# CHAPTER 1
# SETS

Ultimately this book is about sets, structured by other sets, and interrelated by things that are yet again sets. Therefore we begin with a brief discussion of set theoretical notions. Beyond meeting the author's desire for completeness, it develops a series of notations and notes some aberrations from the standard that particularly suit the subject of partial algebras. The avid reader can skip it, and refer back at need; on the other hand, who wants a proper introduction should look elsewhere: to the excellent book (in Dutch) by Van Dalen, Doets and De Swart [DDS], or P.R. Halmos' classical *Naive Set Theory* [H].

## §a Sets and classes

Given a determinate property $P$ of determinate things of some kind or other, we may consider the collection (or class) $\{x|x \text{ has } P\}$ of all things that have the property $P$. Now if $x$ has $P$, we say $x$ is an element, or a member, of the class $\{x|x \text{ has } P\}$; in symbols we write $x \in X$ for '$x$ is an element of $X$', and $x \notin X$ for '$x$ is not an element of $X$' (this is one of many cases where one expresses negation by running through).

In principle, collections may be collected into collections of higher order, for example we could have

$$\{x|x \text{ is a collection of natural numbers}\};$$

but there are limits.

Consider this example: one can collect paintings, but only paintings that have been *painted*. One cannot construct paintings by collecting them.

Now there certainly are more ethereal forms of construction than painting, and to a certain extent *collecting*, as in collecting collections of natural numbers, may be considered one of them. This extent needs to be circumscribed, however; for in set theory, as in art collecting, fraud exists, and its punishment is paradox. For example, in *Russell's Paradox* one defines $K$ as $\{x|x \notin x\}$, and asks whether $K \in K$. If $K \in K$, $K$ has the defining property of $K$, so $K \notin K$; but then since $K$ collects all things that do not belong to themselves, $K \in K$. Apparently, to say the least, $K$ is not the kind of thing that we should consider for elementhood of $K$.

It is customary to distinguish between classes that may safely be considered elements of other classes, and call these *sets*, and classes that can not safely be so considered. The latter are called *proper classes*. A proper class may be viewed as too large to be a set, for example,

$$\{x|x \notin x\}$$

contains, according to a common notion of sets, all sets. The word 'collection' is used as an equivalent of 'class', in particular for classes the elements of which are themselves classes. Collections will often be denoted by letters of a different font: $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and such.

Sets have to be constructed in some way. The axioms of Zermelo, as supplemented by Fraenkel and Skolem, present a catalogue of respectable construction principles. They are the ultimate inspiration of the axioms listed in the sequel.

The first axiom is not a construction principle. It states the characteristic feature of classes:

***Extensionality***: if classes $X$ and $Y$ have the same elements, then $X = Y$.

## §b Sets of nonsets

One can do mathematics, in principle, with nothing but sets; that is, sets with ultimately nothing in them, differing only in the way it is buried. We shall not take such an airy atttitude here. Instead we assume that there are lots of respectable things that are not sets, and that some of them even come in sets. In particular, we have the set $\mathbb{N}$ of natural numbers, the set $\mathbb{Z}$ of integers, the set $\mathbb{Q}$ of rationals, and the set $\mathbb{R}$ of real numbers.

The natural numbers are of fundamental importance. We recall the principle of *mathematical induction*:

(MI) If $P$ is a predicate of natural numbers that applies to 0 and satisfies, for all natural numbers $n$,

$$P(n) \text{ implies } P(n+1),$$

then $P$ applies to all natural numbers.

## §c Pairs

A slow but sure way of constructing sets is

***Pairing***: if $a$ and $b$ are any definite objects, then there exists a set $\{a, b\}$ containing just $a$ and $b$.

We call this set the *pair* of $a$ and $b$. *Note* that we do not consider arbitrary classes to be definite objects. Primitive elements of the kind mentioned in §b are sufficiently definite. A pair of classes is a set only if these classes are sets.

By Extensionality, $\{a, b\} = \{b, a\}$. The special kind of pair with $a = b$ is called a *singleton*. Instead of $\{a, a\}$ we write $\{a\}$.

The basis for all forms of order is the notion of *ordered pair*. The ordered pair with first element $x$ and second element $y$ we denote by $\langle x, y\rangle$. The fundamental property of ordered pairs is

$$\langle x, y\rangle = \langle u, v\rangle \text{ if and only if } x = u \text{ and } y = v. \tag{*}$$

Ordered pairs may be simply modelled by sets. Kuratowski defined $\langle x, y\rangle$ as

$$\{\{x\}, \{x, y\}\};$$

it is easy to see that under this definition the condition (*) is satisfied.

