Set Theory

2008/09

by Peter Koepke

*Das Wesen der Mathematik
liegt in ihrer Freiheit* (Georg Cantor)

*Aus dem Paradies, das Cantor uns geschaffen,
soll uns niemand vertreiben können* (David Hilbert)

1 Introduction

What is Mathematics?

This a grand question and the main challenge for the philosophy of mathematics. A modest attempt would be to say that mathematics deals with *mathematical objects* and describes them in a *mathematical language*. Mathematical objects are usually considered or treated as objectively existing, like numbers or geometric objects, or as created by definitions or existence postulates. It seems that numbers like 0, 1, 2, ... are simply "there" and can be treated like physical objects, although this pragmatic viewpoint becomes problematic on closer inspection. One can define complicated objects like the collection of all twin primes, a collection which appears to be very concrete and very illusive at the same time. Properties of existing and/or created mathematical objects are expressed in a precise formal language which for the layman is a hallmark of mathematics.

By all accounts, existence in mathematics is different from ordinary existence. Even if the collection of twin primes "exists", we do not attempt to grasp it by our usual means of perceptions. We do not go near it to measure or see it, like we would with all other kinds of physical objects. If somebody claims that the collection is infinite, we would only accept a *mathematical proof* of this statement and not any other kind of evidence like physical experiments or observations, historical records, or public opinions. So even if we hold that the collection of twin primes exists, we treat it in a hypothetical and formal way: if *T were* the collection of twin primes then *T would* be infinite.

Depending on personal world views a similar criticism, though not very pragmatic, could im principle be held against ordinary existence of objects like tables and chairs. The natural sciences, confronted with a myriad of different phenomena have always tried to understand them through abstraction and unification. One can see tables and chairs made up from common building blocks like molecules and atoms. This analysis surely misses some aspects of a chair, but has been incredibly helpful for the advancement of knowledge. In physics the search is always on for encompassing world formulas. It is conceivable that such formulas do not speak about concrete physical objects but about some illusive notions like spaces, fields, symmetries, whose existence might be as problematic as that of mathematical objects.

Can we thus proceed to find mathematical "world formulas" which talk about some fundamental objects and which govern the behaviour or familiar higher objects like numbers and figures? A unified foundation for all of mathematics would explain and secure the applicability of one mathematical theory to another one. If number theory and geometry are parts of one bigger consistent theory then it becomes clear why one cannot derive results in number theory which contradict results in geometry.

What should the fundamental objects for mathematics be? Since we cannot expect them to exist in naive ways there will be many degrees of freedom, and the choice between various proposals is a matter of practicality, aesthetics, and agreement among the practitioneers. It is fair to say that *set theory* has been accepted as the basic theory for mathematics almost universally. We shall see that numbers, spaces, functions, figures, etc. can be considered as sets and that appropriate set theoretical assumptions imply that numbers, spaces, etc. have their familiar properties. This is analogous to the explanation of macroscopic friction, say, by assuming the the surfaces of bodies are made up from atoms with mutual attractions and repulsions. The microscopic theory is justified if it can provide macroscopic theories which correspond to our expectations. Set theory satisfies that criterion to a very high degree.

After the establishment of a successful theory, that theory poses questions of its own. In a theory of atoms one can ask what happens when particles collide with high velocities in a vacuum, although this question may seem irrelevant to the friction of tables and chairs. Similarly, set theory has developed its own questions. These are mostly concerned with *infinity*, transcending the concrete infinities given by collections of numbers or space points. One can view the *theory of infinity* as going to extreme frontiers of mathematics and mathematical knowledge.

At these frontiers there are many mathematical questions which cannot be decided by the standard assumptions and methods. In contrast to DAVID HILBERT's "es gibt kein *ignorabimus*" one is able to show rigorously that certain statement cannot be proved or disproved from the axioms. This is analogous to the undecidability of the parallel postulate from the other geometrical axioms. In set theory we shall encounter and prove such phenomena in connection with the axiom of choice. It is a tribute to mathematics that mathematics has often been able to *mathematically prove* its limitations, be it in geometry, algebra, logic, or set theory.

Elements of Set Theory

We want to exemplify some aspects of set theory which will be crucial for the further development.

Set-theoretic Reconstructions of Mathematical Objects

In current mathematics, *many* notions are explicitly defined using sets, e.g., a geometric figure is a set of points, usually given by some definition. The following example indicates that also notions which are not set theoretical *prima facie* can be construed set theoretically:

> f is a real funktion $\equiv$ f is a **set** of ordered pairs $(x, f(x))$ of real numbers;
>
> (x, y) is an ordered pair $\equiv$ (x, y) is a **set** ...$\{x, y\}$... ;
>
> x is a real number $\equiv$ x is a left half of a DEDEKIND cut in $\mathbb{Q}$ $\equiv$ x is a **subset** of $\mathbb{Q}$, such that ... ;
>
> r is a rational number $\equiv$ r is an **ordered pair** of integers, such that ... ;
>
> z is an integer $\equiv$ z is an **ordered pair** of natural numbers ($=$ non-negative integers);
>
> $\mathbb{N} = \{0, 1, 2, ...\}$;
>
> 0 is the empty **set**;
>
> 1 is the **set** $\{0\}$;
>
> 2 is the **set** $\{0, 1\}$; etc. etc.

We shall see that *all* mathematical notions can be reduced to the notion of *set*.

Sets

GEORG CANTOR characterized sets as follows:

> Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten, wohlunterschiedenen Objekten m unsrer Anschauung oder unseres Denkens (welche die "Elemente" von M genannt werden) zu einem Ganzen.

