappa$ is a cardinal $\}$, IC := $\{\kappa|\ \kappa$ is an infinite cardinal $\}$ are proper classes.

Theorem 8.3. *There is an order isomorphic function (functional property)* $\aleph : \mathrm{OR} \to \mathrm{IC}$.[12]

$\aleph_0 = \omega, ..., \omega^2, ..., \omega^\omega, ..., \epsilon, ..., \aleph_1, ..., \aleph_2, ..., \aleph_\omega, ...$

The Continuum Hypothesis $2^{\aleph_0} = \aleph_1$.

Definition 8.4. $\beth : \mathrm{OR} \to \mathrm{IC}$ is defined as
$\beth_0 := \omega$
$\beth_{\alpha+1} := 2^{\beth_\alpha}$
$\beth_\alpha := \sup\{\beth_\beta|\ \beta < \alpha\}$ for limit α.

The Generalized Continuum Hypothesis For all ordinal α, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$, equivalently, $\aleph_\alpha = \beth_\alpha$.

[10] In most advanced book for Axiomatic Set Theory, ordinals are introduced first, then cardinals and related topics are developed later.

[11] In other branches of mathematics such as Algebra and Analysis, cardinal arithmetic is mostly used.

[12] **ZFC:** Formally there is a formula $\aleph(x, y)$ so that a sentence

$$\forall\alpha\exists!\kappa(\aleph(\alpha,\kappa) \wedge \mathrm{IC}(\kappa)) \wedge \forall\alpha\beta(\alpha \in \beta \leftrightarrow \aleph(\alpha) \in \aleph(\beta)) \wedge \forall\kappa(\mathrm{IC}(\kappa) \to \exists\alpha\aleph(\alpha,\kappa))$$

is provable from ZFC. Here $\forall\alpha$..., $\exists\alpha$... stand for $\forall x(\mathrm{OR}(x) \to$ ..., $\exists x(\mathrm{OR}(x) \wedge$ ..., respectively.

9. **Advanced Topics**

$\alpha, \beta, \gamma...$ denote ordinals and $\kappa, \lambda, \mu...$ denote cardinals.

• More on Cardinal Arithmetic

Definition 9.1. For a cardinal κ, we let κ^+, a *successor* of κ, be a successor cardinal in the ordering of CD (i.e. if $\kappa = \aleph_\alpha$, then $\kappa^+ = \aleph_{\alpha+1}$).

Hence we call κ, a *successor cardinal* if it is a successor of some cardinal. Call κ a *limit cardinal* if $\omega < \kappa$ and κ is not a successor.

For a function $f : \lambda \to \kappa$ (both infinite), we say f is *cofinal* if $\operatorname{ran} f$ has no upper bound in κ.

$\operatorname{cf}(\kappa) :=$ the least cardinal in $\{\lambda |\ \exists \text{ cofinal } f : \lambda \to \kappa\}$.

For an infinite κ, we say κ is *regular* if $\operatorname{cf}(\kappa) = \kappa$, and say κ *singular* if not regular.

Clearly $\operatorname{cf}(\kappa) \leq \kappa$, and $\operatorname{cf}(\aleph_\omega) = \omega$.

Theorem 9.2. *Any infinite successor cardinal is regular.*

Definition 9.3. Let κ_i $(i \in I)$ be cardinals. Then $\sum_{i\in I} \kappa_i := \operatorname{Card}(\bigcup \mathcal{A})$ where $\mathcal{A} = \{A_i |\ i \in I\}$ is an indexed family of mutually disjoint sets such that $\kappa_i = \operatorname{Card}(A_i)$.

We define $\kappa^{<\lambda} := \sum_{\mu<\lambda} \kappa^\mu$.

Theorem 9.4. (König's Theorem) *If $\kappa_i < \lambda_i$ $(i \in I)$, then*

$$\sum_{i\in I} \kappa_i < \operatorname{Card}(\prod_{i\in I} \lambda_i).$$

Corollary 9.5. *For an infinite κ, we have $\kappa < \kappa^{\operatorname{cf}(\kappa)}$, and $\kappa < \operatorname{cf}(2^\kappa)$. Hence (in* ZFC*) it follows*

$$2^{\aleph_0} \neq \aleph_\omega.$$

• ZFC: Foundation Axiom

Theorem 9.6. *There do not exist sets $A_1, ..., A_n$ such that $A_1 \in A_2 ... \in A_n \in A_1$. In particular $B \notin B$, for any set B.*

Proof. Suppose not, and let $A = \{A_1, ..., A_n\} \neq \emptyset$. Then and for each $A_i \in A$,

$A_n \in A_1 \cap A \neq \emptyset$ if $i = 1$;

$A_{i-1} \in A_i \cap A \neq \emptyset$ if $1 < i \leq n$.