## §d Union

Let $X$ be a collection of classes. Then $\bigcup X$, the *union* of $X$, is the class of all elements of elements of $X$. Actually, there is no need to require that the elements of $X$ all be classes; if $x \in X$ is not a class, then it has no elements, so there will be no trace of $x$ in $\bigcup X$. We postulate that union is another way of constructing sets:

***Union***: If $X$ is a set, then so is $\bigcup X$.

Clearly, if $\{X\}$ is a set, then $\bigcup\{X\} = X$. The union of a pair $\{X, Y\}$ is usually written as $X \cup Y$.

We denote the class without elements, the *empty* or *void* class, by $\emptyset$. That there is only one such class is a consequence of Extensionality. From the Union Axiom and the existence of sets of primitive objects follows that $\emptyset$ is a set:

**d1 Proposition**. The void class is a set.

**Proof**. Let $A$ be any given set, such as $\mathbb{N}$, of things that are not classes. Then $\emptyset = \bigcup A$. ⊠

Let $x_0, \ldots, x_{n-1}$ be a finite list of things. Then we denote the class that contains exactly these things by $\{x_0, \ldots, x_{n-1}\}$. For example, the class containing exactly the numbers 3, 5, 6 and 9 is $\{3, 5, 6, 9\}$.

**d2 Proposition**. If $x_0, \ldots, x_{n-1}$ are definite objects (in particular, none of them should be a proper class), then $\{x_0, \ldots, x_{n-1}\}$ is a set.

**Proof**. By induction on $n$. That is to say, we construe the proposition as a statement $P(n)$ about a natural number $n$, and then apply the principle (MI) of §b to the predicate $P$.

If $n = 0$, then our list is empty, and $\{x_0, \ldots, x_{n-1}\} = \emptyset$, which we have seen to be a set.

Now suppose lists of $n$ decent objects give rise to sets. Then if we have a list $x_0, \ldots, x_n$ of $n + 1$ acceptable items, $\{x_0, \ldots, x_{n-1}\}$ is a set by hypothesis, and $\{x_n\}$ is a set by the pairing axiom. So again by pairing,

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

is a set; and $\{x_0, \ldots, x_n\} = \{x_0, \ldots, x_{n-1}\} \cup \{x_n\}$. ⊠

## §e Replacement and Regularity

We introduce binary relations, in their primal form of classes of ordered pairs. Then we state the Replacement Axiom, and derive a number of consequences. The Regularity Axiom is a general statement about classes, comparable to Extensionality, but less easy to formulate.

A *binary relation* is a class of ordered pairs. For the time being we will simply speak of *relations*. The *domain* of a relation $R$ is the class $\operatorname{Dom} R$ defined as

$$\{y \mid \exists x \, \langle x, y\rangle \in R\},$$

and the *range* of $R$, notation $\operatorname{Ran} R$, is

$$\{x \mid \exists y \, \langle x, y\rangle \in R\}.$$

For any object $y$ we define the class $y/R$ (pronounce '$y$ over $R$') as

$$\{x \mid \langle x, y\rangle \in R\}.$$

More generally, for any class $Y$ we define $R[Y]$ as

$$\{x \mid \exists y \in Y \, \langle x, y\rangle \in R)\}.$$

Now we introduce another way of constructing sets.

***Replacement***: Let $R$ be a relation and $Y$ a set. If for every $y \in Y$, $y/R$ is a set, then $R[Y]$ is a set.

In words: replacing every element of a set by the elements of a set leaves us with a set. This is a consequential axiom.

Suppose $X$ is a nonvoid class: then the *intersection* of $X$, notation $\bigcap X$, is the class

$$\{y \mid \forall x \in X \, y \in x\}.$$

The class $X$ must be supposed nonvoid, since *everything* is an element of every element in the void class, and we do not generally know what everything is. Again, if $\{X\}$ is a set, then $\bigcap\{X\} = X$. The intersection of a pair $\{X, Y\}$ is usually written $X \cap Y$. Classes $X$, $Y$ such that $X \cap Y = \emptyset$ are called *disjoint*.