If m is an *element* of M one writes $m \in M$. If all mathematical objects are reducible to sets, *both sides* of these relation have to be sets. This means that set theory studies the $\in$-relation $m \in M$ for arbitrary *sets* m and M. As it turns out, this is also sufficient for the purposes of set theory and mathematics. In set theory variables range over the class of all sets, the $\in$-relation is the only undefined structural component, every other notion will be defined from the $\in$-relation. Basically, set theoretical statement will thus be of the form

$$...\forall x...\exists y......x \in y...u \equiv v...,$$

belonging to the first-order predicate language with the only given predicate $\in$. To deal with the complexities of set theory and mathematics one develops a comprehensive and intuitive aparatus of abbreviations and definitions which allows to write familiar statements like

$$e^{i\pi} = -1$$

and to view them as a statement of set theory. The language of set theory may be seen as a low-level, internal language. The language of mathematics possesses high-level "macro" expressions which abbreviate low-level statements in an efficient and intuitive way.

Infinite Cardinalities

The infinite possesses unusual and partially counterintuitive properties. The set $\mathbb{E} = \{0, 2, 4, 6, ...\}$ of even numbers has just as many elements as the set $\mathbb{N} = \{0, 1, 2, 3, ...\}$ of all natural numbers since the map

$$f: \mathbb{N} \leftrightarrow \mathbb{E}, n \mapsto 2 \cdot n$$

is a bijection between $\mathbb{N}$ and $\mathbb{E}$. Note that all these notions like natural number, odd, even, map, bijection will later be redefined as set-theoretical notions. One says that $\mathbb{N}$ and $\mathbb{E}$ have the same *cardinality*, and that they are both *countable*. This equality contradicts the traditional principle that "the part is smaller than the whole".

On December 12, 1873, GEORG CANTOR made a crucial discovery which may be taken as the beginning of set theory:

Theorem 1. *The set $\mathbb{R}$ of real numbers is* not *countable.*

Proof. Assume that $f: \mathbb{N} \leftrightarrow \mathbb{R}$ is bijective. Define a decimal number

$$r = 0, r_0 r_1 r_2 ...$$

by:

$\quad$ $r_i = 0$, if the i-th decimal position of $f(i)$ is a 1, and $r_i = 1$ otherwise.

Since f is surjective, take a natural number n such that $r = f(n)$. By definition $r_n \neq$ the n-th decimal position of $f(n)$ and thus

$$r_n \neq \text{the } n\text{-th decimal position of } r = r_n.$$

Contradiction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

By this fundamental theorem the cardinality of the set $\mathbb{R}$ is strictly bigger than the cardinality of $\mathbb{N}$. CANTOR's *diagonal argument* can be applied to other sets, yielding higher and higher cardinalities. Thus the theory of infinite sets encompasses a rich theory of higher infinities. Infinitary combinatorics has some unusual properties. E.g., the real line $\mathbb{R}$ and the real plane $\mathbb{R} \times \mathbb{R}$ have the same cardinality, higher dimension does not necessarily lead to higher cardinalities.

The Axiom of Choice

One often constructs sequences $(a_i)_{i \in I}$ as follows:

$\quad$ for $i \in I$ choose a_i such that a_i satisfies ... .

This seems to be a straightforward generalization of an unproblematic procedure for finite sets I. The possibility of *infinitely many* simultaneous choices cannot be proved directly and is usually postulated as the *axiom of choice* (AC). The axiom of choice is important in many areas of mathematics. It is crucial for a smooth theory of infinite cardinalities. On the other hand it has problematic, counterintuitive consequences.

The ZERMELO-FRAENKEL *Axioms of Set Theory*

The critical discussion of the axiom of choice and related principles lead to the first axiomatization of set theory by ERNST ZERMELO. We shall deal with the axiom system ZF named after ZERMELO and ABRAHAM FRAENKEL. It is now recognized as the standard and universal axiomatization of set theory. After sufficiently developing ZERMELO-FRAENKEL set theory we shall study axiomatic aspects of set theory.

The Continuum Hypothesis

After proving that the set $\mathbb{R}$ is uncountable, CANTOR formulated the *continuum hypothesis* according to which $\mathbb{R}$ possesses the next cardinality above the cardinality of $\mathbb{N}$:

$$\text{card}(\mathbb{N}) = \aleph_0 \text{ and } \text{card}(\mathbb{R}) = \aleph_1 \,.$$

This may be reformulated as:

> for every infinite set $X \subseteq \mathbb{R}$ of real numbers, either there exists a bijection $X \leftrightarrow \mathbb{N}$ or there exists a bijection $X \leftrightarrow \mathbb{R}$.

The continuum and its generalizations imply strong combinatorial properties.

Independence

The ZF-axioms are subject to the *incompleteness theorems* of KURT GÖDEL: there are set theoretical statements which cannot be proved or disproved from the axioms. This independence does not only affect self-referential statements similar to the *liar paradoxon* ("this sentence cannot be proved from the axioms"), but also natural properties like the axiom of choice or the continuum hypothesis. We shall prove this by constructing differing *models of set theory* in which these statements are true or false.

2 The ZERMELO-FRAENKEL Axioms

CANTOR's naive description of the notion of set suggests that for any mathematical statement in $\varphi(x)$ in one free variable x there is a *set y* such that

$$x \in y \leftrightarrow \varphi(x),$$

i.e., y is the collection of all sets x which satisfy φ . Setting $\varphi(x)$ to be $x \notin x$ this becomes

$$x \in y \leftrightarrow x \notin x \,,$$

and in particular for $x = y$:

$$y \in y \leftrightarrow y \notin y.$$

This *contradiction* is usually denoted RUSSELL's paradox. It shows that the formation of sets as collections of sets by arbitrary formulas is not consistent. ZERMELO's main idea was to restrict the formulas allowed in the formation of sets. The following axiom system extends the original ZERMELO axioms by contributions by FRAENKEL, MIRIMANOFF, and SKOLEM.

