Chapter 3
Sets: finite and infinite

H.C.M. (Harrie) de Swart

Abstract Sets occur abundantly in mathematics and in daily life. But what is a set? Cantor (1845-1918) defined a set as a collection of all objects which have a certain property in common. Russell showed in 1902 that this assumption yields a contradiction, known as Russell's paradox, and hence is untenable. In 1908 Zermelo (1871-1953) weakened Cantor's postulate considerably and consequently had to add a number of additional axioms. We present the set theory of Zermelo-Fraenkel. Next we discuss relations and functions. We use the Hilbert hotel with as many rooms as there are natural numbers to illustrate a number of astonishing properties of sets which are equally large as the set $\mathbb{N}$ of the natural numbers. We shall discover that there are many sets which in a very precise sense are much larger than $\mathbb{N}$. We shall even see that for any set V, finite or infinite, there is a larger set $P(V)$, called the powerset of V. Amazingly, although all sets we experience in the world are finite, we are still able to imagine infinite sets like $\mathbb{N}$ and to see amazing properties of them. This reminds us of the statement by cardinal Cusanus (1400-1453) that in our pursuit of grasping the divine truths we may expect the strongest support of mathematics. Finally we point out that Kant was right that mathematical (true) propositions are not analytic, but synthetic, and that Russell and Frege's logicism, stating that all of mathematics may be reduced to logic, is wrong. What may be true is that mathematics can be reduced to logic plus set theory.

3.1 Russell's Paradox

We all know lots of sets. Here are a few examples: the set of all citizens of the Netherlands, the set of all players in a soccer team, the set of all triangles in a plane.

Another example is the set of the natural numbers 1, 2 and 3. This set is denoted by $\{1, 2, 3\}$. Then $3 \in \{1,2,3\}$ denotes: 3 is an element of the set $\{1, 2, 3\}$; and $7 \notin \{1,2,3\}$ denotes: $\neg(7 \in \{1,2,3\})$, i.e., 7 is not an element of the set $\{1, 2, 3\}$.

H. de Swart, *Philosophical and Mathematical Logic*, Springer Undergraduate Texts in Philosophy, https://doi.org/10.1007/978-3-030-03255-5_3

The numbers 0, 1, 2, 3, ... are called *natural numbers*. We may consider the infinite set of all natural numbers. This set is denoted by $\mathbb{N}$, in other words $\mathbb{N} = \{0,1,2,\ldots\}$. For example, $3 \in \mathbb{N}$ and $1024 \in \mathbb{N}$, but $-3 \notin \mathbb{N}$, $\frac{2}{3} \notin \mathbb{N}$ and $\sqrt{2} \notin \mathbb{N}$.

It turns out that many, if not all, notions from mathematics can be represented by sets. For instance, we shall see that the natural numbers $0,1,2,\ldots$ may be represented by sets. That means that set theory may be conceived as a foundation of mathematics, as a unifying theory in which all mathematics may be represented. So, from now on we shall assume that sets are our universe of discourse.

Cantor's naive comprehension principle But what is a set? G. Cantor (1845 - 1918) answered this question as follows: a set is by definition the collection of all objects which have a certain property A. This principle is now known as the *naive comprehension principle*: Let $A(x)$ express that (set) x has the property A. Then $\{x \mid A(x)\}$ is the set of all (sets) x which have the property A, i.e.,

$$\text{for all (sets) } y,\ y \in \{x \mid A(x)\} \text{ iff } A(y).$$

For instance, let $A(x)$ stand for: x is a natural number. Then Cantor's naive comprehension principle tells us that $\{x \mid x \text{ is a natural number}\}$ is a set, which we may denote by $\mathbb{N}$.

However, in 1902 Bertrand Russell showed in a letter to Frege (see Heijenoort [6], p. 124) that the naive comprehension principle leads to a contradiction. The argument is extremely simple: apply the naive comprehension principle to the property $A(x)$: $x \notin x$. According to Cantor's principle, $\{x \mid x \notin x\}$ is a set V such that for all (sets) y, $y \in V$ iff $y \notin y$. In particular, taking for y the set V itself we get

$$V \in V \text{ iff } V \notin V.$$

Contradiction.

The argument above is known as *Russell's paradox*. Russell's argument shows that set theory with the naive comprehension principle is *inconsistent*. This was quite a shock to the community at the time, because set theory was (and still is) considered to be a foundation for all of mathematics.

One way to escape the paradox was indicated by Zermelo on the grounds of the following observation: the set involved in the derivation of the paradox turns out to be very large – the set of all sets not being an element of themselves. Zermelo noted that the full force of the naive comprehension principle was hardly ever used; one mostly uses it to create subsets of a given set. So, instead of the naive comprehension principle Zermelo put forward his *Aussonderungs Axiom* or separation axiom:

Separation Axiom: if V is a set and $A(x)$ a property, then also $\{x \in V \mid A(x)\}$ is a set, consisting of all elements **in** V which have the property A, i.e., such that for all (sets) y:

$$y \in \{x \in V \mid A(x)\} \text{ iff } y \in V \text{ and } A(y)$$

The separation axiom says that within a given set V we can collect all elements of V, which have a certain property A, into a subset $\{x \in V \mid A(x)\}$ of V. Cantor allowed this principle not only for a given set V, but also for the universe of all sets. And Russell showed that to be contradictory.

If we abandon the naive comprehension principle and adopt the separation axiom instead, we can no longer accept the proof of Russell's paradox. However, we may use the idea of Russell's proof to obtain, with the help of the separation axiom, a positive result. From the separation axiom it follows:

Theorem 3.1. *For any set V there is a set W, namely $W = \{x \in V \mid x \notin x\}$, such that $W \notin V$.*

Proof. Let V be a given set. According to the separation axiom, $W = \{x \in V \mid x \notin x\}$ is a set such that for all sets y, $y \in W$ iff $y \in V$ and $y \notin y$. In particular, since W itself is a set, we get

$$W \in W \text{ iff } W \in V \text{ and } W \notin W.$$

Now suppose $W \in V$; then $W \in W$ iff $W \notin W$. Contradiction. Therefore, $W \notin V$.

Making use of truth-tables (see Chapter 2) one may illustrate this proof as follows. The propositions $W \in W$ and $W \in V$ can be either true (1) or false (0), giving four possible combinations:

$W \in W$	$W \in V$	$W \notin W$	$W \in V \wedge W \notin W$	$W \in W \rightleftarrows W \in V \wedge W \notin W$
1	1	0	0	0
1	0	0	0	0
0	1	1	1	0
0	0	1	0	1

From the Separation Axiom it follows that $W \in W \rightleftarrows W \in V \wedge W \notin W$ is a true (1) proposition. Hence, we are in the 4^{th} line of the truth table. And we can read off from that line that both $W \in W$ and $W \in V$ are false (0). In particular, $W \notin V$. □

From the Separation Axiom it follows that no set may contain all sets, in other words, the universe (or totality) of all sets is not a set.

Corollary 3.1. *The universe (or totality) of all sets is not a set.*

Proof. Suppose the universe of all sets were a set U. Then by definition of U, for all sets W, $W \in U$ (1). But if U were a set, it follows from Theorem 3.1 that there is a set W, namely $W = \{x \in U \mid x \notin x\}$, such that $W \notin U$ (2).
(1) and (2) are contradictory. Hence, the universe of all sets is not a set. □

Russell obtained his paradox from the naive comprehension principle by considering the 'set' $\{x \mid x \notin x\}$. By considering the set $\{x \in V \mid x \notin x\}$, given any set V, we did not obtain a paradox, but the positive and interesting results formulated in Theorem 3.1 and Corollary 3.1 instead.

Another way to escape Russell's paradox is to blame the contradiction on the expression $x \notin x$: $x \notin x$ produced a contradiction, so we must suppress $x \in x$. Russell, in his *theory of types*, has chosen this approach: assign type to variables (sets) and allow expressions such as $x \in y$ only if the type of x is one less than the type of y. So, the expression $x \in x$ is then grammatically not correct.

Since the separation axiom yields only new sets, given any set V in advance, we have to postulate the existence of at least one set, in order to be able to build other sets. E. Zermelo (1871-1953) laid down his system of axioms for sets in 1908. The extension of Fraenkel dates from 1922. Below we present the axioms ZF of Zermelo and Fraenkel. The axioms may be formulated in natural language, but they may also be formulated in the language of predicate logic, letting the variables range over sets and using only two binary predicate symbols: $\in$ (is element of) and $=$ (is equal to).

3.2 Axioms of Zermelo-Fraenkel for Sets

Empty set axiom: There exists a set without elements. In other words, there is a set x such that for all sets y, $y \notin x$.
Formulated in the predicate language just mentioned: $\exists x \forall y[\neg(y \in x)]$

There are many examples of empty sets in daily life: the set of living persons older than 150 years; the set of all persons with blue hair, the set of all natural numbers which are both even and odd, etc. Notice that the existence of the empty set also would follow from the naive comprehension principle: $\{x \mid x \neq x\}$, assuming that each thing is equal to itself.

Sets are, just like triangles and numbers, legitimate mathematical objects. So it makes perfectly good sense to ask whether two sets are identical or not. If two sets x and y are identical (equal), we write $x = y$, if not, $x \neq y$. Identical sets have exactly the same properties; so, if $x = y$, then every element of x is also an element of y and vice versa. One may wonder if, conversely, sets with exactly the same elements are identical. Consider, for example, the set V of all even numbers greater than zero and the set W of all sums of pairs of odd numbers. There is some reason to distinguish V and W: they are given in different ways. On the other hand, we feel (and mathematical practice confirms this) that definitions do not matter so much, it is rather content that counts. So, we make the explicit choice to consider sets as merely being determined by their elements. Hence, 'having the same elements' means 'being equal'.

Axiom of extensionality: Two sets are equal if and only if they have the same elements. As observed above, the 'only if' holds trivially.
Formulated in our predicate language: $x = y \rightleftarrows \forall z[z \in x \rightleftarrows z \in y]$.

The axiom of extensionality has among others the following consequences:

$$\begin{array}{rr} \{3,4,5\} = \{4,3,5\} & \{2,3\} \neq \{3,4\} \\ \{3,3,7\} = \{3,7\} & \{0,1\} \neq \{1,2\} \\ \{2,3\} = \{2,3,3\} & \{2,\{3,4\}\} \neq \{\{2,3\},4\} \end{array}$$

Notice that the only elements of $\{2,\{3,4\}\}$ are: 2 and $\{3,4\}$, while the only elements of $\{\{2,3\},4\}$ are: $\{2,3\}$ and 4. For instance, $2 \in \{2,\{3,4\}\}$, but $2 \notin \{\{2,3\},4\}$; and $\{2,3\} \in \{\{2,3\},4\}$, but $\{2,3\} \notin \{2,\{3,4\}\}$.

Since, by the extensionality axiom, a set is completely determined by its elements, there may be at most one empty set: if there were two sets without elements, they would have the same elements ($0 \rightleftarrows 0 = 1$) and hence, by the axiom of extensionality, be equal. The empty set axiom says that there is at least one empty set. By the axiom of extensionality there is at most one empty set. Hence, there is exactly one empty set. **Notation**: $\emptyset$.
By definition: $\forall y[y \notin \emptyset]$.

Given two sets V and W, we want to be able to construct a set whose elements are exactly V and W themselves. The existence of such a set would also follow from the naive comprehension principle: $\{x \mid x = V \text{ or } x = W\}$. So, we postulate:

Pairing Axiom: Given any sets v and w, there exists a set y, whose elements are exactly v and w.
Formulated in our predicate language: $\forall v \forall w \exists y \forall z[z \in y \rightleftarrows z = v \vee z = w]$.

Again, by the extensionality axiom, given sets v and w, the set whose existence is required by the pairing axiom is unique and is called the *unordered pair* $\{v,w\}$ of v and w. Because $\{v,w\}$ and $\{w,v\}$ have the same elements, they are equal.
So, for all (sets) z, $z \in \{v,w\}$ iff $z = v$ or $z = w$.

$\{v\} := \{v,v\}$ is the *singleton* of v. If v is a set, then so is $\{v\}$, because of the pairing axiom and the definition of $\{v\}$.

Now, with only a few axioms, the existence of infinitely many sets follows:

$$\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \ldots$$

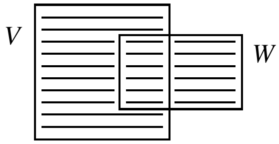
$\emptyset$ (we repeat) is a set without elements. $\{\emptyset\}$, on the other hand, is a set with one element, namely $\emptyset$. Hence, $\emptyset \neq \{\emptyset\}$.
$\{\{\emptyset\}\}$ is the set with $\{\emptyset\}$ as its only element, while $\{\emptyset\}$ has $\emptyset$ as its only element. Hence, $\{\{\emptyset\}\} \neq \{\emptyset\}$, because $\emptyset \notin \{\{\emptyset\}\}$.
The Pairing Axiom also entails the existence of $\{\emptyset,\{\emptyset\}\}$, which is the set with $\emptyset$ and $\{\emptyset\}$ as its only elements.

Given two sets V and W we want to be able to construct the *union* $V \cup W$ of V and W such that for all z, $z \in V \cup W$ iff $z \in V \vee z \in W$. Its existence would follow from the naive comprehension principle: $\{x \mid x \in V \text{ or } x \in W\}$. Notice that in general, $V \cup W$ is a larger set than each of V and W separately.

Union axiom If v and w are sets, then there exists a set y such that for all (sets) z, $z \in y$ iff $z \in v$ or $z \in w$.
Formulated in our predicate language: $\forall v \forall w \exists y \forall z[z \in y \rightleftarrows z \in v \vee z \in w]$

Again, by the extensionality axiom, given sets V and W, the set required by the union axiom is unique and is called the *union* of V and W. **Notation**: $V \cup W$.
So, for all (sets) z,

$$z \in V \cup W \rightleftarrows z \in V \vee z \in W.$$



Example 3.1.
$$\{1,2\} \cup \{5,6\} = \{1,2,5,6\},\ \{1,2\} \cup \{2\} = \{1,2\},$$
$$\{1,2\} \cup \{2,6\} = \{1,2,6\}, \quad \{1,2\} \cup \emptyset \ = \{1,2\}.$$
$$\{1,2\} \cup \{1,2\} = \{1,2\}.$$

The union axiom allows us to construct the union of any two given sets v and w or, put differently, to form the union of all elements of the set $x = \{v,w\}$. A more general version of the union axiom, put forward by Zermelo, was the following.

Sumset Axiom: For every set x there exists a set y, whose elements are exactly the objects occurring in at least one element of x.
Formulated in our predicate language: $\forall x \exists y \forall z[z \in y \rightleftarrows \exists v[v \in x \wedge z \in v]]$.

Again, the extensionality axiom guarantees the uniqueness of the set y, given x. This unique set is called the *sum-set* of x. **Notation**: $\bigcup x$ or $\cup\{y \mid y \in x\}$.
Notice that $v \cup w = \bigcup\{v,w\}$.

Now we are able to define the natural numbers in terms of sets as follows.

Definition 3.1 (Successor function). $0 := \emptyset$.
The *successor function* S is defined by $S(n) = n \cup \{n\}$, also denoted by $n+1$.

Example 3.2. $0 := \emptyset$
$1 := 0 \cup \{0\}$. So, $1 = \{0\} = \{\emptyset\}$.
$2 := 1 \cup \{1\}$. So, $2 = \{0\} \cup \{1\} = \{0,1\} = \{\emptyset, \{\emptyset\}\}$.
$3 := 2 \cup \{2\}$. So, $3 = \{0,1\} \cup \{2\} = \{0,1,2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$.
In general, for any natural number n, $n+1 := n \cup \{n\}$.

One easily checks by induction that for any natural n, defined in this way, $n = \{0, \ldots, n-1\}$ and that the sets 0, 1, 2, 3, ... are distinct pairwise. So, we have identified each natural number n with a certain standard set consisting of n elements. This definition of natural numbers in terms of sets justifies the use of natural numbers in the examples at the beginning of this section.

With very few axioms we have generated up till now infinitely many sets, but all of them are finite. But we also want to be able to deal with the infinite set of all natural numbers, which is so important in mathematics and its many applications. The existence of this set would follow easily from the naive comprehension principle: $\{x \mid x \text{ is a natural number}\}$. Since this naive comprehension principle had to be replaced by the much weaker separation axiom we have to postulate the existence of at least one infinite set.

Axiom of Infinity: There is at least one set y that contains 0, i.e., $\emptyset$, and is such that for every $x \in y$ it also contains Sx, i.e., $x \cup \{x\}$.
Formulated in our predicate language: $\exists y[0 \in y \wedge \forall x[x \in y \rightarrow Sx \in y]]$

The set y whose existence is required by the axiom of infinity has clearly infinitely many members: 0, 1, 2, 3, ... But there might be many of such sets containing in addition other things. So, we take the smallest such set which contains 0 and with every number n its successor $Sn = n+1$ and denote it by $\mathbb{N}$. So, $0 \in \mathbb{N}$, $1 \in \mathbb{N}$, $2 \in \mathbb{N}$, etc. Notice that $\mathbb{N}$ has infinitely many members, but $\{\mathbb{N}\}$ has only one element: $\mathbb{N}$.

In order to be able to construct for instance the set of all even natural numbers, i.e., $\mathbb{N}_{even} = \{n \in \mathbb{N} \mid n \text{ is even}\}$, we need the separation axiom.

Separation Axiom: If x is a set and $A(z)$ a property, then also $\{z \in x \mid A(z)\}$ is a set, consisting of all elements **in** x which have the property A, i.e., such that for all z:

$$z \in \{z \in x \mid A(z)\} \text{ iff } z \in x \text{ and } A(z)$$

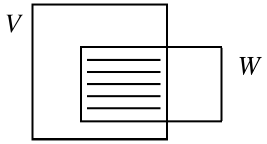
Formulated in our logical predicate language: $\forall x \exists y \forall z[z \in y \rightleftarrows z \in x \wedge A(z)]$ for any formula A in our logical predicate language.

The separation axiom says that within a given set x we can collect all elements of x, which have a given property A, into a subset $\{z \in x \mid A(z)\}$ of x. Notice that the separation axiom is in fact an axiom schema: it yields an axiom for any formula A. By the axiom of extensionality, given a set x and a property A, the set y, whose existence is demanded by the separation axiom, is uniquely determined and shall be denoted by $\{z \in x \mid A(z)\}$.

Given the separation axiom and the axiom of infinity, the existence of the empty set follows immediately: $\emptyset = \{z \in \mathbb{N} \mid z \neq z\}$, if we assume that for all z, $z = z$. Also, given the separation axiom, we may introduce some important set theoretical operations: *intersection* and *relative complement*.

Corollary 3.2 (Intersection). *Given any sets V and W, also the intersection $V \cap W := \{z \in V \mid z \in W\}$ of V and W is a set, such that for all z*

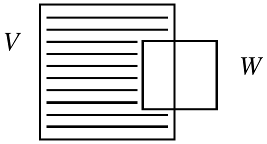
$$z \in V \cap W \rightleftarrows z \in V \wedge z \in W.$$



We may generalize the intersection as follows. If x is a non-empty set, say $v \in x$, then $\bigcap x := \{z \in v \mid \forall y[y \in x \rightarrow z \in y\}$. Notice that $V \cap W = \bigcap\{V, W\}$.

Corollary 3.3 (relative complement). *Given any sets V and W, also the relative complement, $V - W := \{z \in V \mid z \notin W\}$ of W with respect to V, is a set, such that*

$$z \in V - W \rightleftarrows z \in V \wedge z \notin W.$$



Notice that $V \cap W$ and $V - W$ are in general smaller sets than V, while $V \cup W$ in general is a larger set than V. The existence of $V \cap W$ and $V - W$ follows from the separation axiom, while the existence of $V \cup W$ requires the union axiom.

Example 3.3.

$\{1,2\} \cup \{2,3\} = \{1,2,3\}$	$\{1,2\} \cup \emptyset = \{1,2\}$	$\{1,2\} \cup \mathbb{N} = \mathbb{N}$
$\{1,2,3\} \cap \{2,3,4\} = \{2,3\}$	$\{1,2\} \cap \emptyset = \emptyset$	$\{2,3\} \cap \mathbb{N} = \{2,3\}$
$\{1,2,3\} - \{2,3,4\} = \{1\}$	$\{1,2,3\} - \emptyset = \{1,2,3\}$	$\{1,2,3\} - \mathbb{N} = \emptyset$

The reader may easily verify the following statements:
1. $\cap$ and $\cup$ are *idempotent*, i.e., $V \cap V = V$, respectively $V \cup V = V$, for any set V.
2. $\cap$ and $\cup$ are *commutative*, i.e., $V \cap W = W \cap V$, respectively $V \cup W = W \cup V$, for any sets V and W.
3. $\cap$ and $\cup$ are *associative*, i.e., $U \cap (V \cap W) = (U \cap V) \cap W$, respectively $U \cup (V \cup W) = (U \cup V) \cup W$, for any sets U, V, W.
4. $V \cap \emptyset = \emptyset$ and $V \cup \emptyset = V$ for any set V.

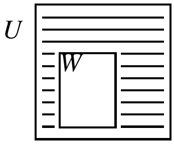
Theorem 3.2 (absorption laws). *For all sets V and W,*
$V \cap (V \cup W) = V$ and $V \cup (V \cap W) = V$.

Proof. By the axiom of extensionality we have to show that the two sets in question have the same elements, i.e., for all z, $z \in V \cap (V \cup W)$ iff $z \in V$ and $z \in V \cup (V \cap W)$ iff $z \in V$. This is straightforward. □

Theorem 3.3 (distributive laws). *For all sets U, V and W,*
$U \cap (V \cup W) = (U \cap V) \cup (U \cap W)$ and $U \cup (V \cap W) = (U \cup V) \cap (U \cup W)$.

Proof. By the axiom of extensionality we have to show that for all z, $z \in U \cap (V \cup W)$ iff $z \in (U \cap V) \cup (U \cap W)$, in other words, $z \in U \wedge (z \in V \vee z \in W)$ iff $(z \in U \wedge z \in V) \vee (z \in U \wedge z \in W)$. This is straightforward and also follows from the distributive laws of propositional logic in Theorem 2.10. □

When it is clear from the context that the complement of a set W is taken relative to a given universe U, $U - W$ is simply called the *complement* of W and denoted by W^c.