**Proposition** (*Comprehension*). If $X$ is a set and $Y$ a class, then $X \cap Y$ is a set.

**Proof**. Let $X$ be a set, and define $R$ to be the relation

$$\{\langle x, x\rangle \mid x \in X \text{ and } x \in Y\}.$$

Then for every $x \in X$, $x/R$ is a set: either $x \notin Y$, and then $x/R = \emptyset$, or $x \in Y$, and $x/R = \{x\}$. So $R[X]$ is a set. But $R[X] = X \cap Y$. ⊠

As a consequence, if $X$ is a set and $P(-)$ a predicate that makes sense for the elements of $X$, then $\{x \mid x \in X \text{ and } P(x)\}$ is a set (*Separation Principle*); for it is the intersection of $X$ and the class $\{x \mid P(x)\}$.

The *difference* $X - Y$ of two classes $X$ and $Y$ is $\{x \mid x \in X \text{ and } x \notin Y\}$. If $X$ is a set, then so is $X - Y$.

We think of collections as constructed, not as given things. If something is an element of a collection, it is so because we put it there. Hence, classes may consist of sets, of sets of sets, and so on; but ultimately we collect things which are not classes. This is expressed by

***Regularity***: If a class has elements, it has elements from which it is disjoint.

(Observe that if $a$ is not a class, we have $X \cap a = \emptyset$.)

## §f The Powerset

A class $X$ is a *subclass* of a class $Y$, and $Y$ is a *superclass* of $X$, if all elements of $X$ are also elements of $Y$; notation $X \subseteq Y$ and $Y \supseteq X$. We also say $X$ *is contained in* $Y$, and that $Y$ *contains*[1] or *extends* $X$. We write $X \subset Y$ ($X$ is a *proper subclass* of $Y$, $X$ is *properly contained* in $Y$) if $X \subseteq Y$ and $X \neq Y$. By Extensionality, in this case there must be things that belong to $Y$ and not to $X$.

[1] This creates an ambiguity between *elements* and *subsets*.

The relation $\subseteq$ of *being contained in* is commonly called *inclusion*.

To denote subclasses of a given class $X$, we often write $\{x \in X | \dots\}$ instead of $\{x | x \in X \;\&\; \dots\}$. So for example, $X - Y$ may be written as

$$\{x \in X | x \notin Y\}.$$

A superclass that happens to be a set is a *superset*. By Separation, subclasses of sets are sets; we speak of *subsets*. We postulate that the totality of subsets is a set:

***Powerset***: Let $X$ be a set. Then the class of all subsets of $X$ is a set.

We use the notation $\mathcal{P}X$ to refer to this set, the *powerset* of $X$.

## §g Relations

The *field* of a relation $R$ is the union of its range and its domain. We denote it by $\mathrm{Fld}\,R$; we might have defined it as $\{x, y | \langle x, y\rangle \in R\}$, or, assuming Kuratowski's definition of ordered pairs, $\bigcup\bigcup R$. If $\mathrm{Fld}\,R \subseteq X$, then $R$ is called a relation *over* $X$. In the particular case that $\mathrm{Fld}\,R = X$, we also say that $R$ is a relation *on* $X$. For example, the well-known ordering $<$ of natural numbers (that is, $\{\langle m, n\rangle | m \in \mathbb{N} \;\&\; n \in \mathbb{N} \;\&\; m < n\}$) is a relation on $\mathbb{N}$, and so is identity, $\{\langle n, n\rangle | n \in \mathbb{N}\}$. For the identity relation it will be useful to have a generally applicable notation that shows the underlying class: we shall write $\Delta_X$ (the *diagonal over* $X$) for $\{\langle x, x\rangle | x \in X\}$.

If $X$ and $Y$ are classes, then their *cartesian*[2] *product* $X \times Y$ is the class

$$\{\langle x, y\rangle | x \in X \;\&\; y \in Y\}.$$

The product of two *sets* is again a set. The *square on* a class $X$ is $X \times X$; where it contrasts with the diagonal, we denote the square on $X$ by $\nabla_X$. The *cartesian hull* of a relation $R$, notation $\mathrm{Ch}R$, is $\mathrm{Ran}\,R \times \mathrm{Dom}\,R$.

Observe that $\mathrm{Dom}(X \times Y) = Y$, $\mathrm{Ran}(X \times Y) = X$, $\mathrm{Ran}\,\Delta_X = \mathrm{Dom}\,\Delta_X = X$, and for the ordering $<$ that we mentioned above, $\mathrm{Ran}(<) = \mathbb{N}$, and $\mathrm{Dom}(<) = \{1, 2, \dots\}$: the set $\mathbb{Z}_+$ of positive integers. If $R$ is a set, then $\mathrm{Dom}\,R$ and $\mathrm{Ran}\,R$ are sets as well.