Definition 2. *The system* ZF *of the* ZERMELO-FRAENKEL *axioms of set theory* *consists of the following axioms:*

  a) *The* set existence axiom (Ex)*:*

$$\exists x \forall y \neg y \in x$$

- *there is a set without elements, the* empty set*.*

b) *The* axiom of extensionality (Ext)*:*

$$\forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

- *a set is determined by its elements, sets having the same elements are identical.*

c) *The* pairing axiom (Pair)*:*

$$\forall x \forall y \exists z \forall w (u \in z \leftrightarrow u = x \lor u = y).$$

- *z is the unordered pair of x and y.*

d) *The* union axiom (Union)*:*

$$\forall x \exists y \forall z (z \in y \leftrightarrow \exists w (w \in x \land z \in w))$$

- *y is the union of all elements of x.*

e) *The* separation schema (Sep) *postulates for every $\in$-formula $\varphi(z, x_1, ..., x_n)$:*

$$\forall x_1 ... \forall x_n \forall x \exists y \forall z (z \in y \leftrightarrow z \in x \land \varphi(z, x_1, ..., x_n))$$

- *this is an infinite scheme of axioms, the set z consists of all elements of x which satisfy φ.*

f) *The* powerset axiom (Pow)*:*

$$\forall x \exists y \forall z (z \in y \leftrightarrow \forall w (w \in z \rightarrow w \in x))$$

- *y consists of all subsets of x.*

g) *The* replacement schema (Rep) *postulates for every $\in$-formula $\varphi(x, y, x_1, ..., x_n)$:*

$$\forall x_1 ... \forall x_n (\forall x \forall y \forall y' ((\varphi(x, y, x_1, ..., x_n) \land \varphi(x, y', x_1, ..., x_n)) \rightarrow y = y') \rightarrow$$
$$\forall u \exists v \forall y (y \in v \leftrightarrow \exists x (x \in u \land \varphi(x, y, x_1, ..., x_n))))$$

- *v is the image of u under the map defined by φ.*

h) *The* axiom of infinity (Inf)*:*

$$\exists x (\exists y (y \in x \land \forall z \neg z \in y) \land \forall y (y \in x \rightarrow \exists z (z \in x \land \forall w (w \in z \leftrightarrow w \in y \lor w = y))))$$

- *by the closure properties of x, x has to be infinite.*

i) *The* foundation schema (Found) *postulates for every $\in$-formula $\varphi(x, x_1, ..., x_n)$:*

$$\forall x_1 ... \forall x_n (\exists x \varphi(x, x_1, ..., x_n) \rightarrow \exists x (\varphi(x, x_1, ..., x_n) \land \forall x' (x' \in x \rightarrow \neg \varphi(x', x_1, ..., x_n))))$$

- *if φ is satisfiable then there are $\in$-minimal elements satisfying φ.*

Note that the axiom system is an infinite *set* of axioms. It seems unavoidable that we have to go back to some previously given set notions to be able to define the collection of set theoretical axioms - another example of the frequent circularity in foundational theories.

We shall discuss the axioms one by one and simultaneously introduce the logical language and useful conventions.

2.1 Set Existence

The *set existence axiom*

$$\exists x \forall y \neg y \in x,$$

like all axioms, is expressed in a language with quantifiers $\exists$ ("there exists") and $\forall$ ("for all"), which is familiar from the ϵ-δ-statements in analysis. The *language of set theory* uses variables $x, y, ...$ which may satisfy the binary relations $\in$ or $=$: $x \in y$ ("x is an *element of* y") or $x = y$. These elementary *formulas* may be connected by the *propositional connectives* $\land$ ("and"), $\lor$ ("or"), $\rightarrow$ ("implies"), $\leftrightarrow$ ("is equivalent"), and $\neg$ ("not"). The use of this language will be demonstrated by the subsequent axioms.

2.2 Extensionality

The *axiom of extensionality*

$$\forall x \forall x'(\forall y(y \in x \leftrightarrow y \in x') \rightarrow x = x')$$

expresses that a set is exactly determined by the collection of its elements. This allows to prove that there is exactly one empty set.

Lemma 3. $\forall x \forall x'(\forall y \, \neg y \in x \land \forall y \, \neg y \in x' \rightarrow x = x')$.

Proof. Consider x, x' such that $\forall y \, \neg y \in x \land \forall y \, \neg y \in x'$. Consider y. Then $\neg y \in x$ and $\neg y \in x'$. This implies $\forall y(y \in x \leftrightarrow y \in x')$. The axiom of extensionality implies $x = x'$. $\square$

Note that this is a *formal proof* in the sense of mathematical logic. The sentences of the proof can be derived from earlier ones by purely formal deduction rules. The rules of natural deduction correspond model common sense figures of argumentation which treat hypothetical objects as if they would concretely exist.

2.3 Pairing

The *pairing axiom*

$$\forall x \forall y \exists z \forall u(u \in z \leftrightarrow u = x \lor u = y)$$

postulates that for all sets x, y there is set z which may be denoted as

$$z = \{x, y\}.$$

This notation abbreviates the formula

$$\forall u(u \in z \leftrightarrow u = x \lor u = y).$$

The language of mathematics which we are about to introduce will consist of many such abbreviations. The abbreviations are chosen for intuitive, pragmatic, or historical reasons. Using the notation for unordered pairs, the pairing axiom may be written as

$$\forall x \forall y \exists z \; z = \{x, y\}.$$

By the axiom of extensionality, the term-like notation has the expected behaviour. E.g.:

Lemma 4. $\forall x \forall y \forall z \forall z' \, (z = \{x, y\} \land z' = \{x, y\} \rightarrow z = z')$.

Proof. Exercise. $\square$

Note that we use a number of notational conventions: variables have to be chosen in a reasonable way, for example the symbols z and z' in the lemma have to be taken different and different from x and y. We also assume some operator priorities to reduce the number of brackets: we let $\land$ bind stronger than $\lor$, and $\lor$ stronger than $\rightarrow$ and $\leftrightarrow$.

