Set Theory
2012/13

by Peter Koepke

Die Mengenlehre ist das Fundament
der gesamten Mathematik
(Felix Hausdorff,
Grundzüge 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. [G. Cantor, Beiträge zur Begründung der transfiniten Mengenlehre, *Mathematische Annalen*, 1895]

Felix Hausdorff in *Grundzüge* 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, beständig mit unendlichen Mengen.

In current mathematics, *many* notions are explicitely defined using sets. 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, 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 and the Language of Mathematics

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 statements will thus be of the form

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

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.

We shall thus be dealing with *two* languages: A minimalistic language whose only undefined symbol is $\in$ and a rich language which contains relation, function, and constant symbols for “all” mathematical notions. We shall call the minimalistic language the *language of set theory* or the $\in$-*language* or briefly $L^{\in}$, and the rich language the *language of mathematics*. The notions of the language of mathematicas will be defined or definable in terms of the $\in$-language so that - in principle - every mathematical notion can be reduced to a formula of the $\in$-language. The language of set theory may be viewed as a low-level, internal language whereas the language of mathematics possesses high-level “macro” expressions which abbreviate low-level statements in efficient and intuitive ways.

So we shall practically work in a rich language, successively being extended definitions. For theoretical arguments, however, it is often convenient to assume that all formulas are $\in$-formulas in the language of set theory. By the previous considerations, these two approaches are equivalent.

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), called *Grundgesetz* V, *Grundgesetz der Wertverläufe*.

BERTRAND RUSSELL noted in 1902 that taking $\varphi(x)$ 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, or contradiction. It was also discoved 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

In 1908, the difficulties around Russell's paradox and also around the axiom of choice led 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 gradually introduce the mathematical language, together with useful notations and conventions.

4.1 Set Existence

The *set existence axiom* is the statement

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

Like all axioms, it 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 $\in$-language $L^{\in}$ will be demonstrated in the development of the subsequent axioms.

The set existence 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 $x = x'$. $\square$

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 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 $\in$-language of set theory by symbols and conventions, in order to reach 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 use implicitly 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 then to be reduced to the "pure" $\in$-language in natural ways. 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.

4.4 Class Terms

We build the language of mathematics using class terms and notations for them. There are axioms for class terms that fix how extended formulas can be reduced to formulas in the $\in$-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.*

In the class term $\{x|\varphi\}$ we may also allow φ to be a formula of the language of mathematics, since every such formula is equivalent to an $L^{\in}$-formula.

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

$$\begin{aligned}&\exists v(v = v \wedge v = \emptyset)\\&\exists v\, v = \emptyset\\&\exists v\, \forall w\, (w \in v \leftrightarrow w \neq w)\\&\exists v \forall w\, w \notin v\end{aligned}$$

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 intuitive equivalent formulations of the axioms of extensionality and pairing. Often one omits initial universal quantifiers and writes pairing more concisely as

$$\{x, y\} \in V.$$

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 well-known boolean properties for the operations $\cup$ and $\cap$. 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) $\bigcap \emptyset = V$.

4.5 Ordered Pairs

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.

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 (HAUSDORFF).

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

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

4.6 Relations and Functions

Ordered pairs allow to introduce *relations* and *functions*. One has to distinguish between *sets* which are relations and functions, and *class terms* which are relations and functions. We extend the term language by parametrized collections of terms.

Definition 13. *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 14. *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.*

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 these definitions are really *infinite schemas* of definitions, with instances for all terms A, B, R.

Definition 16. *Let R, S, A, B 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* $R\restriction A:=\{(x,y)|x\,Ry\wedge x\in A\}$.
e) *The* image *of A under R is* $R[A]:=R''A:=\{y|\exists x\in A\,x\,Ry\}$.
f) *The* preimage *of A under R is* $R^{-1}[A]:=\{x|\exists y\in A\,x\,Ry\}$.
g) *The* composition *of S and R ("S after R") is* $S\circ R:=\{(x,z)|\exists y\,(x\,Ry\wedge y\,Sz)\}$.
h) *The* inverse *of R is* $R^{-1}:=\{(y,x)|x\,Ry\}$.

Relations can play different roles in mathematics.