We often write $Rxy$, $R(x, y)$ or $xRy$ instead of $\langle x, y\rangle \in R$. Observe that

$$R \subseteq \mathrm{Ch}R \subseteq \nabla_{\mathrm{Fld}\,R},$$

and that a relation over a set $X$ is a subset of $X \times X$.

A relation $R$ is *symmetric* if $xRy$ implies $yRx$ for all $x$ and $y$, and *transitive* if for all $x$, $y$ and $z$, $xRy$ & $yRx$ implies $xRz$. A relation that is both symmetric and transitive we call an *equivalence relation*, or simply an *equivalence*. We use lower case Greek letters for equivalences: $\theta$, $\eta$, $\zeta$, etc. Instead of $x\theta y$ we sometimes write $x \equiv_\theta y$ or $x \equiv y\ (\theta)$ or $x \equiv y \pmod{\theta}$; we say $x$ and $y$ *are equivalent modulo* $\theta$.

A *block system* is a collection of pairwise disjoint nonvoid classes. (These classes are the *blocks* of the system.)

Equivalences and block systems are two sides of the same coin. For an element $x$ of the field of an equivalence $\theta$, we call $x/\theta$ the *equivalence class of* $x$.

---

[2] After René Descartes (Renatus Cartesius), 1596-1650, philosopher, discoverer of analytical geometry.

If $X$ is a set, then the set of equivalence classes of elements of $X \cap \operatorname{Fld}\theta$ is called the *quotient* of $X$ over $\theta$, written $X/\theta$.

**g1 Proposition**. (a) If $\theta$ is an equivalence, then $\{x/\theta | x \in \operatorname{Fld}\theta\}$ is a block system.
(b) Let $\mathcal{A}$ be a block system, and define a relation $Q$ by

$$xQy \Leftrightarrow \text{if } x \text{ and } y \text{ belong to the same block of } \mathcal{A}.$$

Then $Q$ is an equivalence.

Let us denote, for the moment, by $\mathcal{A}_\theta$ the block system associated in (a) above to an equivalence $\theta$, and by $\theta_{\mathcal{A}}$ the equivalence associated in (b) to a block system $\mathcal{A}$. For block systems $\mathcal{A}$ and $\mathcal{B}$ we write $\mathcal{A} \le \mathcal{B}$ if every element of $\mathcal{A}$ is contained in an element of $\mathcal{B}$.

**g2 Proposition**. (a) If $\theta$ and $\eta$ are equivalences such that $\theta \subseteq \eta$, then $\mathcal{A}_\theta \le \mathcal{A}_\eta$.
(b) If $\mathcal{A}$ and $\mathcal{B}$ are block systems such that $\mathcal{A} \le \mathcal{B}$, then $\theta_{\mathcal{A}} \subseteq \theta_{\mathcal{B}}$.

If a block system consists of subsets of a given set $A$, we say it is a block system *in* or *of* $A$. We also say of the corresponding equivalence relation that it is an equivalence *in* or *of* $A$.

A relation $R$ is *reflexive* on a class $X$, or *X-reflexive*, if $xRx$ for all $x \in X$. In particular, an equivalence is reflexive on its field: if $xRy$, then $yRx$ by symmetry, hence $xRx$ and $yRy$ by transitivity. An equivalence with field $X$ will be called a *full equivalence (relation) of* $X$, or an *equivalence (relation) on* $X$. For example, for every $X$, the diagonal $\Delta_X$ is an equivalence on $X$. Another well-known example is congruence modulo some natural number $n$: for $x, y \in \mathbb{Z}$, $x \equiv y\ (n)$ if there exists $z \in \mathbb{Z}$ such that $x = y + zn$.

**Example**. Define a relation $\theta$ on the set $\mathbb{N} \times \mathbb{N}$ of pairs of natural numbers by

$$\langle x, y\rangle \theta \langle u, v\rangle \text{ if and only if } x + v = u + y.$$

Then $\theta$ is a full equivalence relation of $\mathbb{N}$. The mapping $\langle x, y\rangle \longmapsto x - y$ establishes a one-to-one correspondence of the equivalence classes with the integers.