We used the "term" $\{x, y\}$ to occur within set theoretical formulas. This abbreviation is than to be expanded in a natural way, so that officially all mathematical formulas are formulas in the "pure" $\in$-language. We want to see the notation $\{x, y\}$ as an example of a *class term*. We define uniform notations and convention for such abbreviation terms.

Definition 5. *A* class term *is of the form* $\{x | \varphi\}$ *where x is a variable and $\varphi \in L^{\in}$. If $\{x | \varphi\}$ and $\{y | \psi\}$ are class terms then*

  – $u \in \{x | \varphi\}$ *stands for* $\varphi \frac{u}{x}$ *, where* $\varphi \frac{u}{x}$ *is obtained from φ by (resonably) substituting the variable x by the variable u ;*

  – $u = \{x | \varphi\}$ *stands for* $\forall v(v \in u \leftrightarrow \varphi \frac{v}{x})$;

  – $\{x | \varphi\} = u$ *stands for* $\forall v(\varphi \frac{v}{x} \leftrightarrow v \in u)$;

    — $\{x|\varphi\} = \{y|\psi\}$ *stands for* $\forall v(\varphi\frac{v}{x} \leftrightarrow \psi\frac{v}{y})$;

    — $\{x|\varphi\} \in u$ *stands for* $\exists v(v \in u \wedge v = \{x|\varphi\})$;

    — $\{x|\varphi\} \in \{y|\psi\}$ *stands for* $\exists v(\psi\frac{v}{y} \wedge v = \{x|\varphi\})$.

A term *is either a variable or a class term.*

We shall further extend this notation give suggestive or traditional names to important formulas and class terms.

Definition 6.

    a) $\emptyset := \{x|x \neq x\}$ *is the* empty set;

    b) $V := \{x|x = x\}$ *is the* universe *(of all sets)*;

    c) $\{x,y\} := \{u|u = x \vee u = y\}$ *is the* unordered pair *of x and y*.

Lemma 7.

    a) $\emptyset \in V$.

    b) $\forall x,y \, \{x,y\} \in V$.

Proof. a) $\emptyset \in V$ abbreviates the formula

$$\exists v(v = v \wedge v = \emptyset).$$

This is equivalent to $\exists v \, v = \emptyset$ which again is an abbreviation for

$$\exists v \, \forall w \, (w \in v \leftrightarrow w \neq w).$$

This is equivalent to $\exists v \forall w \, w \notin v$ which is equivalent to the axiom of set existence. So $\emptyset \in V$ is another way to write the axiom of set existence.
b) $\forall x,y \, \{x,y\} \in V$ abbreviates the formula

$$\forall x,y \exists z(z = z \wedge z = \{x,y\}).$$

This can be expanded equivalently to the pairing axiom

$$\forall x,y \exists z \forall u(u \in z \leftrightarrow u = x \vee u = y). \qquad \square$$

So a) and b) are concise equivalent formulations of the axiom Ex and Pair.
We also introduce *bounded quantifiers* to simplify notation.

Definition 8. *Let A be a term. Then* $\forall x \in A \, \varphi$ *stands for* $\forall x(x \in A \rightarrow \varphi)$ *and* $\exists x \in A \, \varphi$ *stands for* $\exists x \, (x \in A \wedge \varphi)$.

Definition 9. *Let* $x,y,z,...$ *be variables and* $X,Y,Z,...$ *be class terms. Define*

    a) $X \subseteq Y := \forall x \in X \, x \in Y$, *X is a* subclass *of Y;*

    b) $X \cup Y := \{x|x \in X \vee x \in Y\}$ *is the* union *of X and Y;*

    c) $X \cap Y := \{x|x \in X \wedge x \in Y\}$ *is the* intersection *of X and Y;*

    d) $X \setminus Y := \{x|x \in X \wedge x \notin Y\}$ *is the* difference *of X and Y;*

    e) $\bigcup X := \{x|\exists y \in X \, x \in y\}$ *is the* union *of X;*

    f) $\bigcap X := \{x|\forall y \in X \, x \in y\}$ *is the* intersection *of X;*

    g) $\mathcal{P}(X) := \{x|x \subseteq X\}$ *is the* power class *of X;*

    h) $\{X\} := \{x|x = X\}$ *is the* singleton set *of X;*

    i) $\{X,Y\} := \{x|x = X \vee x = Y\}$ *is the* (unordered) pair *of X and Y;*

    j) $\{X_0,...,X_{n-1}\} := \{x|x = X_0 \vee ... \vee x = X_{n-1}\}$.

One can prove the well-known boolean properties for these operations. We only give a few examples.

Proposition 10. $X \subseteq Y \wedge Y \subseteq X \rightarrow X = Y$.

Proposition 11. $\bigcup \{x, y\} = x \cup y$.

Proof. We show the equality by two inclusions:
($\subseteq$). Let $u \in \bigcup \{x, y\}$. $\exists v (v \in \{x, y\} \wedge u \in v)$. Let $v \in \{x, y\} \wedge u \in v$. $(v = x \vee v = y) \wedge u \in v$.
Case 1. $v = x$. Then $u \in x$. $u \in x \vee u \in y$. Hence $u \in x \cup y$.
Case 2. $v = y$. Then $u \in y$. $u \in x \vee u \in y$. Hence $u \in x \cup y$.
    Conversely let $u \in x \cup y$. $u \in x \vee u \in y$.
Case 1. $u \in x$. Then $x \in \{x, y\} \wedge u \in x$. $\exists v(v \in \{x, y\} \wedge u \in v)$ and $u \in \bigcup \{x, y\}$.
Case 2. $u \in y$. Then $x \in \{x, y\} \wedge u \in x$. $\exists v(v \in \{x, y\} \wedge u \in v)$ and $u \in \bigcup \{x, y\}$. $\square$

> **Exercise 1.** Show: a) $\bigcup V = V$. b) $\bigcap V = \emptyset$. c) $\bigcup \emptyset = \emptyset$. d) $\bigcap \emptyset = V$.