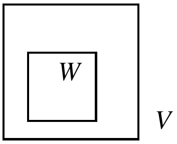


Theorem 3.4. *Let V^c and W^c be the complement of V, respectively W, relative to a given universe U. $(V \cup W)^c = V^c \cap W^c$ and $(V \cap W)^c = V^c \cup W^c$.*

Proof. We leave the proof to the reader as Exercise 3.3. □

In order to be able to formulate the powerset axiom we first have to introduce the notion of *subset*.

Definition 3.2 (Subset). W is a *subset* of V := every element of W is also an element of V, i.e., for every x, if $x \in W$, then also $x \in V$. **Notation**: $W \subseteq V$.



Notice that W is not a subset of V iff not all elements of W are elements of V, in other words, iff there is some $x \in W$ such that $x \notin V$. **Notation**: $\neg(W \subseteq V)$ or $W \not\subseteq V$.

Example 3.4.

$\{2,3\} \subseteq \{1,2,3,4\}$	$\{2,3\} \subseteq \{2,3\}$	$\emptyset \subseteq \{2,3\}$	$\{2,3\} \subseteq \mathbb{N}$
$\{2,3\} \not\subseteq \{3,4,5\}$	$\{1,\{2\}\} \not\subseteq \{1,2\}$	$\{1,2\} \not\subseteq \{1,\{2\}\}$	$\mathbb{N} \not\subseteq \{\mathbb{N}\}$

Definition 3.3 (Proper subset). W is a *proper subset* of V := $W \subseteq V$ and not $W = V$. **Notation**: $W \subset V$.

Example 3.5. $\{2,3\} \subset \{2,3,4\}$ and $\{2,3\} \subset \mathbb{N}$.

Warning: It is important not to confuse $\in$ and $\subseteq$:
$\{2\} \in \{\{2\},3\}$, but $\{2\} \not\subseteq \{\{2\},3\}$, the latter because $2 \in \{2\}$, but $2 \notin \{\{2\},3\}$.
$\{2,3\} \subseteq \{1,2,3\}$, but $\{2,3\} \notin \{1,2,3\}$.

Theorem 3.5. *For any set V, $\emptyset \subseteq V$ and $V \subseteq V$.*

Proof. Suppose that for some V, $\emptyset \not\subseteq V$, i.e., there would be an element $x \in \emptyset$ such that $x \notin V$. Because $\emptyset$ has no elements, this is impossible. Therefore, $\emptyset \subseteq V$. And because every element of V is an element of V, it follows that $V \subseteq V$. □

Example 3.6. $\emptyset \subseteq \emptyset$, but $\emptyset \notin \emptyset$.
$\emptyset \subseteq \{\emptyset\}$, and by definition of $\{\emptyset\}$ also $\emptyset \in \{\emptyset\}$.
$\emptyset \subseteq \{\{\emptyset\}\}$, but $\emptyset \notin \{\{\emptyset\}\}$, since the only element of $\{\{\emptyset\}\}$ is $\{\emptyset\}$.
$\{\emptyset\} \subseteq \{\emptyset\}$, but $\{\emptyset\} \notin \{\emptyset\}$, since the only element of $\{\emptyset\}$ is $\emptyset$.
$\{\emptyset\} \not\subseteq \{\{\emptyset\}\}$, because $\emptyset \in \{\emptyset\}$ while $\emptyset \notin \{\{\emptyset\}\}$, but $\{\emptyset\} \in \{\{\emptyset\}\}$.

Next we will determine for a few small finite sets all their subsets and the set of all their subsets. Let us start with $\emptyset$. The only subset of $\emptyset$ is $\emptyset$ itself. So, the set $P(\emptyset)$ of all subsets of $\emptyset$ is $\{\emptyset\}$.

The only subsets of the set $\{u\}$ are $\emptyset$ with zero elements and $\{u\}$ itself with one element. So, the set $P(\{u\})$ of all subsets of $\{u\}$ is $\{\emptyset,\{u\}\}$.

The subsets of $\{u,v\}$ can have 0, 1 or 2 elements and are, respectively, $\emptyset$ with zero elements, $\{u\}$ and $\{v\}$ with one element, and $\{u,v\}$ itself with two elements. So, the set $P(\{u,v\})$ of all subsets of $\{u,v\}$ is $\{\emptyset,\{u\},\{v\},\{u,v\}\}$. Notice that there are twice as many subsets of $\{u,v\}$ as there are subsets of $\{u\}$: all subsets of $\{u\}$, i.e., $\emptyset$ and $\{u\}$, are also a subset of $\{u,v\}$ and the other subsets of $\{u,v\}$ are obtained by adding the element v to the subsets of $\{u\}$.

The subsets of $\{u,v,w\}$ can have 0, 1, 2 or 3 elements and are, respectively, $\emptyset$ with zero elements, $\{u\}$, $\{v\}$ and $\{w\}$ with one element, $\{u,v\}$, $\{u,w\}$ and $\{v,w\}$ with two elements, and finally $\{u,v,w\}$ itself with three elements. So, the set $P(\{u,v,w\})$ of all subsets of $\{u,v,w\}$ is $\{\emptyset,\{u\},\{v\},\{w\},\{u,v\},\{u,w\},\{v,w\},\{u,v,w\}\}$. Notice that there are twice as many subsets of $\{u,v,w\}$ as there are subsets of $\{u,v\}$: all subsets of $\{u,v\}$, i.e., $\emptyset$, $\{u\}$, $\{v\}$ and $\{u,v\}$, are also a subset of $\{u,v,w\}$ and the other subsets of $\{u,v,w\}$ are obtained by adding the element w to the subsets of $\{u,v\}$.

This brings us to the following observation: each time that one adds one element w to a given finite set V, one obtains twice as many subsets: all the subsets of V plus all subsets of V with the new element w added. From this insight results the following theorem:

Theorem 3.6. *For each natural number n, if V is a finite set with n elements, then V has 2^n subsets.*

Proof. By mathematical induction. For $n=0$: a set V with 0 elements is the empty set $\emptyset$, and this set has $2^0=1$ subset, namely $\emptyset$. Suppose the statement is true for $n=k$, i.e. any set with k elements has 2^k subsets (induction hypothesis). Then a set with $k+1$ elements has twice as many subsets, i.e., $2\cdot 2^k = 2^{k+1}$ subsets. □

For instance, if V has 10 elements, V has $2^{10}=1024$ subsets. And if V has 20 elements, V has $2^{20}=2^{10}\cdot 2^{10}=1024\cdot 1024$ subsets, that is more than one million!

Since sets of subsets occur abundantly in mathematics and since the existence of many of these sets does not follow from the set theoretic axioms introduced up till now, we postulate the following powerset axiom:

Powerset axiom: If V is a set, then also $P(V)=\{X \mid X\subseteq V\}$ is a set. We call $P(V)$ the *powerset* of V.

Formulated in our logical predicate language: $\forall v\exists y\forall x[x\in y \rightleftarrows x\subseteq v]$.

So, the elements of $P(V)$ are the subsets of V, i.e.,

$$X\in P(V) \text{ iff } X\subseteq V.$$

The name *powerset* refers to the fact that if V has n ($n\in\mathbb{N}$) elements, then by Theorem 3.6, $P(V)$ has 2^n elements.

This powerset axiom may look innocent, but is it? We have already seen that if V is a relatively small finite set, then $P(V)$ may become a relatively large set. And what will happen when we apply the P-operator to an infinite set, like $\mathbb{N}$? According to the powerset axiom, not only $P(\mathbb{N})$ is another set, but also $P(P(\mathbb{N}))$, $P(P(P(\mathbb{N})))$, etc. are new sets. As we shall see later on in Section 3.6, these sets become so large that one may ask the question whether we are still able to construct these sets. In fact, the powerset axiom is the only set theoretic axiom which is not by everyone accepted in its full strength, in particular not by the intuitionists; see Chapter 8.

Up till now we have postulated the following axioms for set theory: empty set axiom, axiom of extensionality, pairing axiom, union axiom, sumset axiom, axiom of infinity, separation axiom, and powerset axiom. The set theory *ZF* of Zermelo-Fraenkel contains two more axioms: the axiom of replacement, which is the only contribution of Fraenkel, and the axiom of regularity (or foundation). We only mention these axioms here and refer to exercise 3.8 and to van Dalen, Doets, de Swart [3].

Axiom of Replacement: If for every x in V there is exactly one y such that $\Phi(x,y)$, then there exists a set W which contains precisely the elements y for which there is an $x \in V$ with the property $\Phi(x,y)$. In other words, the image of a set V under an operation (functional property Φ) is again a set.

Axiom of Regularity: Every non-empty set is disjoint from at least one of its elements.

The latter axiom guarantees that for any set x, $x \notin x$ and that there is no sequence $v_1,\ldots,v_n$ of sets such that $v_1 \in v_2$, $v_2 \in v_3$, …, $v_{n-1} \in v_n$ and $v_n \in v_1$ (Exercise 3.8).

There are several set theoretical principles which are consistent with, but independent of the axioms of Zermelo-Fraenkel. The axioms of choice and the continuum hypothesis (see Section 3.6) are not treated here because of their more dubious status. See van Dalen, Doets, de Swart, [3] for an elaborate discussion.

Exercise 3.1. Which of the following propositions are true and which are false?

$\mathbb{N} \in \mathbb{N}$
$\mathbb{N} \in \{\mathbb{N}\}$
$\mathbb{N} \subseteq \mathbb{N}$
$\mathbb{N} \in \{\{\mathbb{N}\}\}$
$\mathbb{N} \subseteq \{\mathbb{N}\}$
$\{1,2\} \in \mathbb{N}$
$\{1,2\} \subseteq \mathbb{N}$
$\{1,2\} \in \{\mathbb{N}\}$

$\{2,3\} \subseteq \{\mathbb{N}\}$
$\{2\} \subseteq \{\mathbb{N}\}$
$\{2\} \subseteq \mathbb{N}$
$2 \in \{1,\{2\},3\}$
$\{2\} \in \{1,\{2\},3\}$
$\{1,\{2\}\} \subseteq \{1,\{2,3\}\}$
$\{1,\{2\}\} \subseteq \{1,\{2\},3\}$
$\{-2,2\} \subseteq \mathbb{N}$

$\emptyset \in \emptyset$
$\emptyset \subseteq \emptyset$
$\emptyset \in \{\emptyset\}$
$\emptyset \subseteq \{\emptyset\}$
$\emptyset \in \{\{\emptyset\}\}$
$\emptyset \subseteq \{\{\emptyset\}\}$
$\{\emptyset\} \in \{\{\emptyset\}\}$
$\{\emptyset\} \subseteq \{\{\emptyset\}\}$

$\{\emptyset\} \in \emptyset$
$\{\emptyset\} \subseteq \emptyset$
$\{\emptyset\} \subseteq \{\emptyset\}$
$\emptyset \subseteq \{\emptyset,\{\emptyset\}\}$
$\emptyset \in \{\emptyset,\{\emptyset\}\}$
$\{\emptyset\} \subseteq \{\emptyset,\{\emptyset\}\}$
$\{\emptyset\} \in \{\emptyset,\{\emptyset\}\}$
$\emptyset \subseteq \{\{\emptyset,\{\emptyset\}\}\}$

Exercise 3.2. Prove or refute: a) $W \subseteq V$ iff $V \cap W = W$; b) $W \subseteq V$ iff $V \cup W = V$.

Exercise 3.3. Prove or refute: for all sets U, V and W,
a) $U - (V \cup W) = (U - V) \cap (U - W)$; b) $U - (V \cap W) = (U - V) \cup (U - W)$.

Exercise 3.4. Prove or refute: for all sets U, V and W,
a) if $U \in V$ and $V \in W$, then $U \in W$; b) if $U \subseteq V$ and $V \subseteq W$, then $U \subseteq W$.

Exercise 3.5. Determine $P(\emptyset)$, $P(P(\emptyset))$ and $P(P(P(\emptyset)))$.

Exercise 3.6. Prove:
(a) If $W \subseteq V$, then $P(W) \subseteq P(V)$; (b) If $P(W) \subseteq P(V)$, then $W \subseteq V$.
(c) If $P(W) = P(V)$, then $W = V$; (d) If $P(W) \in P(V)$, then $W \in V$.

Exercise 3.7. Prove or refute:
a) for all sets W, V, if $P(W) \in PP(V)$, then $W \in P(V)$.
b) for all sets W, V, if $W \in P(V)$, then $P(W) \in PP(V)$.
c) for all sets W, V, if $P(W) \subseteq PP(V)$, then $W \subseteq P(V)$.
d) for all sets W, V, if $W \subseteq P(V)$, then $P(W) \subseteq PP(V)$.

Exercise 3.8. Show that from the axiom of regularity it follows that i) for any set x, $x \notin x$, and ii) there is no sequence $v_1, \ldots, v_n$ such that $v_1 \in v_2$, $v_2 \in v_3$, …, $v_{n-1} \in v_n$ and $v_n \in v_1$.

3.3 Historical and Philosophical Remarks

3.3.1 Mathematics and Theology

In Corollary 3.1 we have seen that from the separation axiom it follows that the universe of all sets itself is not a set. This reminds us of Cardinal Cusanus (1400-1453), who in his *De docta ignorantia* [2] says that in the pursuit of grasping the divine truths we may expect the strongest support from mathematics. Although he illustrated this statement with other examples, it seems fair to say that he might have used Corollary 3.1 as an illustration: the universe of all earthly things (God?) is itself not an earthly thing.

Also the insights about infinite sets to be discovered in Sections 3.5 and 3.6 may be considered as illustrations of his statement. Although we never experience infinite sets in daily life, we are still able to imagine them and even to gain insights into their amazing properties.

3.3.2 Ontology of mathematics

Since the integers, the rational and the real numbers can be defined in terms of sets and natural numbers, it follows that these numbers can ultimately be defined in set-theoretical terms (see van Dalen, Doets, de Swart, [3]). Through practical experience mathematicians have found that most well-known concepts, such as the notion of number, function, triangle, and so on, can be defined in set-theoretical terms. This has led to the slogan ‘Everything is a set’, meaning that all objects from mathematical practice turn out to be representable in terms of sets. Consequently, every mathematical proposition can be reduced to a proposition about sets. It turns out that most, if not all, mathematical theorems – after translation in terms of sets – can be deduced logically from the axioms of set theory.

Set-theoretical Axioms

logical reasoning

mathematical theorems

So one might say that the axioms of *ZF* (Zermelo-Fraenkel) determine the *ontology of mathematics*: all mathematical objects are conceived as sets and the axioms of Zermelo-Fraenkel postulate the existence of certain sets, leaving room for extension with possibly more axioms, and they specify what the characteristic properties of these mathematical objects (sets) are. In this sense the axioms of *ZF* can be considered to be a foundation for (the greater part of) mathematics.

The axioms of Zermelo-Fraenkel (*ZF*) may be described informally. But we have also seen that the set theory of Zermelo-Faenkel may be *formalized* by:
1. first introducing the predicate language with only two binary predicate symbols $=$ and $\in$ with 'is equal to', respectively 'is element of' as intended interpretation, such that all statements about sets may be expressed in this language;
2. and next by specifying the axioms of *ZF* in this language, such that statements about sets (mathematical objects) may be logically deduced from these axioms.

3.3.3 Analytic-Synthetic

In his *Critique of Pure Reason* (1781) Immanuel Kant [7] makes a distinction between analytic and synthetic judgments. Kant calls a judgment *analytic* if its predicate is contained (though covertly) in the subject, in other words, the predicate adds nothing to the conception of the subject. Kant gives 'All bodies are extended' (Alle Körper sind ausgedehnt) as an example of an analytic judgment; I need not go beyond the conception of *body* in order to find extension connected with it. If a judgment is not analytic, Kant calls it *synthetic*; a synthetic judgment adds to our conception of the subject a predicate which was not contained in it, and which no analysis could ever have discovered therein. Kant mentions 'All bodies are heavy' (Alle Körper sind schwer) as an example of a synthetic judgment.

Also in his *Critique of Pure Reason* Kant makes a distinction between *a priori* knowledge and *a posteriori* knowledge. A priori knowledge is knowledge existing altogether independent of experience, while a posteriori knowledge is empirical knowledge, which has its sources in experience.

Sometimes one speaks of *logically necessary* truths instead of analytic truths and of *logically contingent* truths instead of synthetic truths, to be distinguished from physically necessary truths (truths which physically could not be otherwise, true in all physically possible worlds). The distinction between necessary and contingent truth is a *metaphysical* one, to be distinguished from the *epistemological* distinction between *a priori* and *a posteriori* truths.. Although these – the metaphysical and the epistemological – are certainly different distinctions, it was controversial whether they coincide in extension, that is, whether all and only necessary truths are *a priori* and all and only contingent truths are *a posteriori*.

In his *Critique of Pure Reason* Kant stresses that *mathematical judgments are both a priori and synthetic*. ‘Proper mathematical propositions are always *judgments a priori*, and not empirical, because they carry along with them the conception of necessity, which cannot be given by experience.’ Why are mathematical judgments synthetic? Kant considers the proposition $7+5=12$ as an example. ‘The conception of twelve is by no means obtained by merely cogitating the union of seven and five; and we may analyse our conception of such a possible sum as long as we will, still we shall never discover in it the notion of twelve.’ We must go beyond this conception of $7+5$ and have recourse to an intuition which corresponds to counting using our fingers: first take seven fingers, next five fingers extra, and then by starting to count right from the beginning we arrive at the number twelve.

7:	1	1	1	1	1	1	1					
5:								1	1	1	1	1
7 + 5:	1	1	1	1	1	1	1	1	1	1	1	1
	1	2	3	4	5	6	7	8	9	10	11	12

‘Arithmetical propositions are therefore always synthetic, of which we may become more clearly convinced by trying large numbers.’ Geometrical propositions are also synthetic. As an example Kant gives ‘A straight line between two points is the shortest’, and explains ‘For my conception of *straight* contains no notion of *quantity*, but is merely *qualitative*. The conception of the *shortest* is therefore wholly an addition, and by no analysis can it be extracted from our conception of a straight line.’

In more modern terminology, following roughly a ’Fregean’ account of analyticity, one would define a proposition A to be *analytic* iff either
(i) A is an instance of a logically valid formula; e.g., ’No unmarried man is married’ has the logical form $\neg\exists x[\neg P(x) \wedge P(x)]$, which is a valid formula, or
(ii) A is reducible to an instance of a logically valid formula by substitution of synonyms for synonyms; e.g., ’No bachelor is married’.

In his *Two dogmas of empiricism* W.V. Quine [8] is sceptical of the analytic-synthetic distinction. Quine argues as follows. In order to define the notion of analyticity we used the notion of synonymy in clause (ii) above. However, if one tries to explain this latter notion, one has to take recourse to other notions which directly or indirectly will have to be explained in terms of analyticity.

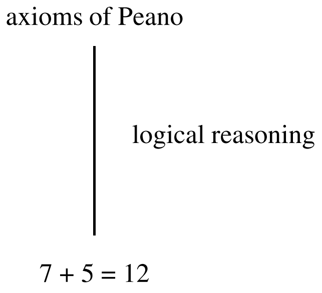
3.3.4 Logicism

Logicism dates from about 1900, its most important representatives being G. Frege in his *Grundgesetze der Arithmetik* I, II (1893, 1903) and B. Russell in his *Principia Mathematica* (1903), together with A.N. Whitehead. The program of the logicists was to reduce mathematics to logic. What do they mean by this? In his Grundgesetze der Arithmetik Frege defines the natural numbers in terms of sets as follows: 1 := the class of all sets having one element, 2 := the class of all sets having two elements, and so on. Next Frege shows that all kinds of properties of natural numbers can be logically deduced from a *naive comprehension principle*: if $A(x)$ is a property of an object x, then there exists a set $\{x \mid A(x)\}$ which contains precisely all objects x which have property A. (See Section 3.1.)

Logicism tried to introduce mathematical notions by means of explicit definitions; mathematical truths would then be logical consequences of these definitions. Mathematical propositions would then be reducible to logical propositions and hence mathematical truths would be analytic, contrary to what Kant said.

The greatest achievement of Logicism is that it succeeded in reducing great parts of mathematics to one single (formal) system, namely, set theory. The logicists believed that by doing this they reduced all of mathematics to logic without making use of any non-logical assumptions, hence showing that mathematical truths are analytic. However, what they actually did was reduce mathematics to logic PLUS set theory. And the axioms of set theory have a non-logical status! The axioms of set theory are – in Kant's terminology – synthetic, and surely not analytic. In his later years Frege came to realize that the axioms of set theory (see Section 3.2) are not a part of logic and gave up Logicism, which he had founded himself. The interested reader is referred to K. Gödel [4], *Russell's mathematical logic*.

Another way to see that a mathematical truth like $7+5=12$ is synthetic is to realize that $7+5=12$ is not a logically valid formula; it is true under the intended interpretation, but not true under all possible interpretations. $7+5=12$ can be logically deduced from the axioms of Peano for (formal) number theory (see Chapter 5), but it cannot be proved by the axioms and rules of formal logic alone.


axioms of Peano

logical reasoning

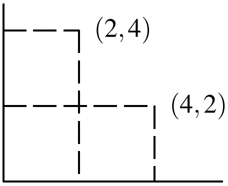
7 + 5 = 12

Again, Peano's axioms are true under the intended interpretation, but are not (logically) valid and hence they do not belong to logic.

3.4 Relations, Functions and Orderings*

3.4.1 Ordered pairs and Cartesian product

In the plane the pairs $(4,2)$ and $(2,4)$ indicate different points.



The order of the numbers 2 and 4 is of importance here, in the same way that the order of letters is of importance in constructing words: 'pin' and 'nip' contain the same letters, but in a different order. A pair of objects, say v and w, in which their order is relevant, is called the *ordered pair* of v and w, written (v,w). Sometimes the notation $< v,w >$ is used. This is different from the ordinary (unordered) pair $\{v,w\}$, which is the same as $\{w,v\}$. Ordered pairs have the characteristic property

$$(v,w) = (x,y) \text{ iff } v = x \text{ and } w = y. \qquad (*)$$

Unordered pairs do not have this property, since $\{v,w\} = \{w,v\}$ even for $v \neq w$.