A block system corresponding with a full equivalence relation of a set $X$ is called a *partition* of $X$.

The *composite* $Q \circ R$ of two relations $Q$ and $R$ is the relation defined by

$$z(Q \circ R)x \text{ if and only if for some } y,\ zQy \text{ and } yRx.$$

The composition symbol $\circ$ is often omitted.

The *skew union* of two relations $Q$ and $R$ is the relation

$$Q \,\text{J}\, R := R \cup (Q - (\operatorname{Ran}Q \times \operatorname{Dom}R));$$

so $y(Q \,\text{J}\, R)x$ if and only if either $yRx$ or $(x \notin \operatorname{Dom}R \ \&\ yQx)$.

The *inverse* of a relation $R$ is the relation $R^{-1}$ defined by

$$R^{-1} = \{\langle x, y\rangle | \langle y, x\rangle \in R\}.$$

For some inverses a special, inverted symbol is in use. Thus, the inverse of $\in$ is $\ni$; the inverse of $\subseteq$ is $\supseteq$, that of $\subset$ is $\supset$.

# §h Operations

Operations and functions are introduced (h1); next, finite sequences and families (h2); and the Axiom of Choice is stated (h3). In (h4) we finally explain why relations were called *binary* at their introduction.

### h1 Operations and functions

A relation $f$ is an *operation* if for every $x$ there is at most one $y$ such that $yfx$. If $f$ is an operation and $yfx$, then $y$ is the *value* that $f$ takes at the *argument* $x$; we write $f\colon x \longmapsto y$, and we denote $y$ by $f(x)$, or $fx$, or $f_x$. If $y$ exists such that $yfx$, we say that *$f$ is defined at $x$*, or that *$f(x)$ is defined*. We use the notation $f(x)\!\downarrow$ to express this, writing $f(x)\!\uparrow$ in the other case, when $f(x)$ is *undefined*. If we list, or pretend to list, the ordered pairs that constitute an operation, we may write

$$\{y_0 \longleftarrow\!\!| x_0, \dots, y_{n-1} \longleftarrow\!\!| x_{n-1}\}$$

instead of $\{\langle y_0, x_0\rangle, \dots, \langle y_{n-1}, x_{n-1}\rangle\}$.

Let $X$ and $Y$ be classes. An operation $f$ is a *mapping* (or just *map*) *from $X$ to $Y$*, or *of $X$ into $Y$*, notation $f\colon X \longrightarrow Y$ or $f\colon Y \longleftarrow X$, if $\mathrm{Dom} f = X$ and $\mathrm{Ran} f \subseteq Y$. The class $Y$ is the *codomain* of the mapping.[3]

A *function* is an operation that is a set. If $f$ is a mapping of a set, it is a function, by Replacement. The collection of all mappings from a set $X$ to a set $Y$ is a subset of $\mathcal{P}(Y \times X)$; it is denoted by $Y^X$.

Let $f$ be an operation. The *kernel* of $f$ is the equivalence relation

$$\ker f := \{\langle x_1, x_2\rangle \in \mathrm{Dom} f \times \mathrm{Dom} f \mid f(x_1) = f(x_2)\}.$$

of $\mathrm{Dom} f$. We call $f$ *injective* if $\ker f = \Delta_{\mathrm{Dom} f}$.

A mapping $f\colon X \longrightarrow Y$ is *surjective* (and $f$ is a mapping *onto $Y$*) if $\mathrm{Ran} f = Y$; and *bijective* if it is both in- and surjective. We write

$$f\colon X \hookrightarrow Y$$

(or $Y \hookleftarrow X$) if $f\colon X \longrightarrow Y$ is injective. For a surjective mapping we write $X \twoheadrightarrow Y$ or $Y \twoheadleftarrow X$; for a bijective mapping, $X \hookrightarrow\!\!\!\!\!\twoheadrightarrow Y$ or $Y \twoheadleftarrow\!\!\!\!\!\hookleftarrow X$.