Combining objects into ordered pairs (x, y) is taken as an undefined fundamental operation of mathematics. We cannot use the unordered pair $\{x, y\}$ for this purpose, since it does not respect the order of entries:

$$\{x, y\} = \{y, x\}.$$

We have to introduce some asymmetry between x and y to make them distinguishable. Following Kuratowski and Wiener we define:

Definition 12. $(x, y) := \{\{x\}, \{x, y\}\}$ *is the* ordered pair *of x and y.*

The definition involves substituting class terms within class terms. We shall see in the following how these class terms are eliminated to yield pure $\in$-formulas.

Lemma 13. $\forall x \forall y \exists z \, z = (x, y)$.

Proof. Consider sets x and y. By the pairing axiom choose u and v such that $u = \{x\}$ and $v = \{x, y\}$. Again by pairing choose z such that $z = \{u, v\}$. We argue that $z = (x, y)$. Note that
    $(x, y) = \{\{x\}, \{x, y\}\} = \{w \mid w = \{x\} \vee w = \{x, y\}\}$.
Then $z = (x, y)$ is equivalent to
    $\forall w (w \in z \leftrightarrow w = \{x\} \vee w = \{x, y\})$,
    $\forall w (w = u \vee w = v \leftrightarrow (w = \{x\} \vee w = \{x, y\})$,
and this is true by the choice of u and v. $\square$

The Kuratowski-pair satisfies the fundamental property of ordered pairs:

Lemma 14. $(x, y) = (x', y') \rightarrow x = x' \wedge y = y'$.

Proof. Assume $(x, y) = (x', y')$, i.e.,
(1) $\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}$.
Case 1. $x = y$. Then
    $\{x\} = \{x, y\}$,
    $\{\{x\}, \{x, y\}\} = \{\{x\}, \{x\}\} = \{\{x\}\}$,
    $\{\{x\}\} = \{\{x'\}, \{x', y'\}\}$,
    $\{x\} = \{x'\}$ and $x = x'$,
    $\{x\} = \{x', y'\}$ and $y' = x$.
Hence $x = x'$ and $y = x = y'$ as required.
Case 2. $x \neq y$. (1) implies
    $\{x'\} = \{x\}$ or $\{x'\} = \{x, y\}$.
The right-hand side would imply $x = x' = y$, contradicting the case assumption. Hence
    $\{x'\} = \{x\}$ and $x' = x$.
Then (1) implies
    $\{x, y\} = \{x', y'\} = \{x, y'\}$ and $y = y'$. $\square$

Exercise 2.

a) Show that $\langle x, y \rangle := \{\{x, \emptyset\}, \{y, \{\emptyset\}\}\}$ also satisfies the fundamental property of ordered pairs (F. HAUSDORFF).

b) Can $\{x, \{y, \emptyset\}\}$ be used as an ordered pair?

Exercise 3. Give a set-theoretical formalization of an ordered-*triple* operation.

Ordered pairs allow to introduce *relations* and *functions* in the usual way. One has to distinguish between *sets* which are relations and functions, and *class terms* which are relations and functions.

Definition 15. *A term R is a* relation *if all elements of R are ordered pairs, i.e., $R \subseteq V \times V$. Also write Rxy or xRy instead of $(x, y) \in R$. If A is a term and $R \subseteq A \times A$ then R is a* relation *on A.*

Note that this definition is really an *infinite schema* of definitions, with instances for all terms R and A. The subsequent extensions of our language are also infinite definition schemas. We extend the term language by parametrized collections of terms.

Definition 16. *Let $t(\vec{x})$ be a term in the variables $\vec{x}$ and let φ be an $\in$ -formula. Then $\{t(\vec{x})|\varphi\}$ stands for $\{z|\exists \vec{x}\,(\varphi \wedge z = t(\vec{x}))\}$.*

Definition 17. *Let R, S, A be terms.*

a) *The* domain *of R is* $\operatorname{dom}(R) := \{x\,|\,\exists y\, x\,R\,y\}$.

b) *The* range *of R is* $\operatorname{ran}(R) := \{y\,|\,\exists x\, x\,R\,y\}$.

c) *The* field *of R is* $\operatorname{field}(R) := \operatorname{dom}(R) \cup \operatorname{ran}(R)$.

d) *The* restriction *of R to A is* $R \!\restriction\! A := \{(x, y)|x\,R\,y \wedge x \in A\}$.

e) *The* image *of A under R is* $R[A] := R''A := \{y\,|\,\exists x \in A\, x\,R\,y\}$.

f) *The* preimage *of A under R is* $R^{-1}[A] := \{x\,|\,\exists y \in A\, x\,R\,y\}$.

g) *The* composition *of S and R ("S after R") is* $S \circ R := \{(x, z)|\exists y\,(x\,R\,y \wedge y\,S\,z)\}$.

h) *The* inverse *of R is* $R^{-1} := \{(y, x)|x\,R\,y\}$.

Relations can play different roles in mathematics.

Definition 18. *Let R be a relation.*

a) *R is* reflexive *iff* $\forall x \in \operatorname{field}(R)\, x\,R\,x$.

b) *R is* irreflexive *iff* $\forall x \in \operatorname{field}(R)\, \neg x\,R\,x$.

c) *R is* symmetric *iff* $\forall x, y\,(x\,R\,y \rightarrow y\,R\,x)$.

d) *R is* antisymmetric *iff* $\forall x, y\,(x\,R\,y \wedge y\,R\,x \rightarrow x = y)$.

e) *R is* transitive *iff* $\forall x, y, z\,(x\,R\,y \wedge y\,R\,z \rightarrow x\,R\,z)$.

f) *R is* connex *iff* $\forall x, y \in \operatorname{field}(R)\,(x\,R\,y \vee y\,R\,x \vee x = y)$.

g) *R is an* equivalence relation *iff R is reflexive, symmetric and transitive.*

h) *Let R be an equivalence relation. Then $[x]_R := \{y|y\,R\,x\}$ is the* equivalence class *of x modulo R.*

It is possible that an equivalence class $[x]_R$ is not a set: $[x]_R \notin V$. Then the formation of the collection of all equivalence classes modulo R may lead to contradictions. Another important family of relations is given by *order relations*.