We can introduce the notion of ordered pair as a primitive notion (i.e., undefined) and introduce the above-mentioned property (*) as an axiom. However, it is a wise rule not to introduce more primitive notions than necessary ('Ockham's razor') and hence we shall define a set, which behaves as an ordered pair, i.e., which satisfies the desired property $(*)$.

Definition 3.4 (Ordered pair). $(v,w) := \{\{v\},\{v,w\}\}$.

This is not the only definition which will work: see Exercise 3.9. We must now show that this definition satisfies (*).

Theorem 3.7. *$(v,w) = (x,y)$ iff $v = x$ and $w = y$.*

Proof. The implication from right to left is trivial. So suppose $(v,w) = (x,y)$, i.e., $\{\{v\},\{v,w\}\} = \{\{x\},\{x,y\}\}$. If two sets are equal, then they have the same elements. Hence, $\{v\} = \{x\}$ and $\{v,w\} = \{x,y\}$ or $\{v\} = \{x,y\}$ and $\{v,w\} = \{x\}$. In the first case it follows that $v = x$ and $w = y$. In the second case we can conclude: $v = x = y$ and $v = w = x$; so, also in this case, $v = x$ and $w = y$. □

The following theorem holds for Definition 3.4 of ordered pairs.

Theorem 3.8. *If $v \in V$ and $w \in W$, then $(v,w) \in PP(V \cup W)$.*

Proof. Suppose $v \in V$ and $w \in W$. Then:

(i) $v \in V \cup W$, so $\{v\} \subseteq V \cup W$, in other words, $\{v\} \in P(V \cup W)$, and

(ii) $w \in V \cup W$, so $\{v,w\} \subseteq V \cup W$, in other words, $\{v,w\} \in P(V \cup W)$.

From (i) and (ii) it follows that $\{\{v\},\{v,w\}\} \subseteq P(V \cup W)$, in other words, $\{\{v\},\{v,w\}\} \in PP(V \cup W)$. □

We can generalize the notion of ordered pair to the notion of ordered *n-tuple*:

Definition 3.5 (Ordered n-tuple). For $n \in \mathbb{N},\ n \geq 1$:

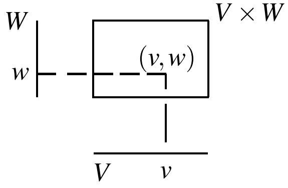
$$(v) := v,$$
$$(v_1, \ldots, v_n, v_{n+1}) := ((v_1, \ldots, v_n), v_{n+1}).$$

By means of mathematical induction one easily verifies that the object $(v_1, \ldots, v_n)$, ($n \in \mathbb{N},\ n \geq 1$), defined above, indeed behaves as an ordered n-tuple.

Theorem 3.9. *$(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ iff $x_1 = y_1$ and $\ldots$ and $x_n = y_n$.*

Proof. For $n = 1$, $(x_1) = x_1$ and $(y_1) = y_1$, so the proposition holds for $n = 1$. Now suppose (induction hypothesis) that the proposition holds for n, i.e., $(x_1, \ldots, x_n)$ = $(y_1, \ldots, y_n)$ iff $x_1 = y_1$ and $\ldots$ and $x_n = y_n$. Next suppose that $(x_1, \ldots, x_n, x_{n+1}) = (y_1, \ldots, y_n, y_{n+1})$, i.e., $((x_1, \ldots, x_n),\ x_{n+1}) = ((y_1, \ldots, y_n),\ y_{n+1})$. Then by Theorem 3.7, $(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ and $x_{n+1} = y_{n+1}$. Hence, by the induction hypothesis, $x_1 = y_1$ and $\ldots$ and $x_n = y_n$ and $x_{n+1} = y_{n+1}$. □

The *Cartesian product* $V \times W$ of two sets V and W is by definition the set of all ordered pairs (v, w) with $v \in V$ and $w \in W$.

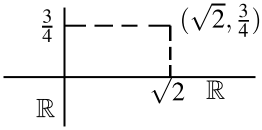


Definition 3.6 (Cartesian Product). $V \times W := \{x \mid$ there is some $v \in V$ and there is some $w \in W$ such that $x = (v, w)\}$, in other words, $V \times W := \{(v, w) \mid v \in V \wedge w \in W\}$.

Example 3.7.

$\{2,3\} \times \{4\} = \{(2,4), (3,4)\}$, $\{2,3\} \times \{4,5\} = \{(2,4), (3,4), (2,5), (3,5)\}$,
$\{1\} \times \{4,5\} = \{(1,4), (1,5)\}$, $\mathbb{R} \times \mathbb{R} = \{(x, y) \mid x \in \mathbb{R} \wedge y \in \mathbb{R}\}$.

So, $\mathbb{R} \times \mathbb{R}$ corresponds to the set of all points in the Euclidean plane:



'There is some $v \in V$ and there is some $w \in W$ such that $x = (v, w)$' can be formulated in our logical symbolism as follows: $\exists v \in V\ \exists w \in W\ [\ x = (v, w)\ ]$.
So, $V \times W = \{x \mid \exists v \in V\ \exists w \in W\ [\ x = (v, w)\ ]\}$.

From Definition 3.6 and Theorem 3.8 we immediately conclude:

Corollary 3.4. *$V \times W = \{x \in PP(V \cup W) \mid \exists v \in V\ \exists w \in W\ [\ x = (v, w)\ ]\}$, or simply*

$$V \times W = \{(v, w) \in PP(V \cup W) \mid v \in V \wedge w \in W\}.$$

From Corollary 3.4, the Axiom of Union, the Powerset Axiom and the Separation Axiom it follows that: if V and W are sets, then so is $V \times W$.

$\{2\} \times \{4\} = \{(2,4)\}$, but $\{4\} \times \{2\} = \{(4,2)\}$. So, it is not true that for all sets V and W, $V \times W = W \times V$; in other words, the operation × is not commutative. The operation × is not associative either (see Exercise 3.11).

Instead of $V \times V$ we usually write V^2.

Example 3.8. $\{3,4\}^2 = \{3,4\} \times \{3,4\} = \{(3,3),(3,4),(4,3),(4,4)\}$.

More generally, we define V^n ($n \in \mathbb{N}$, $n \geq 1$) inductively by:

Definition 3.7. $V^1 := V$, and $V^{n+1} := V^n \times V$.

Example: $\{3,4\}^3 = \{3,4\}^2 \times \{3,4\} = \{((3,3),3),((3,3),4),((3,4),3),((3,4),4),((4,3),3),((4,3),4),((4,4),3),((4,4),4)\}$.

More generally, we define the Cartesian product with finitely many factors:

Definition 3.8. $\mathrm{X}_{i=1}^{1} V_i = V_1$ and $\mathrm{X}_{i=1}^{n+1} V_i = (\mathrm{X}_{i=1}^{n} V_i) \times V_{n+1}$.

Example 3.9. Let $V_1 = \{1,2\}, V_2 = \{3,4\}$ and $V_3 = \{7,8,9\}$.
Then $\mathrm{X}_{i=1}^{3} V_i = (V_1 \times V_2) \times V_3 = (\{1,2\} \times \{3,4\}) \times \{7,8,9\}$.

3.4.2 Relations

We start with a few examples of *binary* relations R between the elements of a set V and the elements of a set W (or: between V and W). Instead of xRy – to be read as: x is in relation R to y – one also writes $R(x,y)$.

Example 3.10.

1. $V = M$(en)	$W = W$(omen)	xRy	:=	x is a son of y
2. $V = \mathbb{N}$	$W = \mathbb{N}$	xRy	:=	$y = x+1$
3. $V = \mathbb{N}$	$W = \mathbb{R}$	xRy	:=	$y = \sqrt{x}$
4. $V = \mathbb{N}^2$	$W = \mathbb{N}^2$	$(m,n)R(p,q)$	:=	$m-n = p-q$
5. $V = \mathbb{N} \times (\mathbb{Z} - \{0\})$	$W = V$	$(m,n)R(p,q)$	:=	$\frac{m}{n} = \frac{p}{q}$
6. $V = \mathbb{N}$	$W = P(\mathbb{N})$	xRy	:=	$x \in y$.

Below are some examples of a *ternary* relation R between the elements of a set V, the elements of a set W and the elements of a set U:
1. $V = M$(en), $W = W$(omen), $U = P$(eople); $R(x,y,z)$:= x and y are parents of z.
2. $V = W = U = \mathbb{N}$; $R(x,y,z) := x+y = z$.
For reasons of efficiency, we will at this point discuss only binary relations.

The adagium ‘everything is a set’ also applies to relations. A relation R between sets V and W can be represented by the set $\{(v,w) \in V \times W \mid vRw\}$. For instance, the relations in Example 3.10, 1 and 2 can be represented by the sets:

1. $\{(x,y) \in M \times W \mid x \text{ is a son of } y\}$
2. $\{(x,y) \in \mathbb{N} \times \mathbb{N} \mid y = x+1\}$

So, we may represent the mathematical notion of 'relation' by a set: each binary relation R between the elements of a set V and those of a set W determines a subset of $V \times W$; and, conversely, each subset of $V \times W$ determines a binary relation between the elements of V and those of W. Hence, the following definition makes sense.

Definition 3.9 (Relation). R is a (binary) *relation* between V and W := $R \subseteq V \times W$.
Notation: $xRy := (x,y) \in R$. One sometimes uses $R(x,y)$ instead of xRy.

For $R \subseteq V \times W$ we define the *domain* and the *range* of R: The domain of R is the set of all elements x in V which are related to at least one element y in W; the range of R is the set of all elements y in W which are related to at least one element x in V.

Definition 3.10 (Domain and Range).

$$\text{Dom}(R) := \{x \in V \mid \exists y \in W[\, xRy \,]\} \textit{ domain} \text{ of } R$$
$$\text{Ran}(R) := \{y \in W \mid \exists x \in V[\, xRy \,]\} \textit{ range} \text{ of } R$$

For the relations in Example 3.10 Dom(R) and Ran(R) are respectively:

	Dom(R)	Ran(R)
1.	the set of all men	the set of all mothers with at least one son
2.	$\mathbb{N}$	$\mathbb{N} - \{0\}$
3.	$\mathbb{N}$	$\{y \in \mathbb{R} \mid \exists x \in \mathbb{N}\, [\, y = \sqrt{x}\,]\}$
4.	$\mathbb{N}^2$	$\mathbb{N}^2$
5.	$\mathbb{N} \times (\mathbb{Z} - \{0\})$	$\mathbb{N} \times (\mathbb{Z} - \{0\})$
6.	$\mathbb{N}$	$P(\mathbb{N}) - \{\emptyset\}$

If $R \subseteq V \times V$, then R is simply a relation on V. Example 3.10, 2 gives a relation on $\mathbb{N}$, Example 3.10, 4 a relation on $\mathbb{N}^2$ and Example 3.10, 5 a relation on $\mathbb{N} \times (\mathbb{Z} - \{0\})$.

Since a relation R between (the elements of) V and (the elements of) W may be represented by the set $\{(x,y) \in V \times W \mid xRy\}$, the set theoretic operations of intersection, union, and complement also apply to relations: $R \cap S$, $R \cup S$ and $\overline{R}$.

Similarly, the set theoretic predicates of inclusion and equality apply to relations R and S: $R \subseteq S$ and $R = S$.

Below we define two special operations on relations: the *converse* $\breve{R}$, also called the *transposition* R^{T}, of R, and the composition $R;S$ of two relations R and S.

Definition 3.11 (Converse relation). Let R be a relation between V and W. Then the *converse* relation $\breve{R}$ of R is the relation between W and V, defined by $w\breve{R}v := vRw$. In set-theoretic terms, $\breve{R} := \{(w,v) \in W \times V \mid (v,w) \in R\}$.

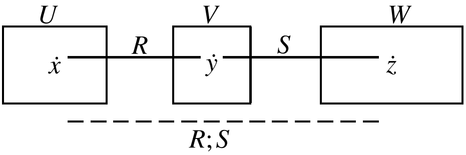
For the relations in Example 3.10, 1 - 4, the converse relations are respectively:

1. $\{(y,x) \in W \times M \mid y \text{ is the mother of } x\}$,
2. $\{(y,x) \in \mathbb{N} \times \mathbb{N} \mid x = y-1\}$,
3. $\{(y,x) \in \mathbb{R} \times \mathbb{N} \mid x = y^2\}$,
4. $\{(p,q),(m,n) \in \mathbb{N}^2 \times \mathbb{N}^2 \mid p-q = m-n\}$.

Note that in example 4, $\breve{R} = R$.

Let R be a relation between sets U and V and S a relation between sets V and W. Then the composition $R;S$ of R and S is the relation between U and W defined by $x(R;S)z :=$ there is some $y \in V$ such that xRy and ySz. In set theoretic terms:

Definition 3.12 (Composition). Let $R \subseteq U \times V$ and $S \subseteq V \times W$.
Then $R;S := \{(x,z) \in U \times W \mid \exists y \in V\ [\ (x,y) \in R \wedge (y,z) \in S\ ]\}$ is called the *composition* of R and S. Instead of $R;S$ one also writes $R \circ S$ and (in case R and S are functions) also $S \circ R$.



Example 3.11. 1. Let R be the relation of Example 3.10, 2, $R \subseteq \mathbb{N} \times \mathbb{N}$, defined by $xRy := y = x+1$, and let S be the relation of Example 3.10, 3, $S \subseteq \mathbb{N} \times \mathbb{R}$, defined by $ySz := z = \sqrt{y}$. Then

$$\begin{aligned} R;S &= \{(x,z) \in \mathbb{N} \times \mathbb{R} \mid \exists y \in \mathbb{N}\ [\ (x,y) \in R \wedge (y,z) \in S\ ]\ \} \\ &= \{(x,z) \in \mathbb{N} \times \mathbb{R} \mid \exists y \in \mathbb{N}\ [\ y = x+1 \wedge z = \sqrt{y}\ ]\ \} \\ &= \{(x,z) \in \mathbb{N} \times \mathbb{R} \mid z = \sqrt{x+1}\ \}. \end{aligned}$$

In other words, $x(R;S)z := z = \sqrt{x+!}$.
2. Let M be the set of all Men and $R \subseteq M \times M$ with $xRy := y$ is the father of x. Then

$$\begin{aligned} R;R &= \{(x,z) \in M \times M \mid \exists y \in M\ [\ (x,y) \in R \wedge (y,z) \in R]\ \} \\ &= \{(x,z) \in M \times M \mid \exists y \in M\ [\ y \text{ is the father of } x \text{ and } z \text{ is the father of } y\ ]\ \} \\ &= \{(x,z) \in M \times M \mid z \text{ is the grandfather of } x\ \}. \end{aligned}$$

In other words: $x(R;R)z := z$ is the grandfather of x.

Finally, we define some special relations: the empty relation O, the universal relation L and the identity relation I.

Definition 3.13. Let V and W be any sets. Then:
$\mathsf{L} := \{(x,y) \mid x \in V \wedge y \in W\}$ is the *universal* relation between V and W. So, $x\mathsf{L}y$ for any $x \in V$ and for any $y \in W$.
$\mathsf{O} := \emptyset$ is the *empty* relation between V and W. So, not $x\mathsf{O}y$, for any $x \in V$ and for any $y \in W$.
$\mathsf{I} := \{(x,x) \mid x \in V\}$ is the *identity* relation on V (or the *diagonal* of $V \times V$). So, $x\mathsf{I}x$ for any $x \in V$.

Notice that in fact we have for any two sets V and W a universal, an empty and an identity relation.

Also notice that in case V and W are finite sets, a relation R between V and W may be represented by a Boolean matrix. For instance, let R be the relation between $V = \{1,2,3\}$ and $W = \{1,2,3,4,5,6\}$ defined by $xRy := y = 2 \cdot x$. Then R may be represented by the following Boolean matrix:

	1	2	3	4	5	6
1		■				
2				■		
3						■

A Boolean matrix interpretation of relations is well suited for many purposes and also used as one of the graphical representations of relations within *RelView*, a software tool for the evaluation of relation-algebraic expressions. The RelView system is an interactive tool for computer-supported manipulation of relations represented as Boolean matrices or directed graphs.

3.4.3 Equivalence Relations

$25 \neq 13$ and $13 \neq 1$, but 25 o'clock = 13 o'clock = 1 o'clock.
$26 \neq 14$ and $14 \neq 2$, but 26 o'clock = 14 o'clock = 2 o'clock.
and so on.

In reading off the clock we call two natural numbers equal if their difference is a multiple of twelve. Therefore, we consider the following relation R on the set $\mathbb{N}$ of the natural numbers: $nRm := n - m$ is a multiple of twelve.

In symbols: $nRm := \exists k \in \mathbb{Z}\ [\, n - m = 12 \cdot k \,]$.

Definition 3.14 (Equivalence relation). A relation R on a set V is an *equivalence relation* on V := R is reflexive, symmetric and transitive, where
R is *reflexive* := for all $x \in V$, xRx;
R is *symmetric* := for all $x,\ y \in V$, if xRy, then yRx;
R is *transitive* := for all $x,\ y,\ z \in V$, if xRy and yRz, then xRz.

Example 3.12. 1. The relation R on the set $\mathbb{N}$, defined by $nRm := n - m$ is a multiple of twelve, is an equivalence relation on $\mathbb{N}$.
2. The relation = on $\mathbb{N}$ is an equivalence relation.
3. The relation R on the set $\mathbb{N}^2$, defined by $(m,n)R(p,q) := m+q = n+p$ (or $m-n = p-q$), is an equivalence relation on $\mathbb{N}^2$.
4. The relation R on the set $\mathbb{N} \times (\mathbb{Z} - \{0\})$, defined by $(m,n)R(p,q) := m \cdot q = n \cdot p$ (or $\frac{m}{n} = \frac{p}{q}$), is an equivalence relation on $\mathbb{N} \times (\mathbb{Z} - \{0\})$.
5. The relation *is parallel to or is equal to* on the set of all straight lines in the Euclidean plane is an equivalence relation.

Definition 3.15 (Equivalence class). Let R be an equivalence relation on a set V. The *equivalence class* $[v]_R$, also called *v modulo R*, of an element v of V with respect to R is by definition the subset of V, consisting of all those elements w in V for which vRw. Instead of $[v]_R$ one sometimes writes v/R.

$$[v]_R := \{w \in V \mid vRw\}$$

v is called a *representative* of the class $[v]_R$. Note that if R is an equivalence relation on V, then for all $v,\ w \in V$, vRw iff $[v]_R = [w]_R$.

Example 3.13. We now give the equivalence classes $[v]_R$ for the equivalence relation R on $\mathbb{N}$ from Example 3.12, 1, where $nRm := n - m$ is a multiple of 12.

$[0]_R = \{0, 12, 24, 36, \ldots\}$, $[12]_R = [0]_R$, $[24]_R = [0]_R$.
$[1]_R = \{1, 13, 25, 37, \ldots\}$, $[13]_R = [1]_R$, $[25]_R = [1]_R$.
$[2]_R = \{2, 14, 26, 38, \ldots\}$, $[14]_R = [2]_R$, $[26]_R = [2]_R$.
$\vdots$
$[11]_R = \{11, 23, 35, 47, \ldots\}$, $[23]_R = [11]_R$, $[35]_R = [11]_R$.
Thus, it would be more appropriate to indicate the numerals on the clock by $[1]_R, [2]_R, \ldots, [11]_R, [12]_R$ instead of $1, 2, \ldots, 11, 12$.

One may show that the integers and the rational numbers can be defined in terms of the natural numbers, making use of the equivalence relations R from Example 3.12, 3 and 4 respectively. So, roughly speaking, one may say that the natural numbers form the basis of all mathematics. For instance, $-1 := [(1,2)]_R$ with $(m,n)R(p,q) := m+q = n+p$ (or $m-n = p-q$) and $\frac{2}{3} := [(2,3)]_R$ with $(m,n)R(p,q) := m \cdot q = n \cdot p$ (or $\frac{m}{n} = \frac{p}{q}$). See van Dalen, Doets, de Swart, [3].

Definition 3.16 (Quotient set). Let R be an equivalence relation on V. The *quotient set* V/R or V *modulo* R is the set of all equivalence classes $[v]_R$ with $v \in V$.
In other words: $V/R := \{[v]_R \mid v \in V\}$.

As an example let us consider the quotient set from Example 3.13 above, where R is the equivalence relation on $\mathbb{N}$ defined by $nRm := n - m$ is a multiple of twelve.

$$\mathbb{N}/R = \{[1]_R, [2]_R, \ldots, [11]_R, [12]_R\}.$$

$\mathbb{N}/R$ has twelve elements, corresponding to the twelve numerals on the clock. The twelve different elements of $\mathbb{N}/R$ are pairwise disjoint, i.e., $[n]_R \cap [m]_R = \emptyset$ for $n \neq m$ and $1 \leq n, m \leq 12$, and together they form the whole set $\mathbb{N}$, more precisely,

$$[1]_R \cup [2]_R \cup \ldots \cup [11]_R \cup [12]_R = \mathbb{N}.$$

Therefore we call $\mathbb{N}/R$ a *partition* of $\mathbb{N}$:

$$\mathbb{N} \left\{ \begin{array}{rcl} [1]_R & = & \{1, 13, 25, 37, \ldots\} \\ & \vdots & \\ [11]_R & = & \{11, 23, 35, 47, \ldots\} \\ [12]_R & = & \{0, 12, 24, 36, \ldots\} \end{array} \right.$$

Definition 3.17 (Partition). A collection U consisting of subsets of V is a *partition* of V := 1) V = the union of all elements of U, and 2) the different elements of U are pairwise disjoint.

Clearly, every partition U consisting of subsets of V defines an equivalence relation R: xRy iff x and y belong to the same element of U. Conversely,

Theorem 3.10. *If R is an equivalence relation on V, then V/R is a partition of V.*

Proof. We have to show: 1) V = the union of all elements in V/R, and 2) the different elements of V/R are pairwise disjoint.
1) Let $v \in V$. Then $v \in [v]_R$. Conversely, if $w \in [v]_R$, then $w \in V$.