Hence both cases contradict Axiom of Foundation. □

By Transfinite Recursion, for each α, we define a set V_α as follows:

$V_0 := \emptyset$,

$V_{\alpha+1} := \mathcal{P}(V_\alpha)$,

$V_\alpha := \bigcup\{V_\beta |\ \beta < \alpha\}$ for limit α.

Each set in V_α is said to be *well-founded.*

Indeed there is a formula $\mathrm{WF}(x, y)$ such that for α, $\mathrm{WF}(\alpha, y)$ says $y \in V_\alpha$. Thus $\forall y \exists \alpha\ \mathrm{WF}(\alpha, y)$ says 'every set is well-founded'.

Theorem 9.7. (In ZF-{Axiom of Foundation}) *Axiom of Foundation holds iff every set is well-founded.*

Appendix: ZFC Axioms for Set Theory

Definition We recursively define *formulas* in set theory.

Variables in set theory are symbols $x, y, z, ..., a, b, c, ..., X, Y, Z, ..., A, B, C, ...$

Formulas:

(1) For variables u, v, we let both $u = v,\ u \in v$ be formulas.
(2) If φ, ψ are formulas, then each of $(\neg\varphi), (\varphi \to \psi)$ is a formula.
(3) If φ is a formula and u is a variable, then $\forall u\varphi$ is a formula.
(4) Nothing else is a formula unless it can be obtained by finitely many applications of (1)(2) and (3).

For example, $\sigma := \forall x((x \in a \to x \in b) \to \forall a\ x \in a)$ is a formula. Here the variable x is bounded by the quantifier $\forall$, where as a, b have unbounded occurrence (even if a once occurs boundedly). We call such variables having unbounded occurrence *free variables* in the formula σ, and we may write σ as $\sigma(a, b)$. It means the formula σ indeed is a *property* of a and b. So in formal set theory, when we say a property P on $x_1, ..., x_n$, it always is referred to a formula $P(x_1, ..., x_n)$ having free variables $x_1, ..., x_n$.

If $\sigma' := \forall a(\neg\forall b\ \sigma(a, b)) = \forall a(\neg\forall b\forall x((x \in a \to x \in b) \to \forall a\ x \in a))$, then now every variable is bounded. We call such a formula having no free variable a *sentence* in set theory. Believe or not, every mathematical statement can be translated into a single set theory sentence. This is one of the reasons why we take set theory as the foundation of mathematics.

Now we introduce axioms and the rules of inferences for set theory. Axioms consist of logical axioms and nonlogical axioms (ZFC axioms).

Logical Axioms: Any formulas of the following forms.

(1) $\varphi \to (\psi \to \varphi)$.[13]
(2) $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$.
(3) $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$.
(4) $\forall u(\varphi \to \psi) \to (\varphi \to \forall u\psi)$ where u is not free in φ.
(5) $\forall u\varphi(u, u_1, ..., u_n) \to \varphi(v, u_1, ..., u_n)$, as far as v remains free at each replacement of free occurrence of u in $\varphi(u, u_1, ..., u_n)$ by v.
(6) $x = y \to (x \in z \to y \in z)$.

ZFC Axioms: We will state these later.

Rules of inferences:

(1) ψ is obtained from φ and $\varphi \to \psi$.
(2) $\forall u\varphi$ is obtained from φ.

We say a formula φ is a *theorem* (or equivalently *is provable* in set theory) if there is a sequence of formulas $\varphi_0, ..., \varphi_n$ such that $\varphi = \varphi_n$ and for each $i \leq n$, φ_i is either an axiom (logical or ZFC), or obtained from previous formulas by the rules of inferences (i.e. for some $j, k < i$, we have $\varphi_k = \varphi_j \to \varphi_i$ or $\varphi_i = \forall u\varphi_j$). Again all the mathematical statements so

[13]For the rest, we omit some of parentheses in the formulas by taking linkage strength convention $(\forall, \exists, \neg > \wedge, \vee > \to, \leftrightarrow)$.

far mathematicians have proved indeed have these proof sequences.

ZFC (Zermelo-Fraenkel + Choice) Axioms

Notation
$\varphi \vee \psi$ abbreviates $\neg\varphi \to \psi$,
$\varphi \wedge \psi$ abbreviates $\neg(\neg\varphi \vee \neg\psi)$,
$\varphi \leftrightarrow \psi$ abbreviates $(\varphi \to \psi) \wedge (\psi \to \varphi)$.
$\exists u\varphi$ abbreviates $\neg\forall u\neg\varphi$.
$\exists! u\varphi(u, v_1, ..., v_n)$ abbreviates $\exists u(\varphi(u, v_1, ..., v_n) \wedge \forall v(\varphi(v, v_1, ..., v_n) \to u = v))$ (There is a unique u satisfying $\varphi(u, v_1, ..., v_n)$).
$B \subseteq A$ abbreviates $\forall x(x \in B \to x \in A)$.
u^+ abbreviates $u \cup \{u\}$.

