Set Theory

2012/13

by Peter Koepke

Die Mengenlehre ist das Fundament
der gesamten Mathematik
(Felix Hausdorff,
Grundzuge der Mengenlehre, 1914)

1 Introduction

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.

Felix Hausdorff in *Grundzuge* formulated shorter:

> Eine Menge ist eine Zusammenfassung von Dingen zu einem Ganzen, d.h. zu einem neuen Ding.

Sets are ubiquitous in mathematics. According to Hausdorff

> Differential- und Integralrechnung, Analysis und Geometrie arbeiten in Wirklichkeit, wenn auch vielleicht in verschleiernder Ausdrucksweise, bestandig mit unendlichen Mengen.

In current mathematics, *many* notions are explicitely defined using sets. The following example indicates that 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, such that ... ;
>
> (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*.

Besides this foundational role, set theory is also the mathematical study of the *infinite*. There are infinite sets like $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$ which can be subjected to the constructions and analyses of set theory; there are various degrees of infinity which lead to a rich theory of infinitary combinatorics.

In this course, we shall first apply set theory to obtain the standard foundation of mathematics and then turn towards "pure" set theory.

2 The Language of Set Theory

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 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 language of abbreviations and definitions which, eventually, allows to write familiar statements like

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

and to view them as statements within 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.

3 Russell's Paradox

Cantor's naive description of the notion of set suggests that for any mathematical statement $\varphi(x)$ 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.

This axiom is a basic principle in Gottlob Frege's *Grundgesetze der Arithmetik, 1893*, Grundgesetz V, Grundgesetz der Wertverlaufe.

Bertrand Russell noted in 1902 that setting $\varphi(x)$ to be $x\notin x$ this becomes

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

and in particular for $x=y$:

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

Contradiction.

This contradiction is usually called Russell's paradox, antinomy, contradiction. It was also discovered slightly earlier by Ernst Zermelo. The paradox shows that the formation of sets as collections of sets by *arbitrary* formulas is not consistent.

4 The Zermelo-Fraenkel Axioms

The difficulties around Russell's paradox and also around the axiom of choice lead Zermelo to the formulation of axioms for set theory in the spirit of the axiomatics of David Hilbert of whom Zermelo was an assistant at the time.

Zermelo's main idea was to restrict Frege's Axiom V to formulas which correspond to mathematically important formations of collections, but to avoid arbitrary formulas which can lead to paradoxes like the one exhibited by Russell.

The original axiom system of Zermelo was extended and detailed by Abraham Fraenkel (1922), Dmitry Mirimanoff (1917/20), and Thoralf Skolem.

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

4.1 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* $\wedge$ ("and"), $\vee$ ("or"), $\rightarrow$ ("implies"), $\leftrightarrow$ ("is equivalent"), and $\neg$ ("not"). The use of this language will be demonstrated by the subsequent axioms.

The axiom expresses the existence of a set which has no elements, i.e., the existence of the *empty set*.

4.2 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 1. $\forall x \forall x' (\forall y \neg y \in x \wedge \forall y \neg y \in x' \rightarrow x = x')$.

Proof. Consider x, x' such that $\forall y \neg y \in x \wedge \forall y \neg y \in x'$. 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 $x = x'$. □

Note that this proof is a usual mathematical argument, and it is also 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 to common sense figures of argumentation which treat hypothetical objects as if they would concretely exist.

4.3 Pairing

The *pairing axiom*

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

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

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

This formula, including the new notation, is equivalent to the formula

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

In the sequel we shall extend the small language of set theory by hundreds of symbols and conventions, in order to get to the ordinary language of mathematics with notations like

$$\mathbb{N}, \mathbb{R}, \sqrt{385}, \pi, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \int_a^b f'(x) \mathrm{dx} = f(b) - f(a), \text{etc.}$$

Such notations 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 2. $\forall x \forall y \forall z \forall z' (z = \{x, y\} \wedge z' = \{x, y\} \rightarrow z = z')$.

Proof. Exercise. □

Note that we implicitly use several 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 $\wedge$ bind stronger than $\vee$, and $\vee$ 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.

The extended language of set theory contains class terms and notations for them. There are axioms for class terms that fix how extended formulas can be reduced to formulas in the unextended $\in$-language of set theory.

Definition 3. *A* class term *is of the form* $\{x|\varphi\}$ *where* x *is a variable and* $\varphi \in L^{\in}$. *The usage of these class terms is defined recursively by the following axioms: If* $\{x|\varphi\}$ *and* $\{y|\psi\}$ *are class terms then*

- $u \in \{x|\varphi\} \leftrightarrow \varphi\frac{u}{x}$, *where* $\varphi\frac{u}{x}$ *is obtained from* φ *by (resonably) substituting the variable* x *by the variable* u*;*
- $u = \{x|\varphi\} \leftrightarrow \forall v\, (v \in u \leftrightarrow \varphi\frac{v}{x})$*;*
- $\{x|\varphi\} = u \leftrightarrow \forall v\, (\varphi\frac{v}{x} \leftrightarrow v \in u)$*;*
- $\{x|\varphi\} = \{y|\psi\} \leftrightarrow \forall v\, (\varphi\frac{v}{x} \leftrightarrow \psi\frac{v}{y})$*;*
- $\{x|\varphi\} \in u \leftrightarrow \exists v(v \in u \wedge v = \{x|\varphi\}$*;*
- $\{x|\varphi\} \in \{y|\psi\} \leftrightarrow \exists v(\psi\frac{v}{y} \wedge v = \{x|\varphi\})$.

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

Definition 4.

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 5.

a) $\emptyset \in V$.

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

Proof. a) By the axioms for the reduction of abstraction terms, $\emptyset \in V$ is equivalent to the following formulas

$$\exists v(v = v \wedge v = \emptyset)$$
$$\exists v\, v = \emptyset$$
$$\exists v\, \forall w\, (w \in v \leftrightarrow w \neq w)$$
$$\exists v \forall w\, w \notin v$$

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

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

So a) and b) are concise equivalent formulations of the axiom Ex and Pair.

We also introduce *bounded quantifiers* to simplify notation.

Definition 6. *Let A be a term. Then* $\forall x\in A\,\varphi\leftrightarrow\forall x(x\in A\rightarrow\varphi)$ *and* $\exists x\in A\,\varphi\leftrightarrow\exists x\,(x\in A\wedge\varphi)$.

Definition 7. *Let $x,y,z,\ldots$ be variables and $X,Y,Z,\ldots$ be class terms. Define*

a) $X\subseteq Y\leftrightarrow\forall x\in X\,x\in Y$, *$X$ 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) $\{X_0,\ldots,X_{n-1}\}:=\{x|x=X_0\vee\ldots\vee x=X_{n-1}\}$.

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

Proposition 8. $X\subseteq Y\wedge Y\subseteq X\rightarrow X=Y$.

Proposition 9. $\bigcup\{x,y\}=x\cup y$.

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 $u\in\bigcup\{x,y\}$. ☐

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 10. $(x,y):=\{\{x\},\{x,y\}\}$ *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 11. $\forall x\forall y\exists z\,z=(x,y)$.

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

$(x,y)=\{\{x\},\{x,y\}\}=\{w|w=\{x\}\vee w=\{x,y\}\}$.

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. ☐

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

Lemma 12. $(x,y)=(x',y')\rightarrow x=x'\wedge y=y'$.

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 $y=y'$. □

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 13. *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 14. *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 15. *Let R,S,A be terms.*

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

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

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

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

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

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

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

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