2) Suppose $[v]_R \neq [w]_R$. Then not vRw. (1)
Now suppose $[v]_R \cap [w]_R \neq \emptyset$. Then for some $u \in V$, $u \in [v]_R$ and $u \in [w]_R$. But then vRu and uRw, and consequently – since R is an equivalence relation – vRw. This is a contradiction of (1). Therefore, $[v]_R \cap [w]_R = \emptyset$ if $[v]_R \neq [w]_R$. □

3.4.4 Functions

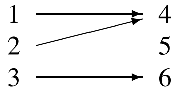
Let V and W be sets. 'f is a (total) *function* or mapping from V to W' means intuitively: f assigns to each $v \in V$ a uniquely determined $w \in W$. *Notation*: $f : V \to W$.

For each $v \in V$, the uniquely determined $w \in W$, which is assigned by f to v, is called the *image* (under f) of v. *Notation*: $w = f(v)$.

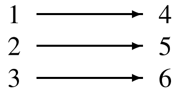
An example from daily life is the function f from the set M of all men to the set W of all women, which assigns to every person x his or her mother $f(x)$.

Example 3.14. Examples of functions $f : V \to W$:

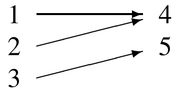
1. $V = \{1, 2, 3\}$, $W = \{4, 5, 6\}$, $f(1) = 4$
$f(2) = 4$
$f(3) = 6$



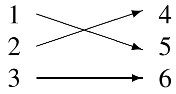
2. $V = \{1, 2, 3\}$, $W = \{4, 5, 6\}$, $f(1) = 4$
$f(2) = 5$
$f(3) = 6$



3. $V = \{1, 2, 3\}$, $W = \{4, 5\}$, $f(1) = 4$
$f(2) = 4$
$f(3) = 5$



4. $V = \{1, 2, 3\}$, $W = \{4, 5, 6\}$, $f(1) = 5$
$f(2) = 4$
$f(3) = 6$

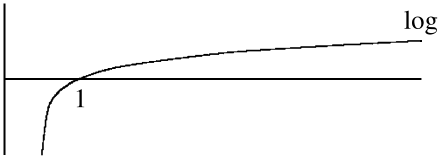


5. $V = \mathbb{N}$, $W = \mathbb{N}$, $\begin{cases} f(n) = 0 \text{ if } n \text{ is even,} \\ f(n) = 1 \text{ if } n \text{ is odd.} \end{cases}$

6. $V = \mathbb{N}$, $W = P(\mathbb{N})$, $f(n) = \{n\}$.

7. $V = \mathbb{N}^2$, $W = \mathbb{Z}$, $f((n,m)) = n - m$.

8. $V = \mathbb{R}_+$ with $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x > 0\}$, $W = \mathbb{R}$, $f(x) = \log(x)$.



If $f : V \to W$, then f determines a set of ordered pairs, namely, $\{(v,w) \in V \times W \mid w = f(v)\}$. This set, known as the *graph* of f, has the property that for each v in V there is a unique element w in W such that (v,w) is in the set (namely $w = f(v)$). Conversely, each subset of $V \times W$ with this special property will determine a function $f : V \to W$.

The graphs of the functions from Example 3.14 are respectively:

1. $\{(1,4),(2,4),(3,6)\}$, 2. $\{(1,4),(2,5),(3,6)\}$,
3. $\{(1,4),(2,4),(3,5)\}$, 4. $\{(1,5),(2,4),(3,6)\}$,
5. $\{(n,m) \in \mathbb{N}^2 \mid (n \text{ is even} \wedge m = 0) \vee (n \text{ is odd} \wedge m = 1)\}$,
6. $\{(n,y) \in \mathbb{N} \times P(\mathbb{N}) \mid y = \{n\}\}$,
7. $\{((n,m),y) \in \mathbb{N}^2 \times \mathbb{Z} \mid y = n - m\}$,
8. $\{(x,y) \in \mathbb{R}_+ \times \mathbb{R} \mid y = \log(x)\}$.

Any function can thus be represented by its graph. In fact, it is common in set theory to identify a function with its graph and thus reduce the notion of function to the notion of set. This is what we will do.

Definition 3.18 (Function). f is a (total) *function* from V to W := f is a relation between V and W, such that for each $v \in V$ there is a unique $w \in W$ such that $(v,w) \in f$. **Notation**: $f : V \to W$.

Because a function $f : V \to W$ is by definition a relation, Definition 3.10 defines the domain Dom(f) and the range Ran(f) of f. It is evident that for $f : V \to W$, Dom(f) = V and Ran(f) = $\{w \in W \mid \exists v \in V\ [\ w = f(v)\ ]\}$. For instance, for the function f in Example 3.14, 1, Ran(f) = $\{4,6\}$; and in Example 3.14, 2, Ran(f) = $\{4,5,6\}$.

We shall maintain the notation introduced earlier, that we write $f(v)$ for the unique $w \in W$ such that $(v,w) \in f$. Thus we have, for all $v \in V$, $w \in W$: $w = f(v)$ if and only if $(v,w) \in f$. From time to time we will write $v \longmapsto f(v)$ for $(v, f(v)) \in f$.

Sometimes it is convenient to have at one's disposal also the notion of *partial function*. Intuitively, a partial function f from V to W assigns to some (not necessarily all) $v \in V$ a uniquely determined $w \in W$.

Definition 3.19 (Partial function). f is a *partial function* from V to W := f is a relation between V and W, such that for all $v \in V$ and w, $w' \in W$, if $(v,w) \in f$ and $(v,w') \in f$, then $w = w'$.

If f is a partial function from V to W, then Dom(f) := $\{v \in V \mid \text{there is a } w \in W \text{ such that } (v,w) \in f\}$. If f is a (total) function from V to W, then Dom(f) = V.

Definition 3.20. If $f : V \to W$ and $V' \subseteq V$, then $f(V') := \{f(v) \mid v \in V'\}$.
If $f : V \to W$ and $W' \subseteq W$, then $f^{-1}(W') := \{v \in V \mid f(v) \in W'\}$.

The notation $f(V')$ may be ambiguous, because a subset of V may at the same time be an element of V.

Remark: Let W be any set. Then $\emptyset \subseteq \emptyset \times W$. Further, because $\emptyset$ has no elements, it follows that for each $v \in \emptyset$ there is a unique $w \in W$ such that $(v,w) \in \emptyset$. Hence, by Definition 3.18, $\emptyset$ is a function from $\emptyset$ to W, in other words $\emptyset : \emptyset \to W$. Since $\emptyset$ is the only relation with Dom($\emptyset$) = $\emptyset$, $\emptyset$ is also the only function from $\emptyset$ to W.

If $f : V \to W$, then $f \subseteq V \times W$ and hence, $f \in P(V \times W)$.

Definition 3.21 (Set of all functions $f : V \to W$).
W^V := the set of all functions $f : V \to W$, i.e., $W^V := \{f \in P(V \times W) \mid f : V \to W\}$. So, if V and W are sets, then by the separation axiom W^V is a set too.

Example 3.15. The set $\{1,2,3\}^{\{5,6\}}$ has $3^2 = 9$ elements $f_1, \ldots, f_9$, the functions $f_1, \ldots, f_9$ being defined by the following scheme:

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
5	1	1	1	2	2	2	3	3	3
6	1	2	3	1	2	3	1	2	3

$$\text{i.e., } \begin{cases} f_1(5) = 1,\ f_2(5) = 1,\ \ldots,\ f_9(5) = 3, \\ f_1(6) = 1,\ f_2(6) = 2,\ \ldots,\ f_9(6) = 3. \end{cases}$$

The reader should check for him or her self that $\{5,6\}^{\{1,2,3\}}$ has $2^3 = 8$ elements.

Theorem 3.11. *If W is a set with m elements and V is a set with n elements ($m, n \in \mathbb{N}$), then W^V has m^n elements.*

So, if W is a set with 10 elements and V has 6 elements, then there are, by this theorem, 10^6, i.e., one million, functions $f : V \to W$.

Proof. Throughout the following argument, let $m \in \mathbb{N}$ be fixed, and let W be a fixed set with m elements. Let $\Phi(n)$:= if V is any set with n elements, then W^V has m^n elements. Then Theorem 3.11 says: for every $n \in \mathbb{N}$, $\Phi(n)$.

By induction it suffices to show: $\Phi(0)$ and for all $k \in \mathbb{N}$, $\Phi(k) \to \Phi(k+1)$.
Induction basis $\Phi(0)$: if V has 0 elements, i.e., $V = \emptyset$, then $\emptyset$ is the only function from V to W; hence, $W^V = \{\emptyset\}$; so $W^\emptyset$ has $m^0 = 1$ element.
Induction step $\Phi(k) \to \Phi(k+1)$: Suppose $\Phi(k)$, i.e., if V is any set with k elements, then W^V has m^k elements. We must now show that $\Phi(k+1)$ holds. So let $\{v_1, \ldots, v_k, v_{k+1}\}$ be a set with $k+1$ elements. By the induction hypothesis $\Phi(k)$ there are m^k different functions from $\{v_1, \ldots, v_k\}$ to W.

	f_1	f_2	$\cdots$	f_{m^k}
v_1	$*$	$*$		$*$
v_2	$*$	$*$		$*$
$\vdots$				
v_k	$*$	$*$		$*$
v_{k+1}				

For each i, $1 \leq i \leq m^k$, there are now m different possible choices for $f_i(v_{k+1})$. Thus, there are $m \cdot m^k = m^{k+1}$ different functions from $\{v_1, \ldots, v_k, v_{k+1}\}$ to W. □

In mathematics (especially analysis) one frequently uses sequences of objects. We can now give an exact formulation of the notion of sequence.

Definition 3.22 (Sequence). An (infinite) *sequence* of elements of V is a function f from $\mathbb{N}$ to V. **Notation**: $f(0), f(1), f(2), \ldots$.
A (finite) *sequence* of elements of V is a function f from $\{0, \ldots, n\}$ to V, for some $n \in \mathbb{N}$. **Notation**: $f(0), \ldots, f(n)$.

The functions $f : V \to W$ in Example 3.14, 2, 4, 6 and 8 have the property that they assign distinct elements of W to distinct elements of V; in other words: for all $v, v' \in V$, if $v \neq v'$, then $f(v) \neq f(v')$, or (equivalently): for all $v, v' \in V$, if $f(v) = f(v')$, then $v = v'$. We call such functions *injective* (one-to-one). Notice that the other functions in Example 3.14 do not have this property.

Definition 3.23 (Injection). $f : V \to W$ is *injective* or an *injection* := for all $v, v' \in V$, if $v \neq v'$, then $f(v) \neq f(v')$. In logical notation: $\forall x \in V\ \forall x' \in V\ [\ x \neq x' \to f(x) \neq f(x')\ ]$. **Notation**: Intuitively, the existence of an injection $f : V \to W$ means that the set V cannot be larger than W; therefore we write $f : V \leq_1 W$ to indicate that $f : V \to W$ is injective.

The functions $f : V \to W$ in Example 3.14, 2, 3, 4, 7 and 8 have the property that each element $w \in W$ is the image (under f) of an element $v \in V$. We call such functions *surjective* (onto). Note that the other functions in Example 3.14 do not have this property.

Definition 3.24 (Surjection). $f : V \to W$ is *surjective* or a *surjection* := for every $w \in W$ there is a $v \in V$ such that $w = f(v)$. In logical notation: $\forall y \in W\ \exists x \in V\ [\ y = f(x)\ ]$. In other words, $f : V \to W$ is surjective if and only if Ran(f) = W.

The functions in Example 3.14, 1 and 5 are neither injective nor surjective. Those in Example 3.14, 2, 4 and 8 have both properties. We call such functions *bijective*.

Definition 3.25 (Bijection). $f : V \to W$ is *bijective* or a *bijection* := f is both injective and surjective. **Notation**: Intuitively, the existence of a bijection $f : V \to W$ means that the sets V and W are equally large; therefore one writes $f : V =_1 W$ to indicate that $f : V \to W$ is bijective.

A bijection $f : V \to W$ gives a **one-one correspondence** *between the elements of V and the elements of W: for each $v \in V$ there is exactly one (f is a function) $w \in W$ such that $w = f(v)$ and for each $w \in W$ there is at least one (f is surjective) and precisely one (f is injective) $v \in V$ such that $w = f(v)$.*

Definition 3.26 (Canonical function). Let R be an equivalence relation on V. The canonical function $f : V \to V/R$ is defined by $f(x) := [x]_R$. It is of course surjective, but in general not injective.

Definition 3.27 (Characteristic function). Let $U \subseteq V$. The *characteristic function* $K_U : V \to \{0,1\}$ *of* U is defined by $K_U(v) = \begin{cases} 1 \text{ if } v \in U, \\ 0 \text{ if } v \notin U. \end{cases}$

In the special case that $U \subseteq \mathbb{N}$, the characteristic function $K_U : \mathbb{N} \to \{0,1\}$ of U may be represented by the infinite sequence $K_U(0), K_U(1), K_U(2), K_U(3), \ldots$ of 0's and 1's (see Definition 3.22). For instance, let $U = \{0,2,4,6,\ldots\}$, then $K_U = 1\ 0\ 1\ 0\ 1\ 0\ 1\ \ldots$.

Since we have defined a function $f : V \to W$ as a set $\{(v,w) \in V \times W \mid w = f(v)\}$ of ordered pairs, the equality relation between functions is thereby determined. Let

$f : V \to W$ and $g : V \to W$. Then, by the axiom of extensionality: $f = g$ iff f and g have the same elements, i.e., for all $v \in V$ and for all $w \in W$, $(v,w) \in f$ iff $(v,w) \in g$. In other words, $f = g :=$ for all $v \in V$ and for all $w \in W$, $w = f(v)$ iff $w = g(v)$. So, for $f,g : V \to W$, $f = g$ iff for all $v \in V$, $f(v) = g(v)$.

In logical notation: $f = g := \forall x \in V[f(x) = g(x)]$.

Theorem 3.12. *The function $K : P(V) \to \{0,1\}^V$, defined by $K(U) := K_U$ (i.e., K assigns to each subset U of V the characteristic function K_U of U) is a bijection.*

Proof. We first show that K is injective. So, suppose $U_1 \neq U_2$, i.e., there is some $v \in V$ such that ($v \in U_1$ and $v \notin U_2$) or ($v \in U_2$ and $v \notin U_1$). Then ($K_{U_1}(v) = 1$ and $K_{U_2}(v) = 0$) or ($K_{U_2}(v) = 1$ and $K_{U_1}(v) = 0$). So, there is a $v \in V$ such that $K_{U_1}(v) \neq K_{U_2}(v)$, and hence $K_{U_1} \neq K_{U_2}$.
Next we show that K is surjective. Suppose $f \in \{0,1\}^V$. Let $U_f := \{v \in V \mid f(v) = 1\}$. Then for all $v \in V$, $K_{U_f}(v) = 1$ iff $v \in U_f$, i.e, for all $v \in V$, $K_{U_f}(v) = 1$ iff $f(v) = 1$. Hence, for all $v \in V$, $K_{U_f}(v) = f(v)$. Therefore, $f = K_{U_f}$. □

Let $f : U \to V$ and $g : V \to W$. Since f and g are (special) relations, the composition $f;g$ of f and g has been defined according to Definition 3.12.

$$\begin{tikzcd}
\boxed{U} \arrow[r, "f"] \arrow[rr, bend right, "f;g"'] & \boxed{V} \arrow[r, "g"] & \boxed{W}
\end{tikzcd}$$

Applying $f;g$ to an element $x \in U$, we first apply f to x and next g to $f(x)$, resulting in $g(f(x))$. So, in the case of the composition of functions $f : U \to V$ and $g : V \to W$ it is attractive to write $g \circ f$ instead of $f;g$, where $(g \circ f)(x) := g(f(x))$.

Definition 3.28 (Composition of functions). Let $f : U \to V$ and $g : V \to W$. Then the *composition* $g \circ f : U \to W$ of f and g is defined by $(g \circ f)(x) = g(f(x))$.

Example 3.16. Let $f : \mathbb{N} \to \mathbb{Z}$ be defined by $f(n) := -n$. Let $g : \mathbb{Z} \to \mathbb{Q}$ be defined by $g(m) := \frac{1}{2}m$. Then $g \circ f : \mathbb{N} \to \mathbb{Q}$ is defined by $(g \circ f)(n) = -\frac{1}{2}n$.

If $f : V \to W$ is a bijection, then there is – because f is surjective – for each $w \in W$ at least one $v \in V$ such that $w = f(v)$, and – because f is injective – there is for each $w \in W$ at most one $w \in V$ such that $w = f(v)$. Hence, if $f : V \to W$ is a bijection, then for each $w \in W$ there is precisely one $v \in V$ such that $w = f(v)$.

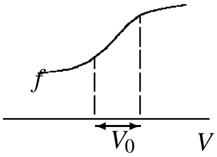
Definition 3.29 (Inverse function). Let $f : V \to W$ be a bijection. Then the *inverse function* $f^{-1} : W \to V$ is defined by $f^{-1}(w) :=$ the unique element v in V such that $w = f(v)$.

Note that the inverse function f^{-1} of a bijection f equals the converse $\check{f}$ of f (see Definition 3.11). If $f : V \to W$ is a bijection, then $f^{-1} \circ f : V \to V$ is the identity function on V and $f \circ f^{-1} : W \to W$ is the identity function on W.

Example 3.17. Let $\mathbb{N}_{even}$ be the set of all even natural numbers and define $f : \mathbb{N} \to \mathbb{N}_{even}$ by $f(n) := 2n$. Then $f : \mathbb{N} \to \mathbb{N}_{even}$ is a bijection and $f^{-1} : \mathbb{N}_{even} \to \mathbb{N}$ is defined by $f^{-1}(m) := \frac{1}{2}m$.

Let $\mathbb{R}_+$ be the set of all real numbers greater than 0 and define $f : \mathbb{R}_+ \to \mathbb{R}$ by $f(x) := \log(x)$ (see Example 3.14, 8). Then $f : \mathbb{R}_+ \to \mathbb{R}$ is a bijection and $f^{-1} : \mathbb{R} \to \mathbb{R}_+$ is defined by $f^{-1}(x) := e^x$.

Definition 3.30. Let $f : V \to W$ and $V_0 \subseteq V$. Then the *restriction* $f\lceil V_0 : V_0 \to W$ is defined by $(f\lceil V_0)(x) := f(x)$.



Example 3.18. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) := \sin \pi x$. Then $f\lceil\mathbb{Z} : \mathbb{Z} \to \mathbb{R}$ is defined by $(f\lceil\mathbb{Z})(m) = \sin \pi m = 0$ (for $m \in \mathbb{Z}$).

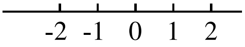
3.4.5 Orderings

We start with giving six examples of an ordering relation R on a given set V.

Example 3.19.
1. $V = P(\{v,w\}) = \{\emptyset, \{v\}, \{w\}, \{v,w\}\}$ with $xRy := x \subseteq y$.

$$\begin{tikzcd}
 & \{v, w\} & \\
\{v\} \arrow[ur, phantom, "\subseteq"] & & \{w\} \arrow[ul, phantom, "\supseteq"] \\
 & \emptyset \arrow[ul, phantom, "\supseteq"] \arrow[ur, phantom, "\subseteq"] &
\end{tikzcd}$$

2. $V = \{1, 2, 3, 4, 6, 8, 12, 24\}$ with $xRy := x$ is a divisor of y.
3. V is the set M of all men with $xRy := x$ is at least as old (in years) as y.
4. $V = \mathbb{Z}$ with $xRy := x \leq y$.



5. $V = \mathbb{N}$ with $xRy := x \leq y$.
6. $V = \mathbb{N} \times \mathbb{N}$ and $(n,m)R(x,y) := n \leq x$ or ($n = x$ and $m \leq y$).
   $(0,0), (0,1), (0,2), \ldots, (1,0), (1,1), (1,2), \ldots, (2,0), \ldots$

The ordering in example 6 is similar to the well-known ordering of words in a dictionary. Therefore we call this ordering the *lexicographic ordering* on $\mathbb{N} \times \mathbb{N}$.

Definition 3.31 (Partial ordering).
A relation R on a set V is a *partial ordering* on V :=
1. R is *reflexive*, i.e., for all $x \in V$, xRx, and
2. R is *anti-symmetric*, i.e., for all $x, y \in V$, if xRy and yRx, then $x = y$, and
3. R is *transitive*, i.e., for all $x, y, z \in V$, if xRy and yRz, then xRz.

The reader should check that all relations in Example 3.19 are a partial ordering on the given set V. Instead of 'R is a partial ordering on V' one sometimes says: V is a set, partially ordered by R, or: R partially orders V, or: (V, R) is a partially ordered set. If it is clear from the context what partial ordering relation is involved, we may write: V is a partially ordered set.

The relations 1 and 2 in Example 3.19 do not have the property that any two elements are comparable via R: for instance, for $v \neq w$, $\{v\} \not\subseteq \{w\}$ and $\{w\} \not\subseteq \{v\}$. The other relations in Example 3.19 do have the property that for all $x, y \in V$, xRy or yRx (or both). In the case that R expresses the (weak) preference of an agent (voter) or a society over the elements of a set V of alternatives or candidates, reading xRy as 'the agent judges x is at least as good as y', 'xRy and yRx' expresses that the agent is indifferent between x and y. Anti-symmetry then expresses that indifference between two distinct elements of V does not occur and transitivity expresses that the preference of the agent is rational.

Definition 3.32 (Complete relation). A relation R on a set V is *complete* := for all $x, y \in V$, xRy or yRx. In other words, any two elements in V are related via R.

Notice that a complete relation on V is by definition reflexive: taking $x = y$, (xRy or yRx) implies xRx.

Definition 3.33 (Weak ordering). A relation R on a set V is a *weak ordering* on V := R is complete and transitive.

The relations in Example 3.19, 3, 4, 5 and 6 are a weak ordering on the given set V. Notice that the third relation is not anti-symmetric: two different men may have the same age; however, the fourth, fifth and sixth are anti-symmetric.