Definition 19. *Let R be a relation.*

a) *R is a* partial order *iff R is reflexive, transitive and antisymmetric.*

b) R *is a* linear order *iff* R *is a connex partial order.*

c) *Let* A *be a term. Then* R *is a* partial order on A *iff* R *is a partial order and* $\text{field}(R) = A$.

d) R *is a* strict partial order *iff* R *is transitive and irreflexive.*

e) R *is a* strict linear order *iff* R *is a connex strict partial order.*

Partial orders are often denoted by symbols like $\leqslant$, and strict partial orders by $<$. A common notation in the context of (strict) partial orders R is to write

$$\exists p\,Rq\,\varphi \text{ and } \forall p\,Rq\,\varphi \text{ for } \exists p\,(p\,Rq \wedge \varphi) \text{ and } \forall p\,(p\,Rq \rightarrow \varphi) \text{ resp.}$$

One of the most important notions in mathematics is that of a *function*.

Definition 20. *Let* F *be a term. Then* F *is a* function *if it is a relation which satisfies*

$$\forall x, y, y'\,(x\,Fy \wedge x\,Fy' \rightarrow y = y').$$

If F *is a function then*

$$F(x) := \{u\,|\,\forall y\,(x\,Fy \rightarrow u \in y)\}$$

is the value *of* F *at* x.

If F is a function and $x\,Fy$ then $y = F(x)$. If there is no y such that $x\,Fy$ then $F(x) = V$; the "value" V at x may be read as "undefined". A function can also be considered as the (indexed) sequence of its values, and we also write

$$(F(x))_{x \in A} \text{ or } (F_x)_{x \in A} \text{ instead of } F\colon A \rightarrow V.$$

We define further notions associated with functions.

Definition 21. *Let* F, A, B *be terms.*

a) F *is a* function from A to B, *or* $F\colon A \rightarrow B$, *iff* F *is a function,* $\text{dom}(F) = A$, *and* $\text{range}(F) \subseteq B$.

b) F *is a* partial function from A to B, *or* $F\colon A \rightharpoonup B$, *iff* F *is a function,* $\text{dom}(F) \subseteq A$, *and* $\text{range}(F) \subseteq B$.

c) F *is a* surjective *function from* A *to* B *iff* $F\colon A \rightarrow B$ *and* $\text{range}(F) = B$.

d) F *is an* injective *function from* A *to* B *iff* $F\colon A \rightarrow B$ *and*

$$\forall x, x' \in A\,(x \neq x' \rightarrow F(x) \neq F(x'))$$

e) F *is a* bijective *function from* A *to* B, *or* $F\colon A \leftrightarrow B$, *iff* $F\colon A \rightarrow B$ *is surjective and injective.*

f) $^{A}B := \{f\,|\,f\colon A \rightarrow B\}$ *is the class of all functions from* A *to* B.

One can check that these functional notions are consistent and agree with common usage:

Exercise 4. Define a relation $\sim$ on V by

$$x \sim y \longleftrightarrow \exists f\,f\colon x \leftrightarrow y.$$

One say that x and y are *equinumerous* or *equipollent*. Show that $\sim$ is an equivalence relation on V. What is the equivalence class of $\emptyset$? What is the equivalence class of $\{\emptyset\}$?

Exercise 5. Consider functions $F\colon A \rightarrow B$ and $F'\colon A \rightarrow B$. Show that

$$F = F' \text{ iff } \forall a \in A\ F(a) = F'(a).$$

2.4 Unions

The *union axiom* reads

$$\forall x \exists y \forall z\,(z \in y \leftrightarrow \exists w\,(w \in x \wedge z \in w)).$$

Lemma 22. *The union axiom is equivalent to* $\forall x \bigcup x \in V$.

Proof. Observe the following equivalences:

$\forall x \bigcup x \in V$
$\leftrightarrow \forall x \exists y (y = y \wedge y = \bigcup x)$
$\leftrightarrow \forall x \exists y \forall z (z \in y \leftrightarrow z \in \bigcup x)$
$\leftrightarrow \forall x \exists y \forall z (z \in y \leftrightarrow \exists w \in x\, z \in w)$

which is equivalent to the union axiom. $\qquad\square$

Note that the union of x is usually viewed as the union of all *elements* of x:

$$\bigcup x = \bigcup_{w \in x} w,$$

where we define

$$\bigcup_{a \in A} t(a) = \{z \,|\, \exists a \in A\, z \in t(a)\}.$$

Combining the axioms of pairing and unions we obtain:

Lemma 23. $\forall x_0, ..., x_{n-1} \{x_0, ..., x_{n-1}\} \in V$.

Note that this is a *schema* of lemmas, one for each ordinary natural number n. We prove the schema by complete induction on n.