Existence of a set $\exists A(A = A)$. This is a theorem obtained only from logical axioms. Intended meaning: There exists a set.

Axiom of Extensionality $\forall A\forall B(A = B \leftrightarrow \forall x(x \in A \leftrightarrow x \in B))$. Intended meaning: Two sets are equal iff they have exactly the same elements.

Axiom Schema[14] **of Comprehension**(=**Subset Axioms**) (text book p. 8) For each formula $\varphi(x)$ having a free variable x, the formula $\forall A\exists B\ \forall x(x \in B \leftrightarrow x \in A \wedge \varphi(x))$. Intended meaning: For each set A, there is a set $B = \{x \in A|\ \varphi(x)\}$.

Existence of the empty set Above axioms imply $\exists A\ \forall x(x \in A \leftrightarrow \neg x = x)$, the existence of $\emptyset = \{x|x \neq x\}$.

Axiom of Pairing $\forall a\forall b\exists A\ \forall z(z \in A \leftrightarrow z = a \vee z = b)$. Intended meaning: For any sets a, b, the set $A = \{a, b\}$ exists.

Axiom of Union $\forall A\exists B\ \forall x(x \in B \leftrightarrow \exists y(x \in y \wedge y \in A))$. Intended meaning: For any set A, there exists $\bigcup A = \{x|\exists y(x \in y \wedge y \in A)\}$.

Axiom of Power Set $\forall A\exists\mathcal{P}\ \forall B(B \in \mathcal{P} \leftrightarrow B \subseteq A)$. Intended meaning: For any set A, there is the power set $\mathcal{P}(A) = \{B|\ B \subseteq A\}$.

Axiom Schema of Replacement[15] (text book p. 112) For each formula $\psi(x, y)$, the formula $\forall x\exists! y\psi(x, y) \to \forall A\exists B\ \forall w(w \in A \to \exists z(z \in B \wedge \psi(w, z)))$. Intended meaning: If $\psi(x, y)$ satisfies the functional property, then for any set A, there is a set B containing

[14] Not a single axiom but a set of axioms.

[15] If we take the following version of Replacement Axioms then those imply above Comprehension Axioms: For each formula $\psi(x, y)$, the formula $\forall x\forall y\forall z(\psi(x, y) \wedge \psi(x, z) \to y = z) \to \forall A\exists B\forall z(z \in B \leftrightarrow \exists w(w \in A \wedge \psi(w, z)))$. Intended meaning: If for each x there is at most one y such that $\psi(x, y)$, then for any set A, there is the set $B = \{z|\ \exists w(w \in A \wedge \psi(w, z))\}$, the image of A under ψ.

$\{z|\ \exists w(w \in A \wedge \psi(w,z))\}$, the image of A under ψ.

Axiom of Infinity $\exists A(\emptyset \in A \wedge \forall x(x \in A \to x^+ \in A))$.[16] Intended meaning: There is an inductive set.

Axiom of Foundation $\forall A(A \neq \emptyset \to \exists x(x \in A \wedge x \cap A = \emptyset))$. Intended meaning: Any nonempty set A has an element disjoint with A.

Choice Axiom $\forall \mathcal{F}[\forall B, C(B, C \in \mathcal{F} \to B \neq \emptyset \wedge (B \neq C \to B \cap C = \emptyset)) \to \exists C\, \forall A(A \in \mathcal{F} \to \exists! z(z \in A \cap C))]$. Intended meaning: If $\mathcal{F}$ is a family of nonempty mutually disjoint sets (i.e. $\mathcal{F}$ is a partition), then there is a set C collecting exactly one element from each set in $\mathcal{F}$.

The reader should be convinced English mathematical statements can be translated into set theoretic sentences, and vice versa. For example, in ZFC, the class $\{x|x \notin x\}$ does not exist. It simply means $\neg\exists y \forall x(x \in y \leftrightarrow x \notin x)$ is provable. Comprehension Axiom only allows the existence of a set $B = \{x \in A|\ x \notin x\}$ from a set A. Now if $B \in B$ then B must satisfy the property $x \notin x$, so $B \notin B$, a contradiction. Hence we have $B \notin B$. But $B \notin B$ need not imply $B \in B$. Thus Russell's paradox is avoidable. Another example is Theorem 0.1. Namely,

$$\forall A \forall B(\forall x(x \in A \to x \in B) \leftrightarrow \forall x(x \in A \leftrightarrow (x \in A \wedge x \in B)))$$

is provable.

[16] The reader can easily figure out the full form of this abbreviated formula.