Definition 3.34 (Linear ordering). R is a *linear* or *total* ordering or simply an *ordering* on V := R is weak ordering on V that in addition is anti-symmetric, i.e.,
1. R is complete: for all $x, y \in V$, xRy or yRx; and hence, in particular, xRx;
2. R is anti-symmetric: for all $x, y \in V$, if xRy and yRx, then $x = y$.
3. R is transitive: for all $x, y, z \in V$, if xRy and yRz, then xRz.

Relation 3 in Example 3.19 is not a linear ordering; the relations 4, 5 and 6 in Example 3.19 are linear orderings on the given sets. Whenever we refer to a subset W of a partially or totally ordered set (V, R), we will usually think of this subset W as being partially, resp. totally ordered by the restriction of R to W, i.e., $R \cap (W \times W)$.

Let R be a weak (preference) ordering on a set V of alternatives, reading xRy as: the agent (voter, judge) weakly prefers x to y, in other words: the agent judges that x is at least as good as y. Then we can express 'the agent strictly prefers x to y' by: xRy and not yRx, which we abbreviate by xPy.

Definition 3.35 (Strict associated ordering). Let R be an ordering on V. The *strict associated ordering P* of R on V is defined by xPy := xRy and not yRx.

Theorem 3.13. *Let R be a (total or linear) ordering on V. Let xPy := xRy and not yRx. Then P satisfies the following properties: 1. for all $x \in V$, not xPx;*

2. P is asymmetric, i.e, for all $x, y \in V$, if xPy, then not yPx;
3. P is transitive; and
4. P is connected, i.e., for all $x, y \in V$, xPy or $x = y$ or yPx.

Proof. Let R be a (total or linear) ordering on V and let $xPy := xRy$ and not yRx.
1. From this definition follows immediately that not xPx.
2. Suppose xPy, i.e., xRy and not yRx. Then certainly not yPx.
3. Suppose xPy and yPz, i.e,, xRy and yRz and hence, by transitivity of R, xRz. Also, not yRx and not zRy. In order to show xPz, we still have to show that not zRx. So, suppose zRx. Then by xRy and the transitivity of R, zRy. Contradiction.
4. It suffices to show: if $x \neq y$, then xPy or yPx. So suppose $x \neq y$. Then, because R is anti-symmetric: not xRy or not yRx (1). Because R is complete: xRy or yRx (2). From (1) and (2) follows: (not xRy and yRx) or (not yRx and xRy), i.e., yPx or xPy. □

The ordered set $(\mathbb{N}, \leq)$ has the property that each non-empty subset of $\mathbb{N}$ has a least (with respect to $\leq$) element. The ordered sets $(\mathbb{Z}, \leq)$ and $(\mathbb{Q}, \leq)$ do not have this property.

Definition 3.36 (Well-ordering). A relation R on a set V is a *well-ordering* on V :=
1. R is an (total) ordering on V, and
2. each non-empty subset of V has a least element (with respect to R), i.e., an element $x \in V$ such that for all $y \in V$, xRy.

So, the set $(\mathbb{N}, \leq)$ is well-ordered, but the sets $(\mathbb{Z}, \leq)$ and $(\mathbb{Q}, \leq)$ are not.

3.4.6 Structures and Isomorphisms

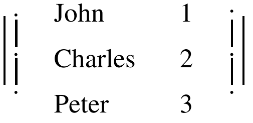
Frequently one is not interested in how the elements of a given set have been constructed, only in how they behave under certain given relations (and operations) on the set. For instance, given a certain set V of people, one may be interested only in how the people in the set behave under the relation 'is father of', or under the relation 'is older than', or under the relation 'is stronger than'; and sometimes one is interested in more than one relation on the same set. This brings us to the notion of structure.

Definition 3.37 (Structure). $\langle V, R_0, \ldots, R_k \rangle$ is a (relational) *structure* := V is a set and $R_0, \ldots, R_k$ are relations on V.

Remark: A more general notion of structure is obtained by considering sets together with certain relations and operations on them; see, for instance, [3].

Example 3.20. Examples of (relational) structures:
1. $\langle$ {Charles, John, Peter}, is older than $\rangle$;
2. $\langle$ {Charles, John, Peter}, is older than, is stronger than $\rangle$;
3. $\langle \mathbb{N}, < \rangle$, where $m < n$:= m is less than n;
4. $\langle \mathbb{N}, <, \mid \rangle$, where $m \mid n$:= m is a divisor of n;
5. $\langle \mathbb{N}_{even}, < \rangle$, where $\mathbb{N}_{even}$ is the set of all even natural numbers;
6. $\langle \mathbb{N}_{even}, <, \mid \rangle$.

Now, let us suppose that John is older than Charles and that Charles is older than Peter. Then there is no difference, as far as order properties are concerned, between the set $\{$Charles, John, Peter$\}$ together with the ordering relation 'is older than' and the set $\{1, 2, 3\}$ together with the ordering relation $<$. In both cases we get the same picture or structure:



where the vertical line denotes in the left picture the relation 'is older than' and in the right picture the relation 'is less than'. For that reason we call the two structures $\langle$ $\{$Charles, John, Peter$\}$, is older than $\rangle$ and $\langle\ \{1, 2, 3\}, <\ \rangle$ *isomorphic.*

Definition 3.38 (Isomorphism). Let $\langle V,\ R_0,\dots,R_k\rangle$ and $\langle W,\ S_0,\dots,S_k\rangle$ be two (relational) structures such that for each $i = 0,\dots,k$, R_i and S_i have the same number, say n_i, of arguments; for convenience, suppose $n_i = 2$ for all i. Let $f : V \to W$.
f is an *isomorphism* from $\langle V,\ R_0,\dots,R_k\rangle$ to $\langle W,\ S_0,\dots,S_k\rangle$:=
1. f is a bijection from V to W, and
2. for all $i = 0,\dots,k$ and for all $v, w \in V$, $R_i(v,w)$ iff $S_i(f(v), f(w))$.

Example 3.21 (Isomorphisms).

1) f : $\{$John, Charles, Peter$\} \to \{1, 2, 3\}$, defined by f(John) = 1, f(Charles) = 2, f(Peter) = 3, is an isomorphism from $\langle\{$John, Charles, Peter$\}$, is older than $\rangle$ to $\langle\{1,2,3\},<\ \rangle$, under the supposition that John is older than Charles and that Charles is older than Peter.
2) $f : \mathbb{N} \to \mathbb{N}_{even}$, defined by $f(n) = 2n$, is an isomorphism from $\langle\mathbb{N},<\rangle$ to $\langle\mathbb{N}_{even},<\rangle$ and likewise an isomorphism from $\langle\mathbb{N},\ <,\ \mid\ \rangle$ to $\langle\mathbb{N}_{even},\ <,\ \mid\ \rangle$, where $\mid$ is the divisibility-relation.

$$\begin{array}{rl} & \mathbb{N} \quad\ :0\ 1\ 2\ 3\ 4\ \dots \\ f\ \downarrow & \\ & \mathbb{N}_{even} :0\ 2\ 4\ 6\ 8\ \dots \end{array}$$

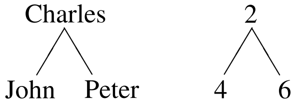
3) Let us suppose that John is the father of Charles and that Charles is the father of Peter. Then the function f, defined in 1), is NOT an isomorphism from $\langle\{$John, Charles, Peter$\}$, is father of $\rangle$ to $\langle\{1,2,3\},<\ \rangle$, since $1 < 3$, i.e., f(John) $< f$(Peter), but not (John is the father of Peter).
4) $f : \mathbb{N} \to \mathbb{Z}$, defined by $f(2n) = n$ and $f(2n-1) = -n$, is a bijection from $\mathbb{N}$ to $\mathbb{Z}$, but it is not an isomorphism from $\langle\mathbb{N},<\rangle$ to $\langle\mathbb{Z},<\rangle$, since $0 < 1$, but not $f(0) < f(1)$.

$$\begin{array}{rl} & \mathbb{N} :\ 0 \quad\ \ 1 \quad 2 \quad\ \ 3 \quad 4\ \dots \\ f\ \downarrow & \\ & \mathbb{Z} :\ 0\ \ -1\ \ 1\ \ -2\ \ 2\ \dots \end{array}$$

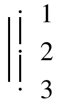
Definition 3.39 (Isomorphic). $\langle V,\ R_0,\dots,R_k\rangle$ is *isomorphic* to $\langle W,\ S_0,\dots,S_k\rangle$:= there is at least one isomorphism f from $\langle V,\ R_0,\dots,R_k\rangle$ to $\langle W,\ S_0,\dots,S_k\rangle$.

Example 3.22 (Isomorphic).

1) Supposing that John and Peter are equally strong and that Charles is stronger than John and Peter, $\langle\{\text{Charles, John, Peter}\}$, is stronger than $\rangle$ is isomorphic to $\langle\{2,4,6\}, |\,\rangle$.



In the left picture the line denotes the relation 'is stronger than' and in the right picture it denotes the relation 'is divisor of'. However, $\langle$ {Charles, John, Peter}, is stronger than $\rangle$ is (under the same supposition as above) NOT isomorphic to $\langle\{1,2,3\},<\rangle$.



2) $f:\mathbb{N}\to\mathbb{N}_{even}$, defined by $f(0)=2, f(1)=0, f(n)=2n$ for $n\geq 2$, is not an isomorphism from $\langle\mathbb{N},<\rangle$ to $\langle\mathbb{N}_{even},<\rangle$, although f is a bijection from $\mathbb{N}$ to $\mathbb{N}_{even}$, because $0<1$, but not $2=f(0)<f(1)=0$.

$$\begin{array}{rl} \mathbb{N} & :\ 0\ 1\ 2\ 3\ 4\ \ldots \\ f\downarrow & \\ \mathbb{N}_{even} & :\ 2\ 0\ 4\ 6\ 8\ \ldots \end{array}$$

Nevertheless, $\langle\mathbb{N},<\rangle$ is isomorphic to $\langle\mathbb{N}_{even},<\rangle$, since there is an isomorphism from $\langle\mathbb{N},<\rangle$ to $\langle\mathbb{N}_{even},<\rangle$, namely $f:\mathbb{N}\to\mathbb{N}_{even}$ defined by $f(n)=2n$ for all $n\in\mathbb{N}$.

$$\begin{array}{rl} \mathbb{N} & :\ 0\ 1\ 2\ 3\ 4\ \ldots \\ f\downarrow & \\ \mathbb{N}_{even} & :\ 0\ 2\ 4\ 6\ 8\ \ldots \end{array}$$

Exercise 3.9. We provide alternative notions of ordered pair:
a) $(v,w) := \{\{v,\emptyset\},\ \{w,\{\emptyset\}\}\}$, and b) $(v,w) := \{\{v,\emptyset\},\{w\}\}$.
Prove that for these definitions it holds that $(v,w)=(x,y)$ iff $v=x$ and $w=y$.

Exercise 3.10. Prove that the operation × (Cartesian Product) is distributive with respect to union and intersection, i.e., $U\times(V\cup W)=(U\times V)\cup(U\times W)$, and $U\times(V\cap W)=(U\times V)\cap(U\times W)$.

Exercise 3.11. Give an example to show that the operation × (Cartesian Product) is not associative, i.e., that not for all sets U, V and W, $U\times(V\times W)=(U\times V)\times W$.

Exercise 3.12. Let $R=\{(0,1),(0,3),(0,4),(2,1),(1,2),(4,7)\}$. Compute Dom($R$), Ran($R$) and $\breve{R}$. Is R a function? Let $S=\{(1,4),(3,2),(5,0)\}$. Compute $R;S$ and $S;R$.

Exercise 3.13. a) Let U be a partition of V and define for $v,w\in V$, $vSw :=$ there is a set W in U such that both $v,w\in W$. Show that S is an equivalence relation on V.
b) Let R be an equivalence relation on V. Then $V/R=\{[v]_R \mid v\in V\}$ is the partition of V belonging to R (see Theorem 3.10). Let S be the equivalence relation belonging to V/R according to a). Prove that R and S are identical.

Exercise 3.14. Check whether each of the following relations on $\mathbb{Z}$ is an equivalence relation or not.
a) $R = \{(x,y) \in \mathbb{Z}^2 \mid x+y < 3\}$ b) $R = \{(x,y) \in \mathbb{Z}^2 \mid x \text{ is a divisor of } y\}$
c) $R = \{(x,y) \in \mathbb{Z}^2 \mid x+y \text{ is even }\}$ d) $R = \{(x,y) \in \mathbb{Z}^2 \mid x = y \text{ or } x = -y\}$

Exercise 3.15. Prove that in each of the following cases $\{V_r \mid r \in \mathbb{R}\}$ is a partition of $\mathbb{R} \times \mathbb{R}$. Describe geometrically the members of this partition. Find the equivalence relations corresponding to the partitions (see Exercise 3.13).
a) $V_r = \{(x,y) \in \mathbb{R}^2 \mid y = x+r\}$, b) $V_r = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = r\}$.
Hint: $y = x+r$ is the equation of a line and $x^2+y^2 = r$ is the equation of a circle.

Exercise 3.16. For each $n \in \mathbb{Z}$ let $V_n = \{m \in \mathbb{Z} \mid \exists q \in \mathbb{Z}\, [\, m = n+5q \,]\}$. Prove that $\{V_n \mid n \in \mathbb{Z}\}$ is a partition of $\mathbb{Z}$.

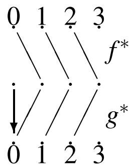
Exercise 3.17. Give an example of a relation, which is transitive and symmetric, but not reflexive.

Exercise 3.18. Spot the flaw in the following argument: Let R be transitive and symmetric. Then xRy and yRz implies xRz for all x, y and z. Also $xRy \to yRx$ holds for all x and y. Now take any x and y such that xRy; then, by the preceding lines, xRx. Hence R is reflexive.

Exercise 3.19. Draw diagrams for the following partially ordered sets:
a) The set of all subsets of a set with 3 elements, partially ordered by $\subseteq$.
b) The set of natural numbers $1,\ldots,25$, partially ordered by divisibility.

Exercise 3.20. Determine which of the following sets are relations, functions, injections, surjections or bijections from $\{1,2,3,4\}$ to $\{1,2,3,4\}$:
a) $R_1 = \{(3,1),(4,2),(4,3),(2,3)\}$, b) $R_2 = \{(2,3),(1,2),(3,2),(4,3)\}$,
c) $R_3 = \{(2,1),(1,2),(4,3),(3,4)\}$, d) $R_1;R_2$ and e) $\breve{R}_3$.

Exercise 3.21. Let $f : U \to V$ and $g : V \to W$. Prove: a) if $g \circ f$ is injective, then f is injective; and b) if $g \circ f$ is surjective, then g is surjective.
Let $f^* : \mathbb{N} \to \mathbb{N}$ be defined by $f^*(n) = n+1$ and let $g^* : \mathbb{N} \to \mathbb{N}$ be defined by $g^*(0) = 0$ and $g^*(n+1) = n$. Prove, using f^* and g^*, that not for all f and g:
c) if $g \circ f$ is injective, then g is injective; d) if $g \circ f$ is surjective, then f is surjective;
e) if $g \circ f$ is bijective, then f or g is bijective.



Exercise 3.22. Let $f : V \to W$. Prove: $\breve{f}$ is a function from W to V iff f is bijective.

Exercise 3.23. $f : U \to V$ and $g : V \to W$. Prove: a) If f and g are injective, then $g \circ f$ is injective; b) If f and g are surjective, then $g \circ f$ is surjective; c) If f and g are bijective, then $g \circ f$ is bijective.

Exercise 3.24. Prove that $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, defined by $f(n,m) := 2^m(2n+1) - 1$, is injective.

Exercise 3.25. Prove: a) $\langle \mathbb{N}, < \rangle$ is not isomorphic to $\langle \mathbb{Z}, < \rangle$, i.e., there is no isomorphism from $\langle \mathbb{N}, < \rangle$ to $\langle \mathbb{Z}, < \rangle$; and b) $\langle \mathbb{Z}, < \rangle$ is not isomorphic to $\langle \mathbb{Q}, < \rangle$.

Exercise 3.26. Prove: $\langle \{2,4,6,12\}, / \rangle$ is isomorphic to $\langle P(\{1,2\}), \subseteq \rangle$.

3.5 The Hilbert Hotel; Denumerable Sets

All sets we experience in daily life are finite. That is why we think that a proper part is smaller than its whole. For instance, $\{2, 3\}$ is a smaller set than $\{1, 2, 3\}$. We shall see that this law for finite sets does not hold anymore for infinite sets.

The numbers $0,1,2,3,\ldots$ are called *natural numbers*. $\mathbb{N} = \{0,1,2,\ldots\}$ is the set of all natural numbers. So, for example, $3 \in \mathbb{N}$, $5 \in \mathbb{N}$ and $1024 \in \mathbb{N}$, while $-3 \notin \mathbb{N}$, $\frac{2}{3} \notin \mathbb{N}$ and $\sqrt{2} \notin \mathbb{N}$. The numbers $\ldots,-3,-2,-1,0,1,2,3,\ldots$ are called *integers*. $\mathbb{Z} = \mathbb{N} \cup \{-1,-2,-3,\ldots\}$ is the set of all integers. Note that each natural number is an integer, but not conversely. Examples: $2 \in \mathbb{Z}$, $-2 \in \mathbb{Z}$, $0 \in \mathbb{Z}$, $3 \in \mathbb{Z}$, $\frac{2}{3} \notin \mathbb{Z}$ and $\sqrt{2} \notin \mathbb{Z}$.

Numbers of the form $\frac{p}{q}$, where $p \in \mathbb{Z}$, $q \in \mathbb{N}$, $q \neq 0$ (and p and q relatively prime) are called *rational numbers*. $\mathbb{Q}^+ = \mathbb{N} \cup \{\frac{1}{1} \cdot \frac{1}{2}, \frac{1}{3}, \ldots\} \cup \{\frac{2}{1}, \frac{2}{2}, \frac{2}{3}, \ldots\} \cup \{\frac{3}{1}, \frac{3}{2}, \frac{3}{3}, \ldots\} \cup \ldots$ is the set of all positive rational numbers. Examples: $\frac{1}{4} \in \mathbb{Q}^+$, $2 \in \mathbb{Q}^+$, $0 \in \mathbb{Q}^+$, $\frac{3}{5} \in \mathbb{Q}^+$, $\sqrt{2} \notin \mathbb{Q}^+$, $\pi \notin \mathbb{Q}^+$. By $\mathbb{Q}$ we mean the set of all positive and negative rational numbers. Note that all integers and hence also all natural numbers are rational.

There are many, many, numbers which are not rational. Already the Greeks knew that $\sqrt{2}$ cannot be written as a quotient of the form $\frac{p}{q}$. The same holds for many other numbers, such as $\sqrt{5}$, $\log 2$, π and Euler's constant e. By $\mathbb{R}$ we mean the set of all *real numbers*. This set contains all natural numbers, all integers and all rational numbers, but also all limits of convergent sequences of rational numbers, such as $\sqrt{2}$, $\log 2$, π and e. For a precise definition of real numbers in terms of sets, see van Dalen, Doets, de Swart [3], section 12.

In this section we shall see that the set $\mathbb{N}$ of all natural numbers is as large as the set $\mathbb{Z}$ of all integers and also as large as the set $\mathbb{Q}$ of all rational numbers. In Section 3.6 we shall see that the set $\mathbb{R}$ of all real numbers is larger than each of the sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$, which are equally large.

From a classical or platonistic point of view the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are *actually infinite*, i.e., the Creator has created these sets, just like the planets, as a completed totality, prior to and independently of any human process of generation and as though they can be spread out completely for our inspection. Mathematicians are like astronomers who try to discover properties of the objects which have been created in its full totality by the Creator.

Since for an intuitionist like L.E.J. Brouwer (1891-1966) mathematical objects are my own mental constructions, from an intuitionistic point of view the infinite is treated only as *potential* or *becoming* or *constructive*, i.e., the set $\mathbb{N}$ of the natural numbers is identified with the construction process for its elements: start with 0 and add 1 to each natural number which has already been constructed before. And it was one of the main achievements of Brouwer to solve the problem how we can talk constructively about the non-denumerable set $\mathbb{R}$ of the real numbers; see Chapter 8.

What does it mean that 'set V has just as many elements as set W'? The proper formulation of this question makes use of the notion of *one-one correspondence* or *matching*. For example, the set {Plato, Augustine, Wittgenstein} has just as many elements as the set {chair 1, chair 2, chair 3}, simply because we can match these sets in a suitable way:

$$\begin{array}{rcl} \text{Plato} & \text{——} & \text{chair 1} \\ \text{Augustine} & \text{——} & \text{chair 2} \\ \text{Wittgenstein} & \text{——} & \text{chair 3} \end{array}$$

From an intuitive point of view a one-one correspondence between two sets V and W is a prescription or function f that associates with every element v in V exactly one element $f(v)$ in W in such a way that conversely for every element w in W there is exactly one v in V with $w = f(v)$. More technically, a one-one correspondence between V and W is a bijective function from V to W; see Section 3.4.

Early scientists were rather puzzled by the effects of the matching-concept. In 1638 Galileo noticed that we can match the set of squares of the positive integers and the set of positive integers itself:

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & \dots & n & \dots \\ \vert & \vert & \vert & \vert & & \vert & \\ 1 & 4 & 9 & 16 & \dots & n^2 & \dots \end{array}$$

This was considered paradoxical, in view of Euclid's proposition that 'the whole is greater than its part' (circa 300 B.C.). However, if one thinks of billiard-balls, being labeled 1, 2, 3, 4, ... on one occasion and the same balls being labeled 1, 4, 9, 16, ... on another occasion, it becomes quite obvious that the sets in question can be matched and hence have as many elements.

This is essentially Gödel's defense that the following definition is *the* natural one for comparing sets in magnitude, also in the case of infinite sets; see Gödel [5], *What is Cantor's continuum problem?*.

Definition 3.40 (Equipollent). V is *equally great as* or *equipollent to* W (V and W are of the *same cardinality*) iff there exists a one-one correspondence between (the elements of) V and (the elements of) W. **Notation**: $V =_1 W$.

One easily may verify the following:

Theorem 3.14. *For all sets U, V and W,*
i) $V =_1 V$; ii) if $V =_1 W$, then $W =_1 V$; iii) if $U =_1 V$ and $V =_1 W$, then $U =_1 W$.