Proof. For $n = 0, 1, 2$ the lemma states that $\emptyset \in V$, $\forall x \{x\} \in V$, and $\forall x, y \{x, y\} \in V$ resp., and these are true by previous axioms and lemmas. For the induction step assume that the lemma holds for n, $n \geqslant 1$. Consider sets $x_0, ..., x_n$. Then

$$\{x_0, ..., x_n\} = \{x_0, ..., x_{n-1}\} \cup \{x_n\}.$$

The right-hand side exists in V by the inductive hypothesis and the union axiom. $\qquad\square$

2.5 Separation

It is common to form a subset of a given set consisting of all elements which satisfy some condition. This is codified by the *separation schema*. For every $\in$-formula $\varphi(z, x_1, ..., x_n)$ postulate:

$$\forall x_1 ... \forall x_n \forall x \exists y \forall z\, (z \in y \leftrightarrow z \in x \wedge \varphi(z, x_1, ..., x_n)).$$

Using class terms the schema can be reformulated as: for every term A postulate

$$\forall x\, A \cap x \in V.$$

The crucial point is the restriction to the given set x. The unrestricted, FREGEan version $A \in V$ for every term A leads to the RUSSELL antinomy. We turn the antinomy into a consequence of the separation schema:

Theorem 24. $V \notin V$.

Proof. Assume that $V \in V$. Then $\exists x\, x = V$. Take x such that $x = V$. Let R be the RUSSELLian class:

$$R := \{x \,|\, x \notin x\}.$$

By separation, $y := R \cap x \in V$. Note that $R \cap x = R \cap V = R$. Then

$$y \in y \leftrightarrow y \in R \leftrightarrow y \notin y,$$

contradiction. $\qquad\square$

This simple but crucial theorem leads to the distinction:

Definition 25. *Let A be a term. Then A is a* proper class *iff $A \notin V$.*

Set theory deals with sets and proper classes. Sets are the favoured objects of set theory, the axiom mainly state favorable properties of sets and set existence. Sometimes one says that a term A *exists* if $A \in V$. The intention of set theory is to construe important mathematical classes like the collection of natural and real numbers as sets so that they can be treated set-theoretically. Zermelo observed that this is possible by requiring some set existences together with the *restricted* separation principle.

Exercise 6. Show that the class $\{\{x\}|x \in V\}$ of *singletons* is a proper class.

2.6 Power Sets

The *power set axiom* in class term notation is

$$\forall x \, \mathcal{P}(x) \in V.$$

The power set axiom yields the existence of function spaces.

Definition 26. *Let A, B be terms. Then*

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

is the cartesian product *of A and B.*

Exercise 7.

By the specific implementation of Kuratowski ordered pairs:

Lemma 27. $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B))$.

Proof. Let $(a,b) \in A \times B$. Then

$$
\begin{aligned}
a, b &\in A \cup B \\
\{a\}, \{a, b\} &\subseteq A \cup B \\
\{a\}, \{a, b\} &\in \mathcal{P}(A \cup B) \\
(a,b) = \{\{a\}, \{a, b\}\} &\subseteq \mathcal{P}(A \cup B) \\
(a,b) = \{\{a\}, \{a, b\}\} &\in \mathcal{P}(\mathcal{P}(A \cup B))
\end{aligned}
$$

$\square$

Theorem 28.

a) $\forall x, y \; x \times y \in V.$

b) $\forall x, y \; {}^{x}y \in V.$

Proof. Let x, y be sets. a) Using the axioms of pairing, union, and power sets, $\mathcal{P}(\mathcal{P}(x \cup y)) \in V$. By the previous lemma and the axiom schema of separation,

$$x \times y = (x \times y) \cap \mathcal{P}(\mathcal{P}(x \cup y)) \in V.$$

b) ${}^{x}y \subseteq \mathcal{P}(x \times y)$ since a function $f : x \to y$ is a subset of $x \times y$. By the separation schema,

$${}^{x}y = {}^{x}y \cap \mathcal{P}(x \times y) \in V.$$
$\square$

Note that to "find" the sets in this theorem one has to apply the power set operation repeatedly. We shall see that the universe of all sets can be obtained by iterating the power set operation.

The power set axiom leads to higher *cardinalities*. The theory of cardinalities will be developed later, but we can already prove Cantor's theorem:

Theorem 29. *Let $x \in V$.*

a) *There is an injective map $f : x \to \mathcal{P}(x)$.*

b) *There does not exist an injective map $g : \mathcal{P}(x) \to x$.*

Proof. a) Define the map $f \colon x \to \mathcal{P}(x)$ by $u \mapsto \{u\}$. This is a set since

$$f = \{(u, \{u\}) \mid u \in x\} \subseteq x \times \mathcal{P}(x) \in V.$$

f is injective: let $u, u' \in x$, $u \neq u'$. By extensionality,

$$f(u) = \{u\} \neq \{u'\} = f(u').$$

b) Assume there were an injective map $g \colon \mathcal{P}(x) \to x$. Define the CANTORean set

$$c = \{u \mid u \in x \wedge u \notin g^{-1}(u)\} \in P(x).$$

Let $u_0 = g(c)$. Then $g^{-1}(u_0) = c$ and

$$u_0 \in c \leftrightarrow u_0 \notin g^{-1}(u_0) = c.$$

Contradiction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2.7 Replacement

If every element of a set is definably *replaced* by another set, the result is a set again. The *schema of replacement* postulates for every term F:

$$F \text{ is a function } \to \forall x\, F[x] \in V.$$

Lemma 30. *The replacement schema implies the separation schema.*

Proof. Let A be a term and $x \in V$.
Case 1. $A \cap x = \emptyset$. Then $A \cap x \in V$ by the axiom of set existence.
Case 2. $A \cap x \neq \emptyset$. Take $u_0 \in A \cap x$. Define a map $F \colon x \to x$ by

$$F(u) = \begin{cases} u, & \text{if } u \in A \cap x \\ u_0, & \text{else} \end{cases}$$

Then by replacement

$$A \cap x = F[x] \in V$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2.8 Infinity

All the axioms so far can be realized in a domain of finite sets, see exercise ???. The true power of set theory is set free by postulating the existence of *one* infinite set and continuing to assume the axioms. The *axiom of infinity* expresses that the set of "natural numbers" exists. To this end, some "number-theoretic" notions are defined.