Definition 17. *Let R be a relation.*

a) *R is* reflexive *iff* $\forall x\in\mathrm{field}(R)\; x\,Rx$.
b) *R is* irreflexive *iff* $\forall x\in\mathrm{field}(R)\;\neg x\,Rx$.
c) *R is* symmetric *iff* $\forall x,y\,(x\,Ry\rightarrow y\,Rx)$.
d) *R is* antisymmetric *iff* $\forall x,y\,(x\,Ry\wedge y\,Rx\rightarrow x=y)$.
e) *R is* transitive *iff* $\forall x,y,z\,(x\,Ry\wedge y\,Rz\rightarrow x\,Rz)$.
f) *R is* connex *iff* $\forall x,y\in\mathrm{field}(R)\,(x\,Ry\vee y\,Rx\vee x=y)$.
g) *R is an* equivalence relation *iff R is reflexive, symmetric and transitive.*
h) *Let R be an equivalence relation. Then $[x]_R:=\{y|y\,Rx\}$ 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 undesired results. Another important family of relations is given by *order relations*.

Definition 18. *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* $\mathrm{field}(R)=A$.
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

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

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

Definition 19. *Let F be a term. Then F is a* function *if it is a relation which satisfies*

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

If F is a function then

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

is the value *of F at x.*

If F is a function and $x F y$ then $y = F(x)$. If there is no y such that $x F y$ 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

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

We define further notions associated with functions.

Definition 20. *Let F, A, B be terms.*

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

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

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

d) *F is an* injective *function from A to B iff $F: A \rightarrow B$ 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: $A \leftrightarrow B$, iff F: $A \rightarrow B$ is surjective and injective.*

f) ${}^{A}B := \{f | f: A \rightarrow B\}$ *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

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

One says 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 $F: A \rightarrow B$ and $F': A \rightarrow B$. Show that

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

4.7 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 21. *The union axiom is equivalent to $\forall x \bigcup x \in V$ or simply $\bigcup x \in V$.*

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. □

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 22. $\forall x_0, \ldots, x_{n-1} \{x_0, \ldots, x_{n-1}\} \in V$.

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 $x_0, \ldots, x_n$. Then

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

The right-hand side exists in V by the inductive hypothesis and the union axiom. □

4.8 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 $\varphi(z, x_1, \ldots, x_n)$ postulate:

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

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

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

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 an important consequence of the axioms:

Theorem 23. *$V \notin V$.*

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. □

This simple but crucial theorem leads to the distinction:

Definition 24. *Let A be a term. Then A is a* set *iff $A \in V$, and A is a* proper class *iff $A \notin V$.*

Set theory deals with sets and proper classes. Sets are the primary objects of set theory, the axiom mainly postulate 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.

4.9 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.

By the specific implementation of KURATOWSKI ordered pairs:

Lemma 25. $A\times B\subseteq\mathcal{P}(\mathcal{P}(A\cup B))$.

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

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

□

Theorem 26.

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,

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

□

Note that to "find" the sets in this theorem one has to apply the power set operation repeatedly. We shall later 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 27. *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 $f{:}\,x\to\mathcal{P}(x)$ by $u\mapsto\{u\}$. This is a set since

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

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 $g{:}\,\mathcal{P}(x)\to x$. Define the CANTORean set

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

similar to the class R in RUSSELL's paradox.

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. $\square$

This expresses that $\mathcal{P}(x)$ is "strictly larger" than x.

4.10 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:

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

Lemma 28. *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 $F: x \rightarrow x$ 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. $\square$

4.11 Infinity

All the axioms so far can be realized in a domain of finite sets. 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* basically expresses that the set of "natural numbers" exists. To this end, some "number-theoretic" notions are defined.

Definition 29.

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 30. *Define*

a) $1 := 0 + 1$;

b) $2 := 1 + 1$;

c) $3 := 2 + 1$; ...

From the context it will always 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 the important number-theoretic properties of “three”.

4.12 Foundation

The *axiom schema of foundation* provides structural information about the set theoretic universe V. It postulates for any term A:

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

Viewing $\in$ as a 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 31. *Let n be a natural number $\geqslant 1$. Then there are* no $x_0, \ldots, x_{n-1}$ *such that*

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

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

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

$A \neq \emptyset$ 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. □

Exercise 7. Show that $x \neq x + 1$.

Exercise 8. Show that the successor function $x \mapsto x + 1$ is injective.

Exercise 9. Show that the term $\{x, \{x, y\}\}$ may be taken as an ordered pair of x and y.

Theorem 32. *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. □

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

Definition 33. *A relation R on a domain D is called* wellfounded*, iff for all terms A*

$$\emptyset \neq A \wedge A \subseteq D \rightarrow \exists x \in A \, A \cap \{y \mid y R x\} = \emptyset .$$

Exercise 10. Formulate and prove a principle for R-induction on D which coressponds to the assumption that R is wellfounded on D.