Proof. i) The identity function which associates with every element $x \in V$ this same x, is a one-one correspondence (bijection) between V and V. ii) Let $f : V \to W$ be a one-one correspondence (bijection) between V and W, then the inverse function

$f^{-1}: W \to V$ (see Definition 3.29) is a one-one correspondence between W and V.
iii) Let $f: U \to V$ be a one-one correspondence between U and V and $g: V \to W$ a one-one correspondence between V and W. Then the composition $g \circ f: U \to W$ (see Definition 3.28) is a one-one correspondence between U and W. □

Definition 3.41 (Finite). a) V is *finite* iff there is some natural number $n \in \mathbb{N}$ such that $V =_1 \{x \in \mathbb{N} \mid x < n\}$. b) V is *infinite* iff V is not finite.

Example 3.23. $\{Plato, Augustine, Wittgenstein\} =_1 \{1,2,3\}$; not $\{1,2,3\} =_1 \{1,2\}$.

Example 3.24. $\mathbb{N} =_1 P$, where P is the set of prime numbers. In book IX of Euclid's 'Elements' (300 B.C.) it is shown that there are infinitely many prime numbers. Euclid proceeds by constructing for each finite set of primes a prime which does not belong to it. Using the fact that there are infinitely many primes, we can find a bijection from $\mathbb{N}$ to P by running through $\mathbb{N}$ and checking whether each number is a prime. This is basically the method known as the sieve of Erathostenes.

Theorem 3.15. $\mathbb{N} =_1 \mathbb{N}_{even}$*, where* $\mathbb{N}_{even} := \{x \in \mathbb{N} \mid x \textit{ is even}\}$.

Proof. The correspondence or function f that associates with each natural number n in $\mathbb{N}$ the even natural number $f(n) = 2n$ in $\mathbb{N}_{even}$ is one-one: f associates with each natural number n in $\mathbb{N}$ exactly one even natural number in $\mathbb{N}_{even}$, namely $f(n) = 2n$, in such a way that conversely for every even natural number $m = 2n$ in $\mathbb{N}_{even}$ there is exactly one natural number n in $\mathbb{N}$, namely $n = \frac{m}{2}$, such that $f(n) = m$.

$$\begin{array}{lccccccccc}
\mathbb{N}: & 0 & 1 & 2 & 3 & 4 & \ldots & n & \ldots \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & & \downarrow & \\
\mathbb{N}_{even} & 0 & 2 & 4 & 6 & 8 & \ldots & 2n & \ldots
\end{array}$$

□

Hence, the proposition 'the whole is greater than its part' (Euclid) turns out to be false for infinite sets: $\mathbb{N}_{even}$ is a proper subset of $\mathbb{N}$, but $\mathbb{N}_{even}$ is still equipollent to $\mathbb{N}$. However, it is easy to see that 'a proper part is smaller than the whole' is true for finite sets.

$\mathbb{N} =_1 V$ means that there is a one-one correspondence v between the sets $\mathbb{N}$ and V:

$$\begin{array}{lccccccc}
\mathbb{N}: & 0 & 1 & 2 & 3 & 4 & 5 & \ldots \\
V: & v(0) & v(1) & v(2) & v(3) & v(4) & v(5) & \ldots
\end{array}$$

If V is equipollent to $\mathbb{N}$, we say that V is *denumerable*: $V = \{v(0), v(1), v(2), \ldots\}$.

Definition 3.42 (Denumerable; Enumerable; Countable).
V is *denumerable* $:= \mathbb{N} =_1 V$.
V is *enumerable* or *countable* $:= V$ is finite or denumerable.
A one-one correspondence v between $\{0, 1, \ldots, n\}$ or $\mathbb{N}$ respectively and V is called an *enumeration* of V.

Remark 3.1. The usage of the terminology is not firmly established. Instead of 'denumerable' some authors use *countably infinite*.

Suppose somewhere in heaven is a hotel, called the *Hilbert hotel*, after the German mathematician and philosopher David Hilbert (1862 – 1943), with as many rooms as there are natural numbers. We also suppose that in every room there is exactly one guest: g_0, g_1, g_2, g_3, ...

room:	0	1	2	3	4	5	...
	g_0	g_1	g_2	g_3	g_4	g_5	

So, the Hilbert hotel is full in the sense that there is a one-one correspondence g between the set of room numbers $\{0,1,2,\ldots\}$ and the set $\{g_0,g_1,g_2,\ldots\}$ of guests.

At a certain day two new guests, g_{-1} and g_{-2}, arrive at the reception and both ask for a private room; neither the two new guests nor the existing guests want to share a room with somebody else. The receptionist, who studied mathematics and philosophy, had to think a little while, but found an easy solution: let all the existing guests move two rooms; then the first two rooms are becoming free and can be given to the two new guests. The result is the following room assignment:

room:	0	1	2	3	4	5	...
	g_{-2}	g_{-1}	g_0	g_1	g_2	g_3	

We see that the two sets $\{g_0,g_1,g_2,\ldots\}$ and $\{g_{-2},g_{-1},g_0,g_1,g_2,\ldots\}$ are equally large: the number of rooms did not change. We also see that $\mathbb{N} = \{0,1,2,\ldots\}$ is as large as $\{-2,-1\} \cup \mathbb{N}$, in other words: there is a one-one correspondence f between these two sets: $f(0) = -2$, $f(1) = -1$ and $f(n+2) = n$ for $n \geq 0$.

Theorem 3.16. *a)* $\mathbb{N} =_1 \{-2,-1\} \cup \mathbb{N}$, *in other words,* $\{-2,-1\} \cup \mathbb{N}$ *is denumerable. b) More generally: if* W *is a finite set* $\{w_0,\ldots,w_{k-1}\}$, $k \geq 1$, *and* V *is a denumerable set, then* $W \cup V$ *is denumerable.*

Proof. a) The function f from $\mathbb{N}$ to $\{-2,-1\} \cup \mathbb{N}$, defined by $f(0) = -2$, $f(1) = -1$, and $f(n+2) = n$ for $n \geq 0$, is a one-one correspondence between these two sets: f assigns to every $n \in \mathbb{N}$ exactly one element in $\{-2,-1\} \cup \mathbb{N}$, such that conversely for every element m in $\{-2,-1\} \cup \mathbb{N}$ there is exactly one $n \in \mathbb{N}$, namely $n = m+2$, with $m = f(n)$.

b) Suppose $W = \{w_0,\ldots,w_{k-1}\}$, $k \geq 1$, and V is denumerable, i.e., there is a one-one correspondence v between $\mathbb{N}$ and V. Hence, $V = \{v_0,v_1,v_2,\ldots\}$. Then the function f from $\mathbb{N}$ to $W \cup V$, defined by $f(0) = w_0$, ..., $f(k-1) = w_{k-1}$, and for each $n \in \mathbb{N}$, $f(n+k) = v_n$, is a one-one correspondence between $\mathbb{N}$ and $W \cup V$. □

So far, so good! But at a certain day all denumerably many guests, except g_0, of the Hilbert hotel want to invite a personal friend and to give him or her a private room; again nobody is willing to share his room with somebody else. Say guest g_i wants to invite g_{-i} for each $i \geq 1$. This situation is pictured in the following schema:

room:	0	1	2	3	4	5	…
	g_0	g_1	g_2	g_3	g_4	g_5	
		g_{-1}	g_{-2}	g_{-3}	g_{-4}	g_{-5}	⋯

The receptionist looks concerned; she could host finitely many new guests, but now she is asked to host countably many new guests, each wanting a separate room. But ... after some thinking she found a solution: let all old guests move to the room with number twice the old room number; by doing that all rooms with an odd number become empty and the new guests can be hosted in these odd numbered rooms. So, the new room assignment looks as follows:

room:	0	1	2	3	4	5	…
	g_0	g_{-1}	g_1	g_{-2}	g_2	g_{-3}	

The receptionist is proud, the guests are happy and the Hilbert hotel is doing good business. Guest g_0 can stay in room number 0, guest g_1 moves to room number 2, guest g_2 moves to room number 4, guest g_3 moves to room number 6, etc. By doing this the rooms 1, 3, 5, … with an odd number become available and the new guests $g_{-1}, g_{-2}, g_{-3}, \ldots$ can occupy these rooms.

We see that the set $\mathbb{N}$ is as large as the set $\{-1,-2,-3,\ldots\} \cup \mathbb{N}$, this is $\mathbb{Z}$, while at the same time $\mathbb{N}$ is a proper subset of $\mathbb{Z}$.

Theorem 3.17. *a)* $\mathbb{N} =_1 \mathbb{N} \cup \{-1,-2,-3,\ldots\} = \mathbb{Z}$*; i.e.,* $\mathbb{Z}$ *is denumerable.*
b) More generally: if V and W are denumerable, then also $V \cup W$ is denumerable.

Proof. a) With the even natural numbers 0, 2, 4, … in $\mathbb{N}$ we can associate respectively the numbers 0, 1, 2, … in $\mathbb{Z}$ and with the odd natural numbers 1, 3, 5, … in $\mathbb{N}$ we can associate respectively the numbers $-1,-2,-3,\ldots$ in $\mathbb{Z}$. More precisely, the function f from $\mathbb{N}$ to $\mathbb{Z}$, defined by $f(2n) = n$ and $f(2n-1) = -n$ is a one-one correspondence between $\mathbb{N}$ and $\mathbb{Z}$.
b) Suppose V and W are denumerable, i.e., there are one-one correspondences v and w between $\mathbb{N}$ and V, respectively W. Hence, $V = \{v(0), v(1), v(2), \ldots\}$ and $W = \{w(0), w(1), w(2), \ldots\}$. Then $v(0), w(0), v(1), w(1), v(2), w(2), \ldots$ is an enumeration of $V \cup W$. More precisely, the function f from $\mathbb{N}$ to $V \cup W$, defined by $f(2n) = v(n)$ and for $n \geq 1$, $f(2n-1) = w(n-1)$ is a one-one correspondence between $\mathbb{N}$ and $V \cup W$. □

So far the Hilbert hotel had overcome all difficulties. The real problem started only the next day, when each of the denumerably many guests announced that he or she wants to accommodate denumerably many friends. Each guest g_i, $i \in \mathbb{N}$, wants to invite denumerably many friends $g_{i1}, g_{i2}, g_{i3}, \ldots$. How should the receptionist provide everybody with a private room?

room:	0	1	2	3	4	5	…
	g_0	g_1	g_2	g_3	g_4	g_5	
	g_{01}	g_{11}	g_{21}	g_{31}	g_{41}	g_{51}	$\cdots$
	g_{02}	g_{12}	g_{22}	g_{32}	g_{42}	g_{52}	$\cdots$
	g_{03}	g_{13}	g_{23}	g_{33}	g_{43}	g_{53}	$\cdots$
	⋮	⋮	⋮	⋮	⋮	⋮	…

Although the first thought of the receptionist was that this may be impossible, after thinking approximately fifteen minutes she found a solution. For convenience, she identifies g_0 with g_{00}, $g_!$ with g_{10}, g_2 with g_{20}, etc. Let $V_0 = \{g_{00}, g_{01}, g_{02}, \ldots\}$, $V_1 = \{g_{10}, g_{11}, g_{12}, \ldots\}$, $V_2 = \{g_{20}, g_{21}, g_{22}, \ldots\}$, etc. Then the diagram below at the left hand side shows all the guests who have to be accommodated in a private room, i.e., $V_0 \cup V_1 \cup V_2 \cup \ldots$. Making a systematic 'walk' through the schema of guests, as indicated in the diagram below at the right hand side, gives an enumeration of all the guests in $V_0 \cup V_1 \cup V_2 \cup \ldots$. The receptionist assigns to guest g_{ij}, the jth friend of g_i, room number $\frac{1}{2}(i+j)(i+j+1)+j$. So, guest $g_0 = g_{00}$ gets room 0, guest $g_1 = g_{10}$ gets room 1, guest g_{01} gets room 2, guest $g_2 = g_{20}$ gets room 3, guest g_{11} gets room 4, guest g_{02} gets room 5, guest $g_3 = g_{30}$ gets room 6, guest g_{21} gets room 7, guest g_{12} gets room 8, guest g_{03} gets room 9, guest $g_4 = g_{40}$ gets room 10, etc.

More precisely, the function f from $V_0 \cup V_1 \cup V_2 \cup \ldots$ to $\mathbb{N}$, defined by $f(g_{ij}) = \frac{1}{2}(i+j)(i+j+1)+j$, is a one-one correspondence between the two sets in question: f assigns to every g_{ij} exactly one natural (room) number $f(g_{ij})$, such that conversely for every natural number n in $\mathbb{N}$ there is exactly one guest g_{ij} with $n = f(g_{ij})$.

V_0	V_1	V_2	V_3	V_4	$\ldots$
g_{00}	g_{10}	g_{20}	g_{30}	g_{40}	$\cdots$
g_{01}	g_{11}	g_{21}	g_{31}	g_{41}	$\cdots$
g_{02}	g_{12}	g_{22}	g_{32}	g_{42}	$\cdots$
g_{03}	g_{13}	g_{23}	g_{33}	g_{43}	$\cdots$
⋮	⋮	⋮	⋮	⋮	

0	1	3	6	10	…
2	4	7	11		…
5	8	12			…
9	13				…
⋮	⋮	⋮	⋮	⋮	

Theorem 3.18. *a) The union of denumerably many denumerable sets $V_0, V_!, V_2, \ldots$ is denumerable.*
b) The set $\mathbb{Q}^+$ of all rational numbers greater or equal than 0 is denumerable.

Proof. a) Let $V_0 = \{v_{00}, v_{01}, v_{02}, \ldots\}$, $V_1 = \{v_{10}, v_{11}, v_{12}, \ldots\}$, $V_2 = \{v_{20}, v_{21}, v_{22}, \ldots\}$, etc. be denumerably many denumerable sets. Then the function f from $V_0 \cup V_1 \cup V_2 \cup$

... to $\mathbb{N}$, defined by $f(v_{ij}) = \frac{1}{2}(i+j)(i+j+1)+j$, is a one-one correspondence between the two sets in question: f assigns to every v_{ij} exactly one natural number $f(v_{ij}) \in \mathbb{N}$, in such a way that conversely for every natural number $n \in \mathbb{N}$ there is exactly one $v_{ij} \in V_0 \cup V_1 \cup V_2 \cup \ldots$ with $f(v_{ij}) = n$.
b) Identifying g_{ij} with the rational number $\frac{i}{j}$, leaving out g_{i0} for all $i \in \mathbb{N}$ and taking away all double occurrences of the same rational number, such as $\frac{1}{2} = \frac{2}{4} = \frac{3}{6} = \ldots$, we obtain an enumeration of all rational numbers $\frac{i}{j} \geq 0$ with $i, j \in \mathbb{N}$ and $j > 0$. □

Corollary 3.5. $\mathbb{Q}$ *is denumerable.*

Proof. $\mathbb{Q} = \mathbb{Q}^+ \cup \mathbb{Q}^-$, where $\mathbb{Q}^- = \{x \in \mathbb{Q} \mid x < 0\}$. According to Theorem 3.18 $\mathbb{Q}^+$ is denumerable. In the same way one may prove that $\mathbb{Q}^-$ is denumerable. And by Theorem 3.17 the union of two denumerable sets is again denumerable. □

Exercise 3.27. a) Prove that a) $\mathbb{Z} =_1 \mathbb{N}_{even}$; b) $\mathbb{N}_{even} =_1 \mathbb{N}_{odd}$, where $\mathbb{N}_{even} = \{x \in \mathbb{N} \mid x \text{ is even}\}$ and $\mathbb{N}_{odd} = \{x \in \mathbb{N} \mid x \text{ is odd}\}$.

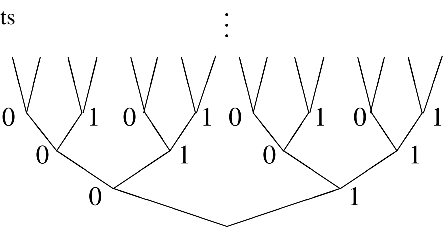
Exercise 3.28. a) Prove that the set $\{0,1\}^*$ of all finite sequences of 0's and 1's is denumerable.
b) Let Σ be an alphabet, i.e., a finite set of symbols. And let Σ^* be the set of all words over Σ, i.e., Σ^* is the set of all finite sequences of elements of Σ. Prove that Σ^* is denumerable. Hint: Note that a) is a special case of b) by taking $\Sigma = \{0,1\}$.
c) Conclude that the set of all expressions in English is denumerable.

Exercise 3.29. Let V be a enumerable set. Prove that the set V^* of all finite sequences of elements of V is denumerable.

3.6 Non-enumerable Sets

In Section 3.5 we have seen that the infinite sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ have the same cardinality, i.e., are equally large, although clearly $\mathbb{N}$ is a proper subset of $\mathbb{Z}$ and $\mathbb{Z}$ is a proper subset of $\mathbb{Q}$. One might be inclined to think that all infinite sets are equally large. Nothing is less true! We shall see in this section that there are many sets which are larger than the sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$. But first we have to explain what we mean by 'being larger than'. The natural definition of $V <_1 W$, V is smaller than W, is:
1. V may be embedded into W, i.e., there is a function f from V to W such that for all $x, y \in V$, if $x \neq y$, then $f(x) \neq f(y)$ (f is injective), but
2. there is no one-one correspondence between the (elements of) V and W. More precisely, for every function f from V to W there will be at least one element $w \in W$ such that there is no $v \in V$ with $f(v) = w$. In other words, there is no surjection $f : V \to W$ and hence there cannot be a bijection $f : V \to W$.

A first example of a non-enumerable set is the set $\{0,1\}^{\mathbb{N}}$, i.e., the set of all functions $f : \mathbb{N} \to \{0,1\}$. Since a function $f : \mathbb{N} \to \{0,1\}$ may be identified with the infinite sequence $f(0), f(1), f(2), \ldots$ of zero's and one's, the set $\{0,1\}^{\mathbb{N}}$ is also called the set of all infinite sequences of zero's and one's.



It is easy to see that $\mathbb{N}$ can be embedded into $\{0,1\}^{\mathbb{N}}$: let $F : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ be defined by $F(n)$ = the infinite sequence of zero's and one's with $F(n)(i) = 0$ for $i \neq n$ and $F(n)(n) = 1$, i.e., $F(n)$ is the sequence of zero's and one's with a 1 only at the n^{th} place. Evidently, an infinite sequence with two or more one's does not belong to the range of F. In Theorem 3.19 we shall prove that any function $F : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ will 'forget' some elements of $\{0,1\}^{\mathbb{N}}$, more precisely, that for any such function F there is an infinite sequence s of zero's and one's such that $s \neq F(i)$ for all $i \in \mathbb{N}$.

Definition 3.43 (Smaller than). V is smaller than W :=
a) There exists an embedding from V into W, i.e., there exists a function $f : V \to W$ such that for all $x, y \in V$, if $x \neq y$, then $f(x) \neq f(y)$ (f is injective),
b) but there is no surjection, and hence no bijection or one-one correspondence, $f : V \to W$. **Notation**: $V <_1 W$,

Theorem 3.19. $\mathbb{N} <_1 \{0,1\}^{\mathbb{N}}$.

Proof. a) There is an injection $F : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$; for instance, the function F with $F(n)(i) = 0$ for all $i \neq n$ and $F(n)(n) = 1$.
b) We show that each $F : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ is not surjective, in other words, that for each such function F there is an infinite sequence s in $\{0,1\}^{\mathbb{N}}$ such that $s \neq F(i)$ for all $i \in \mathbb{N}$. So, let $F : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$. Then for all $i \in \mathbb{N}$, $F(i)$ is an infinite sequence of zero's and one's. The sequences $F(i)$ may be represented in, for instance, the following diagram:

$$\begin{array}{rcccccc}
 & & 0 & 1 & 2 & 3 & \dots \\
 & & & & & & \\
F(0) & = & 0 & 1 & 0 & 0 & \dots \\
F(1) & = & 1 & 0 & 1 & 0 & \dots \\
F(2) & = & 0 & 0 & 1 & 1 & \dots \\
F(3) & = & 1 & 1 & 0 & 0 & \dots \\
\vdots & & & & & & \\
s & = & 1 & 1 & 0 & 1 & \dots
\end{array}$$

Construct the infinite sequence s by interchanging the zero's and one's at the diagonal $F(0)(0)$, $F(1)(1)$, $F(2)(2)$, $F(3)(3), \ldots$, i.e., define $s(i) := 1 - F(i)(i)$. Then for all $i \in \mathbb{N}$, s differs from $F(i)$ at place i, in other words $s(i) \neq F(i)(i)$. So, $s \neq F(i)$ for all $i \in \mathbb{N}$ and therefore $F : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ is not surjective. □

Remark 3.2. The method used in the proof of Theorem 3.19 is called *diagonalisation* or the *diagonal method* of Cantor.

Next we will show that $P(\mathbb{N}) =_1 \{0,1\}^{\mathbb{N}}$, from which it follows by Theorem 3.19 that $\mathbb{N} <_1 P(N)$.

Theorem 3.20. *For any set V, $P(V) =_1 \{0,1\}^V$.*

Proof. The function $K : P(V) \to \{0,1\}^V$, defined by $K(U) = K_U$, where K_U is the characteristic function of U, is a bijection from $P(V)$ to $\{0,1\}^V$. First we show that K is injective. So, suppose $U_1 \neq U_2$, say $v \in U_1$ and $v \notin U_2$. Then $K_{U_1}(v) = 1$ and $K_{U_2}(v) = 0$ and therefore $K(U_1) \neq K(U_2)$. To show that K is surjective, suppose $f \in \{0,1\}^V$. Taking $U := \{v \in V \mid f(v) = 1\}$, it follows that $f = K(U)$, since $f(v) = 1$ iff $v \in U$, i.e., iff $K_U(v) = 1$. □

Now that we have established that $P(\{1,\ldots,n\}) =_1 \{0,1\}^{\{1,\ldots,n\}}$ we can use Theorem 3.11 to determine the number of elements of $P(\{1,\ldots,n\})$, namely 2^n, the number of elements of $\{0,1\}^{\{1,\ldots,n\}}$. So, we know that for finite sets V the power set of V is much larger than the set V itself. A similar proposition, Theorem 3.21, holds for infinite sets, only we cannot expect to prove it by just counting. Cantor provided us with a revolutionary technique for this purpose: *diagonalisation*.