Definition 31.

    a) $0 := \emptyset$ *is the number* zero.

    b) For any term t, $t + 1 := t \cup \{t\}$ is the successor *of t.*

These notions are reasonable in the later formalization of the natural numbers. The axiom of infinity postulates the existence of a set which contains 0 and is closed under successors

$$\exists x\, (0 \in x \wedge \forall n \in x\ n + 1 \in x).$$

Intuitively this says that there is a set which contains all natural numbers. Let us define set-theoretic analogues of the standard natural numbers:

Definition 32. *Define*

    a) $1 := 0 + 1$*;*

    b) $2 := 1 + 1$*;*

c) $3 := 2 + 1; \dots$

From the context it will be clear, whether "3", say, is meant to be the standard number "three" or the set theoretical object

$$
\begin{aligned}
3 &= 2 \cup \{2\} \\
&= (1+1) \cup \{1+1\} \\
&= (\{\emptyset\} \cup \{\{\emptyset\}\}) \cup \{\{\emptyset\} \cup \{\{\emptyset\}\}\} \\
&= \{\emptyset, \{\emptyset\}, \{\emptyset\} \cup \{\{\emptyset\}\}\}.
\end{aligned}
$$

The set-theoretic axioms will ensure that this interpretation of "three" has all important number-theoretic properties of "three".

2.9 Foundation

The *axiom schema of foundation* provides structural information about the set theoretic universe V. It can be reformulated by postulating, for any term A:

$$
A \neq \emptyset \rightarrow \exists x \in A \; A \cap x = \emptyset.
$$

Viewing $\in$ as some kind of order relation this means that every non-empty class has an $\in$-minimal element $x \in A$ such that the $\in$-predecessors of x are not in A. Foundation excludes circles in the $\in$-relation:

Lemma 33. *Let n be a natural number $\geqslant 1$. Then there are no $x_0, \dots, x_{n-1}$ such that*

$$
x_0 \in x_1 \in \dots \in x_{n-1} \in x_0.
$$

Proof. Assume not and let $x_0 \in x_1 \in \dots \in x_{n-1} \in x_0$. Let

$$
A = \{x_0, \dots, x_{n-1}\}.
$$

$A \neq \emptyset$ since $n \geqslant 1$. By foundation take $x \in A$ such that $A \cap x = \emptyset$.
Case 1. $x = x_0$. Then $x_{n-1} \in A \cap x = \emptyset$, contradiction.
Case 2. $x = x_i$, $i > 0$. Then $x_{i-1} \in A \cap x = \emptyset$, contradiction. $\square$

Exercise 8. Show that $x \neq x + 1$.

Exercise 9. Show that the successor function $x \mapsto x + 1$ is injective.

Exercise 10. Show that the term $\{x, \{x, y\}\}$ may be taken as an ordered pair of x and y.

Theorem 34. *The foundation scheme is equivalent to the following,* Peano-*type, induction scheme: for every term B postulate*

$$
\forall x \, (x \subseteq B \rightarrow x \in B) \rightarrow B = V.
$$

This says that if a "property" B is inherited by x if all elements of x have the property B, then every set has the property B.

Proof. $(\rightarrow)$ Assume B were a term which did not satisfy the induction principle:

$$
\forall x \, (x \subseteq B \rightarrow x \in B) \text{ and } B \neq V.
$$

Set $A = V \setminus B \neq \emptyset$. By foundation take $x \in A$ such that $A \cap x = \emptyset$. Then

$$
u \in x \rightarrow u \notin A \rightarrow u \in B,
$$

i.e., $x \subseteq B$. By assumption, B is inherited by $x : x \in B$. But then $x \notin A$, contradiction.
$(\leftarrow)$ Assume A were a term which did not satisfy the foundation scheme:

$$
A \neq \emptyset \text{ and } \forall x \in A \; A \cap x \neq \emptyset.
$$

Set $B = V \setminus A$. Consider $x \subseteq B$. Then $A \cap x = \emptyset$. By assumption, $x \notin A$ and $x \in B$. Thus $\forall x \, (x \subseteq B \rightarrow x \in B)$. The induction principle implies that $B = V$. Then $A = \emptyset$, contradiction. $\square$

This proof shows, that the induction principle is basically an equivalent formulation of the foundation principle. The $\in$-relation is taken as some binary relation without reference to specific properties of this relation. This leads to:

Exercise 11. A relation R on a domain D is called *wellfounded*, iff for all terms A

$$\emptyset \neq A \wedge A \subseteq D \to \exists x \in A \; A \cap \{y \mid yRx\} = \emptyset.$$

Formulate and prove a principle for R-induction on D which coressponds to the assumption that R is well-founded on D.

2.10 Axiom Systems

Using class terms, the ZF can be formulated concisely:

Theorem 35. *The* ZF *axioms are equivalent to the following system; we take all free variables of the axioms to be universally quantified:*

a) *Ex:* $\emptyset \in V$.

b) *Ext:* $x \subseteq y \wedge y \subseteq x \to x = y$.

c) *Pair:* $\{x, y\} \in V$.

d) *Union:* $\bigcup x \in V$.

e) *Sep:* $A \cap x \in V$.

f) *Pow:* $\mathcal{P}(x) \in V$.

g) *Rep: F is a function* $\to F[x] \in V$.

h) *Inf:* $\exists x \, (0 \in x \wedge \forall n \in x \; n + 1 \in x)$.

i) *Found:* $A \neq \emptyset \to \exists x \in A \; A \cap x = \emptyset$.

This axiom system can be used as a foundation for nearly all of mathematics. Axiomatic set theory considers various axiom systems of set theory.

Definition 36. *The axiom system* ZF^- *consists of the* ZF*-axioms except the power set axiom. The system* EML *("elementary set theory") consists of the axioms* Ex, Ext, Pair, *and* Union.