From Theorem 3.19 and Theorem 3.20 follows Cantor's theorem:

Corollary 3.6 (Cantor's Theorem). $\mathbb{N} <_1 P(\mathbb{N})$. *So, there are more subsets of* $\mathbb{N}$ *than there are natural numbers.*

Proof. By Theorem 3.19 there is an injection $f : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$. By Theorem 3.20 there is a bijection $g : P(\mathbb{N}) \to \{0,1\}^{\mathbb{N}}$. Then $g^{-1} \circ f : \mathbb{N} \to P(\mathbb{N})$ is an injection. Suppose there were a surjection $f : \mathbb{N} \to P(\mathbb{N})$. Then $g \circ f : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ would be a surjection (see Exercise 3.23), which contradicts Theorem 3.19. □

More generally, we shall prove that any set V is smaller than its powerset $P(V)$. It is easy to see that any set V can be embedded into its powerset $P(V)$: with every element $v \in V$ corresponds the set $\{v\} \in P(V)$, more precisely, the function f from V to $P(V)$, defined by $f(v) = \{v\}$, assigns to different elements in V different elements of $P(V)$. Clearly, this function is not a one-one correspondence between V and $P(V)$: for instance, if W is a subset of V with two or more elements, then there is no $v \in V$ such that $f(v) = W$. Even stronger, below we shall show that there cannot exist a one-one correspondence between V and $P(V)$. Consequently, for any set V, the powerset $P(V)$ of V is larger than V itself. Note that we already verified this for finite sets, see Theorem 3.6.

Theorem 3.21. *For any set V, $V <_1 P(V)$.*

Proof. Clearly, the function f from V to $P(V)$, defined by $f(v) = \{v\}$, assigns to different elements of V different elements of $P(V)$. So, f embeds V into $P(V)$. Next we have to show that there cannot exist a one-one correspondence between V and $P(V)$. So, suppose g is any function from V to $P(V)$. Then we have to show that there is a set $W \in P(V)$, i.e., $W \subseteq V$, such that for no $v \in V$, $W = g(v)$. Take $W = \{v \in V \mid v \notin g(v)\}$. Then indeed, there is no $v \in V$ such that $W = g(v)$.

For suppose for some $v_0 \in V$, $W = \{v \in V \mid v \notin g(v)\} = g(v_0)$. Then for all x, $x \in W$ iff $x \in g(v_0)$, i.e., for all $x \in V$, $x \notin g(x)$ iff $x \in g(v_0)$. In particular, taking $x = v_0$, $v_0 \notin g(v_0)$ iff $v_0 \in g(v_0)$. Contradiction. □

The preceding theorem is an eye-opener: it says in particular that

$$\mathbb{N} <_1 P(\mathbb{N}) <_1 P(P(\mathbb{N})) <_1 P(P(P(\mathbb{N}))) <_1 \ldots$$

So, there are many degrees of infinity: the degree of infinity of $\mathbb{N}$ is smaller than the one of $P(\mathbb{N})$, which in its turn is smaller than the one of $P(P(\mathbb{N}))$, etc.

Definition 3.44 (Interval). For $a,b \in \mathbb{R}$ let $[a,b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$; $(a,b) := \{x \in \mathbb{R} \mid a < x < b\}$; $[a,b) := \{x \in \mathbb{R} \mid a \leq x < b\}$; and $(a,b] := \{x \in \mathbb{R} \mid a < x \leq b\}$. $[a,b]$ is called the *closed* (at both sides) interval between a and b, while (a,b) is called the *open* (at both sides) interval between a and b.

Next we will prove that $\mathbb{N}$ is not only smaller than $P(\mathbb{N})$, but also smaller than [0, 1], the set of all real numbers between 0 and 1. We will present a direct proof here for historical reasons. The proof below is Poincaré's proof. The first direct proof was presented by Cantor.

Theorem 3.22. $\mathbb{N} <_1$ [0, 1]. *So, there are more real numbers between 0 and 1 than there are natural numbers.*

Proof. It is easy to construct an embedding from $\mathbb{N}$ into [0, 1]. For instance, $f : \mathbb{N} \to$ [0, 1], defined by $f(0) = 0$ and $f(n) = \frac{1}{n}$ for $n \geq 1$, is an injection. Next we have to show that there cannot exist a surjection $g : \mathbb{N} \to$ [0, 1]. To do so, we shall prove that for any function $g : \mathbb{N} \to$ [0, 1] we can construct a real number b between 0 and 1 such that $b \neq g(n)$ for any $n \in \mathbb{N}$. So, let $g : \mathbb{N} \to$ [0, 1]. Given this g, we can construct a chain $S_0, S_1, S_2, \ldots$ of segments (in $\mathbb{Q}$), where each segment is contained in the preceding one and the length of the segments is decreasing to 0, such that for every $n \in \mathbb{N}$, $g(n)$ is not an element of S_n.

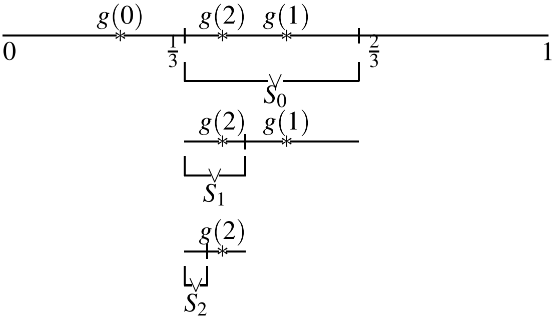
Note that [0, 1] = [0, $\frac{1}{3}$] ∪ [$\frac{1}{3}$, $\frac{2}{3}$] ∪ [$\frac{2}{3}$, 1]. At least one of those three subsets does not contain $g(0)$, say S_0.

Suppose $S_0, \ldots, S_n$ have already been defined, such that
1. for all i, $0 \leq i \leq n$, $g(i)$ is not an element of S_i,
2. for all i, $0 \leq i < n$, $S_{i+1} \subseteq S_i$, and
3. for all i, $0 \leq i \leq n$, the length of S_i equals 3^{-i-1}.

Let S_n = $[p_n, q_n]$. Now S_n is the union of

$$[p_n, \tfrac{2p_n+q_n}{3}],\ [\tfrac{2p_n+q_n}{3}, \tfrac{p_n+2q_n}{3}] \text{ and } [\tfrac{p_n+2q_n}{3}, q_n].$$

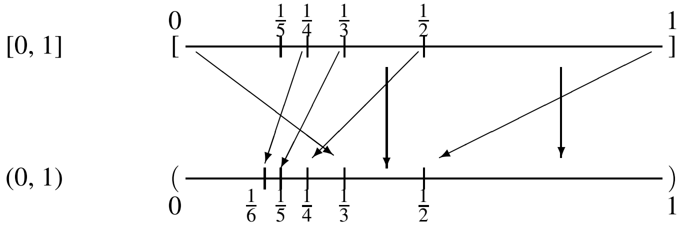
At least one of those three subsets of S_n does not contain $g(n+1)$, say S_{n+1}. This chain of segments $S_0, S_1, S_2, \ldots$ determines a real number b (which in general will not be a rational number), such that for every $n \in \mathbb{N}$, b occurs in S_n, and hence, $b \in$ [0, 1]. Now for every $n \in \mathbb{N}$, $g(n)$ does not occur in S_n, while b does occur in S_n. Hence, for every $n \in \mathbb{N}$, $b \neq g(n)$. □



Theorem 3.22 tells us that [0, 1] is not enumerable, more precisely, for each enumeration $g : \mathbb{N} \to$ [0, 1] of elements of [0, 1], a real number b (between 0 and 1) can be constructed such that b does not occur in that enumeration, i.e., for all $n \in \mathbb{N}$, $b \neq g(n)$. On the other hand, we can define only countably many individual real numbers (between 0 and 1). This restriction is inherent to our language. Next we are going to show that [0, 1] is equipollent to $\mathbb{R}$ and consequently, by Theorem 3.22, that $\mathbb{N} <_1 \mathbb{R}$. In order to do so, we first show:

Theorem 3.23. [0, 1] $=_1$ (0, 1).

Proof. Consider the following denumerable subset of [0, 1]: $\{1,0,\frac{1}{2},\frac{1}{3},\frac{1}{4},\ldots\}$. Now let f : [0, 1] $\to$ (0, 1) be defined as follows:



$f(1) = \frac{1}{2}$, $f(\frac{1}{n}) = \frac{1}{n+2}$ if $n \geq 2$,
$f(0) = \frac{1}{3}$, $f(x) = x$ if $x \notin \{1,0,\frac{1}{2},\frac{1}{3},\frac{1}{4},\ldots\}$.
Clearly, f is a bijection from [0, 1] to (0, 1). Therefore [0, 1] $=_1$ (0, 1). □

In the proof of Theorem 3.23 we have used the fact that the uncountable sets [0, 1] and (0, 1) have an denumerable subset, $\{1,0,\frac{1}{2},\frac{1}{3},\frac{1}{4},\ldots\}$ and $\{\frac{1}{2},\frac{1}{3},\frac{1}{4},\ldots\}$ respectively. More generally, one can show:

Theorem 3.24. *If V contains a denumerable subset, then there is a proper subset of V, which is equipollent to V. Hence, Euclid's axiom 'the whole is greater than its proper part' does not hold for such sets V.*

Proof. Let $\{x_0, x_1, x_2, \ldots\}$ be a denumerable subset of V. Then $V - \{x_0\}$ is a proper subset of V, which is equipollent to V. For the function $g : V \to V - \{x_0\}$, defined by $g(x) = x$ if $x \notin \{x_0, x_1, x_2, \ldots\}$, and $g(x_i) = x_{i+1}$ for all $i \in \mathbb{N}$, is a bijection from V to $V - \{x_0\}$. □

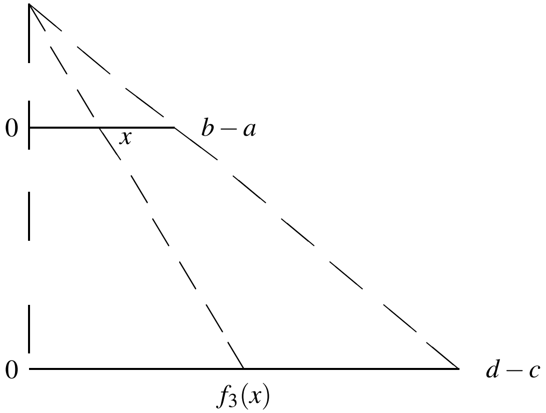
By using an argument similar to the proof of Theorem 3.23 (see Exercise 3.30) one can show:

Theorem 3.25. *For $a, b \in \mathbb{R}$, $[a,b] =_1 (a,b] =_1 [a,b) =_1 (a,b)$.*

Amazingly, the length of an interval of real numbers does not change the cardinality (number of elements) of the interval. Compare an interval of real numbers with an elastic. By stretching the elastic out, its length becomes larger, but the number of points in the elastic does not change.

Theorem 3.26. *For $a, b, c, d \in \mathbb{R}$ with $a < b$ and $c < d$, $(a,b) =_1 (c,d)$.*

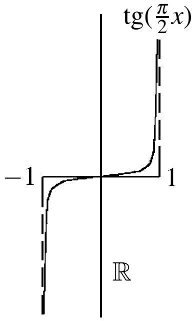
Proof. First we translate a and c to 0.
$f_1 : (a,b) \to (0, b-a)$, defined by $f_1(x) = x - a$, is a bijection from (a,b) to $(0, b-a)$.
$f_2 : (c,d) \to (0, d-c)$, defined by $f_2(x) = x - c$, is a bijection from (c,d) to $(0, d-c)$.
Next we stretch (or shrink) $(0, b-a)$: $f_3 : (0, b-a) \to (0, d-c)$, defined by $f_3(x) = \frac{d-c}{b-a}x$, is a bijection from $(0, b-a)$ to $(0, d-c)$. Then $f_2^{-1} \circ f_3 \circ f_1 : (a,b) \to (c,d)$ is a bijection from (a,b) to (c,d). □



Next we show that any interval (of finite length) of real numbers is equipollent to the set $\mathbb{R}$ of all real numbers.

Theorem 3.27. $(-1,1) =_1 \mathbb{R}$.

Proof. $f : (-1,1) \to \mathbb{R}$, defined by $f(x) = tg(\frac{\pi}{2}x)$, is a bijection from $(-1,1)$ to $\mathbb{R}$. □



$(-1,1)$ is again a proper subset of $\mathbb{R}$, which is equipollent to $\mathbb{R}$. So, there are as many real numbers between -1 and 1 as there are real numbers on a straight line.

By Theorem 3.23, [0, 1] $=_1$ (0, 1), by Theorem 3.26, (0, 1) $=_1$ (-1, 1) and by Theorem 3.27, (-1, 1) $=_1$ $\mathbb{R}$. Hence, [0, 1] $=_1$ $\mathbb{R}$. Since, according to Theorem 3.22, $\mathbb{N} <_1$ [0, 1], it follows that:

Theorem 3.28. $\mathbb{N} <_1 \mathbb{R}$

By Theorem 3.19 and Theorem 3.22 we know that $\{0,1\}^{\mathbb{N}}$, i.e., the set of all infinite sequences of zero's and one's, and [0, 1], i.e., the set of all real numbers between 0 and 1, each are larger than $\mathbb{N}$. But how do the cardinalities of these two sets compare; in other words, is one larger than the other or are they equipollent?

It is known that each real number in (0, 1] has a unique non-terminating decimal extension. For instance, 1 = 0.999. . ., $\frac{1}{3}$ = 0.333. . . and 0.5 = 0.4999. . .. Hence, (0, 1] $=_1$ $\{0,1,\ldots,9\}^{\mathbb{N}}$. One can also show that $\{0,1,\ldots,9\}^{\mathbb{N}} =_1 \{0,1\}^{\mathbb{N}}$ (see [3], section 18). Hence, it follows that:

Theorem 3.29. $\mathbb{R} =_1 (0,1] =_1 \{0,1,\ldots,9\}^{\mathbb{N}} =_1 \{0,1\}^{\mathbb{N}} =_1 P(\mathbb{N})$.

Traditionally $\mathbb{R}$ is called the *continuum*. The term is, however, also metaphorically used for $\{0,1\}^{\mathbb{N}}$, which is usually written as $2^{\mathbb{N}}$.

Summarizing: The sets in each column below are equipollent and are strictly smaller than any set in a column to the right of it.

$$\begin{array}{lllll}
\mathbb{N} & \{0,1\}^{\mathbb{N}} & PP(\mathbb{N}) & PPP(\mathbb{N}) & \ldots \\
\mathbb{Z} & P(\mathbb{N}) & P(\mathbb{R}) & PP(\mathbb{R}) & \\
\mathbb{Q} & (a,b) & & & \\
 & \mathbb{R} & & &
\end{array}$$

One may say that there are infinitely many degrees of infinity. As far as our limited experience goes, it turns out that (leaving aside larger sets, such as $PP(\mathbb{N}), PPP(\mathbb{N})$) most familiar infinite sets are either denumerable or equipollent to the continuum. A natural question to ask is whether there are sets which are larger than $\mathbb{N}$ and smaller than $\mathbb{R}$.

Cantor conjectured in 1878 that each infinite subset of $\mathbb{R}$ is either denumerable or equipollent to the continuum. This conjecture is known as the *Continuum Hypothesis* (CH). A precise formulation reads:

Cantor's Continuum Hypothesis: there is no set $V \subseteq \mathbb{R}$ such that $\mathbb{N} <_1 V <_1 \mathbb{R}$.

So far the continuum hypothesis has withstood all attempts to settle it. From the work of Gödel (1938) and Cohen (1963) we know that the Continuum Hypothesis is consistent with, but at the same time independent of, the basic axioms of set theory (such as given by Zermelo and Fraenkel). The matter of its truth or falsity in the intended universe of set theory however remains unsettled. Gödel, in his paper *What is Cantor's Continuum Problem?* in [5] has analysed the evidence, which turns out to be rather in favour of a rejection.

Exercise 3.30. Prove that (0, 1] $=_1$ [0, 1], (0, 1] $=_1$ (0, 1) and (0, 1] $=_1$ [0, 1).

Exercise 3.31. Let Σ be an alphabet, i.e., a finite set of symbols. L is a language over $\Sigma := L \subseteq \Sigma^*$, where Σ^* is the set of all finite sequences of elements of Σ. Prove that the set of languages over Σ is uncountable.

Exercise 3.32. Prove that [0, 1] $=_1$ [0, 2] and that (0, 1) $=_1$ (0, 3).

Exercise 3.33. V is *Dedekind infinite* := there is an injective function with domain V and whose range is a proper subset of V. Prove: V is infinite iff V is Dedekind infinite. Hint: If V is infinite, then by the axiom of choice V has an denumerable subset. Next use Theorem 3.24. For the axiom of choice see van Dalen, e.a. [3].

3.7 Solutions

Solution 3.1.

$\mathbb{N} \notin \mathbb{N}$
$\mathbb{N} \in \{\mathbb{N}\}$
$\mathbb{N} \subseteq \mathbb{N}$
$\mathbb{N} \notin \{\{\mathbb{N}\}\}$
$\mathbb{N} \not\subseteq \{\mathbb{N}\}$
$\{1,2\} \notin \mathbb{N}$
$\{1,2\} \subseteq \mathbb{N}$
$\{1,2\} \notin \{\mathbb{N}\}$

$\{2,3\} \not\subseteq \{\mathbb{N}\}$
$\{2\} \not\subseteq \{\mathbb{N}\}$
$\{2\} \subseteq \mathbb{N}$
$2 \notin \{1,\{2\},3\}$
$\{2\} \in \{1,\{2\},3\}$
$\{1,\{2\}\} \not\subseteq \{1,\{2,3\}\}$
$\{1,\{2\}\} \subseteq \{1,\{2\},3\}$
$\{-2,2\} \not\subseteq \mathbb{N}$

$\emptyset \notin \emptyset$
$\emptyset \subseteq \emptyset$
$\emptyset \in \{\emptyset\}$
$\emptyset \subseteq \{\emptyset\}$
$\emptyset \notin \{\{\emptyset\}\}$
$\emptyset \subseteq \{\{\emptyset\}\}$
$\{\emptyset\} \in \{\{\emptyset\}\}$
$\{\emptyset\} \not\subseteq \{\{\emptyset\}\}$

$\{\emptyset\} \notin \emptyset$
$\{\emptyset\} \not\subseteq \emptyset$
$\{\emptyset\} \subseteq \{\emptyset\}$
$\emptyset \subseteq \{\emptyset,\{\emptyset\}\}$
$\emptyset \in \{\emptyset,\{\emptyset\}\}$
$\{\emptyset\} \subseteq \{\emptyset,\{\emptyset\}\}$
$\{\emptyset\} \in \{\emptyset,\{\emptyset\}\}$
$\emptyset \subseteq \{\{\emptyset,\{\emptyset\}\}\}$

Solution 3.2. a) $W \subseteq V$ iff $V \cap W = W$. Proof: We have to show that
(i) if $W \subseteq V$, then $V \cap W = W$, and conversely, (ii) if $V \cap W = W$, then $W \subseteq V$.
Proof of (i): Suppose $W \subseteq V$. In order to show that $V \cap W = W$ it suffices – by the axiom of extensionality – to show that $V \cap W$ and W have the same elements. Clearly, each element of $V \cap W$ is also an element of W. Conversely, that each element of W also is an element of $V \cap W$ follows from the assumption that $W \subseteq V$.
Proof of (ii): Suppose $V \cap W = W$. To show: $W \subseteq V$. So, let $x \in W$. Then it follows from $V \cap W = W$ that $x \in V \cap W$. Hence, $x \in V$.
b) $W \subseteq V$ iff $V \cup W = V$ is shown in a similar way.

Solution 3.3. a) To show: $U - (V \cup W) = (U - V) \cap (U - W)$.

$$\begin{aligned}
\text{Proof: } x \in U - (V \cup W) &\rightleftarrows\ x \in U \wedge \neg\,(x \in V \cup W)\\
&\rightleftarrows\ x \in U \wedge \neg\,(x \in V \vee x \in W)\\
&\rightleftarrows\ x \in U \wedge (x \notin V \wedge\ x \notin W)\\
&\rightleftarrows\ (x \in U \wedge\ x \notin V) \wedge (x \in U \wedge\ x \notin W)\\
&\rightleftarrows\ x \in (U - V)\ \wedge\ x \in (U - W)\\
&\rightleftarrows\ x \in (U - V)\ \cap\ (U - W).
\end{aligned}$$

b) $U - (V \cap W) = (U - V) \cup (U - W)$ is shown in a similar way.

Solution 3.4. a) Let $U = \emptyset$, $V = \{\emptyset\}$ and $W = \{\{\emptyset\}\}$. Then $U \in V$ and $V \in W$, but $U \notin W$. b) Proof: Suppose that $U \subseteq V$ and $V \subseteq W$, i.e., $\forall x[x \in U \rightarrow x \in V]$ and $\forall x[x \in V \rightarrow x \in W]$. Then it follows that $\forall x[x \in U \rightarrow x \in W]$, i.e., $U \subseteq W$.

Solution 3.5. $\emptyset$ has only one subset: $\emptyset$. So, $P(\emptyset) = \{\emptyset\}$.
$P(\emptyset) = \{\emptyset\}$ has $2^1 = 2$ subsets: $\emptyset$ and $\{\emptyset\}$. So, $P(P(\emptyset)) = \{\emptyset, \{\emptyset\}\}$.
$P(P(\emptyset)) = \{\emptyset, \{\emptyset\}\}$ has $2^2 = 4$ subsets: $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$ and $\{\emptyset, \{\emptyset\}\}$.
So, $P(P(P(\emptyset))) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

Solution 3.6. (a) Suppose that $W \subseteq V$. Then $\forall x[x \subseteq W \rightarrow x \subseteq V]$, in other words, $\forall x[x \in P(W) \rightarrow x \in P(V)]$ and this means precisely that $P(W) \subseteq P(V)$.
(b) Suppose $P(W) \subseteq P(V)$, i.e., $\forall x[x \in P(W) \rightarrow x \in P(V)]$, in other words, $\forall x[x \subseteq W \rightarrow x \subseteq V]$. Now we know $W \subseteq W$. Hence also $W \subseteq V$.
(c) Suppose $P(W) = P(V)$. Then $P(W) \subseteq P(V)$ and $P(V) \subseteq P(W)$. Hence, applying (b) twice, $W \subseteq V$ and $V \subseteq W$. Hence $W = V$.
(d) Suppose $P(W) \in P(V)$, i.e., $P(W) \subseteq V$. Now $W \in P(W)$, and so $W \in V$.
Warning: The converse of (d), if $W \in V$, then $P(W) \in P(V)$, does not hold. Counterexample: Let $W := \{\emptyset\}$ and $V := \{\{\emptyset\}\}$. Then $P(W) = \{\emptyset, \{\emptyset\}\}$ and $P(V) = \{\emptyset, \{\{\emptyset\}\}\}$. So $P(W) \not\in P(V)$, while $W \in V$.

Solution 3.7. a) Proof: Suppose $P(W) \in PP(V)$. This is equivalent to $P(W) \subseteq P(V)$. Since $W \in P(W)$, it follows that $W \in P(V)$.
b) Proof: Suppose $W \in P(V)$, i.e., $W \subseteq V$. Then $\forall x[x \subseteq W \rightarrow x \subseteq V]$, i.e., $\forall x[x \in P(W) \rightarrow x \in P(V)]$, i.e., $P(W) \subseteq P(V)$, or equivalently, $P(W) \in P(P(V))$.
c) Proof: Suppose $P(W) \subseteq PP(V)$, i.e., $\forall x[x \in P(W) \rightarrow x \in PP(V)]$. $W \subseteq W$, so $W \in P(W)$; therefore, $W \in PP(V)$; in other words, $W \subseteq P(V)$.
d) Proof: Suppose $W \subseteq P(V)$. Then $\forall x[x \subseteq W \rightarrow x \subseteq P(V)]$, i.e., $\forall x[x \in P(W) \rightarrow x \in P(P(V))]$, or, equivalently, $P(W) \subseteq P(P(V))$.

Solution 3.8. i) $\{v\} \neq \emptyset$. So, by the regularity axiom, there is some $z \in \{v\}$ such that $z \cap \{v\} = \emptyset$, i.e., $v \cap \{v\} = \emptyset$. Now suppose $v \in v$. Then $v \in v$ and $v \in \{v\}$; so, $v \cap \{v\} \neq \emptyset$. Contradiction. Therefore, by the regularity axiom it follows that $v \not\in v$.
ii) $\{v_1, \ldots, v_n\} \neq \emptyset$. So, by the regularity axiom, there is some $z \in \{v_1, \ldots, v_n\}$ such that $z \cap \{v_1, \ldots, v_n\} = \emptyset$. Now suppose $v_1 \in v_2 \wedge v_2 \in v_3 \wedge \ldots v_{n-1} \in v_n \wedge v_n \in v_1$. Then there is no $z \in \{v_1, \ldots, v_n\}$ such that $z \cap \{v_1, \ldots, v_n\} = \emptyset$. Contradiction.

Solution 3.9. a) From right to left is trivial. From left to right: Suppose $(v, w) = (x, y)$, i.e., $\{\{v, \emptyset\}, \{w, \{\emptyset\}\}\} = \{\{x, \emptyset\}, \{y, \{\emptyset\}\}\}$. So, these two sets have the same elements; hence, (i) $\{v, \emptyset\} = \{x, \emptyset\}$ and $\{w, \{\emptyset\}\} = \{y, \{\emptyset\}\}$, or (ii) $\{v, \emptyset\} = \{y, \{\emptyset\}\}$ and $\{w, \{\emptyset\}\} = \{x, \emptyset\}$. In case (i) $v = x$ and $w = y$. In case (ii) it follows from $\emptyset \neq \{\emptyset\}$ that $v = \{\emptyset\}$ and $y = \emptyset$; $w = \emptyset$ and $x = \{\emptyset\}$. Hence, $v = x$ and $w = y$.
b) From right to left is trivial. So, suppose $(v, w) = (x, y)$, i.e., $\{\{v, \emptyset\}, \{w\}\} = \{\{x, \emptyset\}, \{y\}\}$. So, these two sets have the same elements. Hence, (i) $\{v, \emptyset\} = \{x, \emptyset\}$ and $\{w\} = \{y\}$, or (ii) $\{v, \emptyset\} = \{y\}$ and $\{w\} = \{x, \emptyset\}$. In case (i), $v = x$ and $w = y$. In case (ii), $v = y = \emptyset$ and $w = x = \emptyset$; so, again $v = x$ and $w = y$.

Solution 3.10.

$$\begin{aligned}
(u, v) \in U \times (V \cup W) \text{ iff } & u \in U \text{ and } v \in V \cup W \\
& u \in U \text{ and } (v \in V \text{ or } v \in W) \\
& (u \in U \text{ and } v \in V) \text{ or } (u \in U \text{ and } v \in W) \\
& (u, v) \in U \times V \text{ or } (u, v) \in U \times W \\
& (u, v) \in (U \times V) \cup (U \times W).
\end{aligned}$$

Solution 3.11. Counterexample: Let $U = \{1\}$, $V = \{2\}$, $W = \{3\}$. Then $U \times (V \times W) = \{1\} \times \{(2,3)\} = \{(1,(2,3))\}$, which is different from $(U \times V) \times W = \{(1,2)\} \times \{3\} = \{((1,2),3)\}$, since $(1,(2,3)) \neq ((1,2),3)$.

Solution 3.12. Dom(R) = $\{0,1,2,4\}$, Ran(R) = $\{1,2,3,4,7\}$. R is not a function, because $0 \in$ Dom(R) and there is more than one $z \in$ Ran(R) such that $(0,z) \in R$. $\breve{R} = \{(1,0),(3,0),(4,0),(1,2),(2,1),(7,4)\}$. $R;S = \{(0,4),(0,2),(2,4)\}$. $S;R = \{(1,7),(3,1),(5,1),(5,3),(5,4)\}$.

Solution 3.13. a) Let U be a partition of V. To prove: S is reflexive, symmetric and transitive. (1) S is reflexive. Suppose $v \in V$. Then there is precisely one set $W \in U$ such that $v \in W$; hence, vSv. (2) S is symmetric. Suppose $v,w \in V$ and vSw, i.e., there is a set W in U such that both v and w are elements of W. Then also w and v are elements of W; hence, wSv. (3) S is transitive. Suppose $u,v,w \in V$ and uSv and vSw. Then for some W_1 in U both $u \in W_1$ and $v \in W_1$. Also for some W_2 in U both $v \in W_2$ and $w \in W_2$. Since U is a partition of V and $v \in W_1 \cap W_2$, it follows that $W_1 = W_2$. So, $u \in W_1$ and $w \in W_1$; therefore uSw.
b) vSw is defined as follows: there is a set $[u]_R$ in V/R such that $v,w \in [u]_R$, i.e., vRu and wRu. To prove: vSw iff vRw. From left to right: suppose vSw, i.e., vRu and wRu for some $u \in V$. Then vRu and uRw. Hence, vRw. From right to left: Suppose vRw. Then vRv and wRv. Hence, there is a set $[u]_R$ in V/R, namely $[v]_R$, such that $v \in [u]_R$ and $w \in [u]_R$, i.e., vSw.

Solution 3.14. a) R is neither reflexive nor transitive. b) R is not symmetric. c) R is an equivalence relation. d) R is an equivalence relation.

Solution 3.15. To prove: (1) $\cup\{V_r \mid r \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R}$. (2) The elements of $\{V_r \mid r \in \mathbb{R}\}$ are pairwise disjoint. a) Proof for $V_r = \{(x,y) \in \mathbb{R}^2 \mid y = x + r\}$: (1) For any $x,y \in \mathbb{R}$ take $r := y - x$. Then $(x,y) \in V_r$. (2) Suppose $r \neq r'$. Then, clearly, $V_r \cap V_{r'} = \emptyset$. Geometrically, V_r as defined above is a straight line cutting the y-axis in r and the x-axis in $-r$. The equivalence relation R is defined by $(x_1,y_1)R(x_2,y_2)$:= for some $r \in \mathbb{R}$, $(x_1,y_1) \in V_r$ and $(x_2,y_2) \in V_r$. Hence, $(x_1,y_1)R(x_2,y_2)$ iff $y_1 - x_1 = y_2 - x_2$.
b) The proof that $\{V_r \mid r \in \mathbb{R}\}$ with $V_r := \{(x,y) \in \mathbb{R}^2 \mid r = x^2 + y^2\}$ is a partition of $\mathbb{R} \times \mathbb{R}$ is analogous to the proof given in a). In this case V_r is a circle with centre $(0,0)$ and radius r. The equivalence relation R is defined by $(x_1,y_1)R(x_2,y_2) := x_1{}^2 + y_1^2 = x_2^2 + y_2^2$.

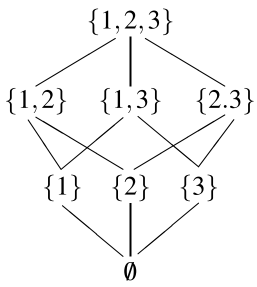
Solution 3.16. To prove: (1) $\cup\{V_n \mid n \in \mathbb{Z}\} = \mathbb{Z}$. (2) The different elements of $\{V_n \mid n \in \mathbb{Z}\}$ are pairwise disjoint. Proof: (1) Take any $m \in \mathbb{Z}$. Then there is an $n \in \mathbb{Z}$ such that $m = n + 5 \cdot q$ for some $q \in \mathbb{Z}$. So, $m \in V_n$.

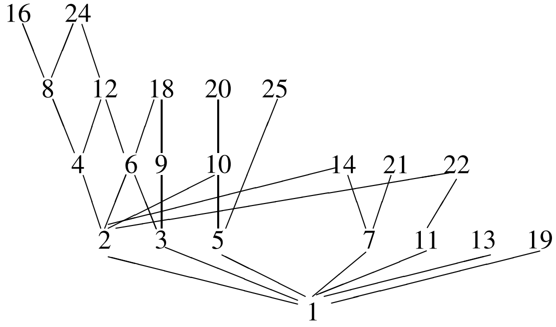
(2) $V_0 = \{\ldots, -10, -5, 0, 5, 10, \ldots\}$, $V_5 = V_0$,
$V_1 = \{\ldots, -9, -4, 1, 6, 11, \ldots\}$, $V_6 = V_1$,
$V_2 = \{\ldots, -8, -3, 2, 7, 12, \ldots\}$, $V_7 = V_2$,
$V_3 = \{\ldots, -7, -2, 3, 8, 13, \ldots\}$, $V_8 = V_3$,
$V_4 = \{\ldots, -6, -1, 4, 9, 14, \ldots\}$, $V_9 = V_4$, etc.

Solution 3.17. For any set V consider the empty relation $R_\emptyset$ on V, i.e., for all $x,y \in V$, not $xR_\emptyset y$. Clearly, $R_\emptyset$ is not reflexive, but $\forall x,y \in V[xR_\emptyset y \rightarrow yR_\emptyset x]$ is logically true, since $xR_\emptyset y$ is false; in a similar way one sees that $R_\emptyset$ is transitive.

Solution 3.18. The argument presupposes there is at least one pair (x,y) such that xRy. This argument is not valid if $R = \emptyset$.

Solution 3.19.





Solution 3.20. a) R_1 is a relation between $\{1,2,3,4\}$ and $\{1,2,3,4\}$. b) R_2 is a function from $\{1,2,3,4\}$ to $\{1,2,3,4\}$. c) R_3 is a bijection from $\{1,2,3,4\}$ to $\{1,2,3,4\}$. d) $R_1;R_2 = \{(2,2),(3,2),(4,2),(4,3)\}$ is a relation between $\{1,2,3,4\}$ and $\{1,2,3,4\}$. e) $\breve{R}_3$ is a bijection from $\{1,2,3,4\}$ to $\{1,2,3,4\}$.

Solution 3.21. a) Proof: Suppose $g \circ f : U \to W$ is injective, $x \neq x'$ and $f(x) = f(x')$. Then, because $g : V \to W$ is a function, $g(f(x)) = g(f(x'))$. But $g \circ f$ is injective. So, we have a contradiction.
b) Proof: Suppose $g \circ f : U \to W$ is a surjection. Then for every $w \in W$ there is $u \in U$ such that $w = g(f(u))$. Hence, for every $w \in W$ there is $v \in V$, namely $v = f(u)$, such that $w = g(v)$. In other words: $g : V \to W$ is surjective.
c) Counterexample: $g^* \circ f^* : \mathbb{N} \to \mathbb{N}$ is an injection; but $g^*(0) = 0$ and $g^*(1) = 0$; hence, $g^* : \mathbb{N} \to \mathbb{N}$ is not an injection.
d) Counterexample: $g^* \circ f^* : \mathbb{N} \to \mathbb{N}$ is a surjection, but there is no $n \in \mathbb{N}$ such that $0 = f^*(n)$.
e) Counterexample: $g^* \circ f^* : \mathbb{N} \to \mathbb{N}$ is a bijection, but $f^* : \mathbb{N} \to \mathbb{N}$ is not a surjection and $g^* : \mathbb{N} \to \mathbb{N}$ is not an injection.

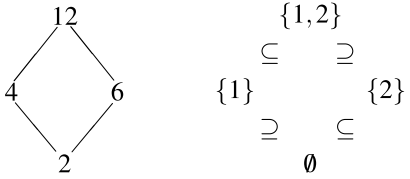
Solution 3.22. Let $f : V \to W$. $\breve{f} : W \to V$:= for all $w \in W$ there is precisely one $v \in V$ such that $\breve{f}(w) = v$, or equivalently, $f(v) = w$. Hence, $\breve{f} : W \to V$ iff $f : V \to W$ is a bijection.

Solution 3.23. a) Proof: Suppose $f : U \to V$ and $g : V \to W$ are injective. Then for any $x, x' \in U$, if $x \neq x'$, then $f(x) \neq f(x')$ and $g(f(x)) \neq g(f(x'))$.
b) Proof: Suppose $f : U \to V$ and $g : V \to W$ are surjective. Then for every $w \in W$ there is $v \in V$ such that $w = g(v)$. Also for every $v \in V$ there is $u \in U$ such that $v = f(u)$. Hence, for every $w \in W$ there is $u \in U$ such that $w = g(f(u))$.
c) This follows immediately from a) and b).

Solution 3.24. $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, defined by $f(n,m) = 2^m(2n+1) - 1$, is injective. Proof: suppose that $2^m(2n+1) - 1 = 2^{m'}(2n'+1) - 1$. Then, supposing that $m \geq m'$, $2^{m-m'} = \frac{2n'+1}{2n+1}$. But $2^{m-m'}$ is even, except when $m = m'$; and an odd number divided by an odd number is again an odd number. So, $2^{m-m'} = 1$ and $m = m'$. Consequently, also $n = n'$.

Solution 3.25. a) Suppose f were an isomorphism from $\langle \mathbb{N}, < \rangle$ to $\langle \mathbb{Z}, < \rangle$. Let $f(0) = z$ with $z \in \mathbb{Z}$. Since f is an isomorphism, for all $k \in \mathbb{N}$, $z \leq f(k)$. So, f is not surjective, since the elements of $\mathbb{Z}$ smaller than z are not in the range of f.
b) Suppose f were an isomorphism from $\langle \mathbb{Z}, < \rangle$ to $\langle \mathbb{Q}, < \rangle$. Let $f(0) = q_1$ and $f(1) = q_2$ with $q_1,\ q_2 \in \mathbb{Q}$. Then between q_1 and q_2 there is a rational number q with $q_1 < q < q_2$. But there is no integer i in $\mathbb{Z}$ between 0 and 1 such that $f(i) = q$. Hence, f is not surjective.

Solution 3.26. Let $f : \{2,4,6,12\} \rightarrow P(\{1,2\})$ be defined as follows: $f(2) = \emptyset$, $f(4) = \{1\}$, $f(6) = \{2\}$ and $f(12) = \{1,2\}$. Then f is a bijection and for all $n,m \in \{2,4,6,12\}$, n/m iff $f(n) \subseteq f(m)$.



Solution 3.27. a) The function f from $\mathbb{Z}$ to $\mathbb{N}_{even}$, defined by $f(n) = 4n$ and $f(-n) = 4n - 2$ for any $n \in \mathbb{N}$, is a one-one correspondence between $\mathbb{Z}$ and $\mathbb{N}_{even}$:

$$\begin{array}{lcccccccc}
\mathbb{Z}: & 0 & -1 & 1 & -2 & 2 & -3 & 3 & \dots \\
 & | & | & | & | & | & | & | & \\
\mathbb{N}: & 0 & 1 & 2 & 3 & 4 & 5 & 6 & \dots \\
 & | & | & | & | & | & | & | & \\
\mathbb{N}_{even}: & 0 & 2 & 4 & 6 & 8 & 10 & 12 & \dots
\end{array}$$

b) The function f from $\mathbb{N}_{even}$ to $\mathbb{N}_{odd}$, defined by $f(2n) = 2n + 1$ for all $n \in \mathbb{N}$, is a one-one correspondence between the two sets in question:

$$\begin{array}{lcccccccc}
\mathbb{N}_{even}: & 0 & 2 & 4 & 6 & 8 & 10 & 12 & \dots \\
 & | & | & | & | & | & | & | & \\
\mathbb{N}_{odd}: & 1 & 3 & 5 & 7 & 9 & 11 & 13 & \dots
\end{array}$$

Solution 3.28. a) Let $\{0,1\}^n$ be the set of all finite sequences of 0's and 1's of length n ($n \in \mathbb{N}$). For each $n \in \mathbb{N}$, $\{0,1\}^n$ has 2^n elements. Now $\{0,1\}^*$ is the union of all sets $\{0,1\}^n$ with $n \in \mathbb{N}$. Hence, $\{0,1\}^*$ is the union of denumerably many finite sets and hence denumerable. b) Let Σ_n ($n \in \mathbb{N}$) be the set of all words over Σ of length n. Let k be the number of symbols (characters) in Σ. Then Σ_n has k^n elements. Now Σ^* is the union of all Σ_n with $n \in \mathbb{N}$. Hence, Σ^* is the union of denumerably many finite sets and hence denumerable.

Solution 3.29. Suppose V is enumerable. Let V_n ($n \in \mathbb{N}$) be the set of all finite sequences of elements of V of length n. For each $n \in \mathbb{N}$, V_n is enumerable. Now V^*, the set of all finite sequences of elements of V, is the union of all V_n with $n \in \mathbb{N}$. Hence, V^* is the union of denumerably many enumerable sets and hence, by Theorem 3.18, V^* is denumerable.

Solution 3.30. (i) $f : (0,\ 1] \to [0,\ 1]$, defined by $f(1) = 0$, $f(\frac{1}{n}) = \frac{1}{n-1}$ for $n \in \mathbb{N}$, $n \geq 2$, $f(x) = x$ if $x \notin \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, is a bijection.
(ii) $f : (0,\ 1] \to (0,\ 1)$, defined by $f(\frac{1}{n}) = \frac{1}{n+1}$ for $n \in \mathbb{N}$, $n \geq 1$, $f(x) = x$ if $x \notin \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, is a bijection.
(iii) $f : (0,\ 1] \to [0,\ 1)$, defined by $f(1) = 0$, $f(x) = x$ if $x \neq 1$, is a bijection.

Solution 3.31. By Exercise 3.28, Σ^* is denumerable. Hence, by Theorem 3.21, $P(\Sigma^*)$ is uncountable. And $P(\Sigma^*)$ is precisely the set of all languages over Σ, since L is a language over Σ iff $L \in P(\Sigma^*)$.

Solution 3.32. a) $f : [0,\ 1] \to [0,\ 2]$, defined by $f(x) = 2x$, is a bijection.
b) $f : (0,\ 1) \to (0,\ 3)$, defined by $f(x) = 3x$, is a bijection.

Solution 3.33. Suppose V is infinite. Then by the axiom of choice V has a denumerable subset $\{x_0, x_1, x_2, \ldots\}$. By Theorem 3.24, $g : V \to V - \{x_0\}$, defined by $g(x_i) = x_{i+1}$ and $g(x) = x$ if $x \notin \{x_0, x_1, x_2, \ldots\}$, is a bijection with domain V and range $V - \{x_0\}$. Conversely, suppose V is Dedekind infinite, i.e., there is an injective function with domain V and whose range is a proper subset of V. Then V cannot be finite. Therefore, V is infinite.

References

1. Arrow, K., E.Maskin, *Social Choice and Individual Values*. Yale University Press, 1951, 2012.
2. Cusanus, N. (Nikolaus von Kues), *De docta ignorantia*. In: *Philosophisch-Theologische Schriften* Band I, II, III. Herder Verlag, Wien, 1964, 1966, 1967.
3. Dalen, D. van, H.C. Doets, H.C.M. de Swart, *Sets: Naive, Axiomatic and Applied*. Pergamon Press, Oxford, 1978.
4. Gödel, K., *Russell's mathematical logic*. In: P.A. Schilpp (ed.), *The philosophy of Bertrand Russell* (Tudor, N.Y., 1944), and in: P. Benacerraf and H. Putnam, *Philosophy of Mathematics, Selected readings* (1964, 1983).
5. Gödel, K., What is Cantor's continuum problem? *American Mathematical Monthly*, vol. 54, 1947, pp. 515-525.
6. Heijenoort, J. van, *From Frege to Gödel*. A source book in mathematical logic 1879-1931. Harvard University Press, Cambridge, Mass., 1967.
7. Kant, I., *The Critique of Pure Reason*. William Benton, Publisher, Encyclopaedia Brittanica, 1952, 1978, in particular pp. 14-18 (Great Books of the Western World 42).
8. Quine, W.V., *Two Dogmas of Empiricism*. In: W.V. Quine, *From a Logical Point of View*. Harvard University Press, Cambridge, Mass., 1953, 1961, 1980.