*Composition* of operations is a special case of relation composition. The composite $f \circ g$ of operations $f$ and $g$ is again an operation, and

$$\textit{if } (f \circ g)(x)\!\downarrow \textit{ or } f(g(x))\!\downarrow\textit{, then } (f \circ g)(x) = f(g(x)); \tag{*}$$

using Kleene's complete equality (2§a), we may abbreviate (*) to:

$$(f \circ g)(x) \simeq f(g(x)).$$

**Proposition 1**. (a) If $f\colon X \longrightarrow Y$ and $g\colon Y \longrightarrow Z$, then $gf\colon X \longrightarrow Z$.
(b) If $f\colon X \longrightarrow Y$ and $g\colon Y \twoheadrightarrow Z$, then $gf\colon X \twoheadrightarrow Z$.
(c) If $f\colon X \hookrightarrow Y$ and $g\colon Y \hookrightarrow Z$, then $gf\colon X \hookrightarrow Z$.
(d) If $f\colon X \hookrightarrow\!\!\!\!\!\twoheadrightarrow Y$ and $g\colon Y \hookrightarrow\!\!\!\!\!\twoheadrightarrow Z$, then $gf\colon X \hookrightarrow\!\!\!\!\!\twoheadrightarrow Z$.

[3] A mapping involves an operation, and often mappings are referred to by the names of operations. Yet the two are not to be confused: a mapping has a codomain, and an operation a range; and the same operation gives rise to many mappings.

**Lemma** (*triangle completion lemma*). Let $f\colon X \twoheadrightarrow Y$ and $g\colon X \longrightarrow Z$ be mappings.
(i) There exists a mapping $h\colon Y \longrightarrow Z$ such that $h \circ f = g$ if and only if $\ker f \subseteq \ker g$.
(ii) If it exists, this mapping $h$ is uniquely determined by the condition $h \circ f = g$. Moreover,
(iii) $h$ is injective if and only if $\ker f = \ker g$, and
(iv) surjective if and only if $g$ is surjective.

The situation is sketched in the diagram below.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "g"'] & Y \arrow[dl, dashed, "h"] \\ Z & \end{tikzcd}$$

**Proof**. (i) ($\Rightarrow$) If $f(x_1) = f(x_2)$, then certainly $g(x_1) = h(f(x_1)) = h(f(x_2)) = g(x_2)$; so $\ker f \subseteq \ker g$.
($\Leftarrow$) Since $f$ is surjective, for every $y \in Y$ there exist $x \in X$ such that $y = f(x)$. Define: $h(y) = g(x)$. This will obviously work, *if* it is unambiguous. And it is: if $f(x_1) = f(x_2)$, then $g(x_1) = g(x_2)$ since $\ker f \subseteq \ker g$.

(ii) Suppose that $h_1$ and $h_2$ both satisfy the condition: $h_i \circ f = g$ for $i = 1, 2$. Take any $y \in Y$. Then $x \in X$ exist for which $y = f(x)$, and for any such $x$, $h_1(y) = h_1(f(x)) = g(x) = h_2(f(x)) = h_2(y)$.

(iii) ($\Rightarrow$) Suppose $A \in X/\ker f$. Then in particular $A \neq \emptyset$; say $a \in A$. It will suffice to show that $A = a/\ker(g)$. Inclusion from left to right is immediate by $\ker f \subseteq \ker g$. For the other direction, suppose $g(x) = g(a)$. Then $h(f(x)) = h(f(a))$, so $f(x) = f(a)$ since $h$ is injective.
($\Leftarrow$) Suppose $h(y_1) = h(y_2)$. There are $x_i$, $i = 1, 2$, such that $y_i = f(x_i)$. Then $g(x_1) = h(f(x_1)) = h(f(x_2)) = g(x_2)$, so $x_1$ and $x_2$ are in the same $\ker f$-equivalence class, which implies $y_1 = y_2$.

(iv) ($\Rightarrow$) If $g$ is not surjective, then neither is $h$.
($\Leftarrow$) Take any $z \in Z$. Since $g$ is surjective, there exist $x \in X$ such that $g(x) = z$. Take any such $x$. The block $x/\ker(g)$ contains a block of $\ker f$, say $x'/\ker f \subseteq x/\ker g$. Then $h(f(x')) = g(x') = g(x) = z$. ☒

In a typical application of this lemma, $f$ is the *quotient map* belonging to a full equivalence $\alpha$ of $X$, mapping $x \in X$ to $x/\alpha$. Note that the kernel of this $f$ is $\alpha$.

A mapping that is injective, is called an *injection*; likewise we speak of *surjections* and *bijections*. Observe that if $f\colon X \longrightarrow Y$ is bijective, the inverse $f^{-1}$ is a bijective mapping of $Y$ onto $X$. If there exists a bijection from $X$ onto $Y$, we say $X$ and $Y$ are *equipollent*. A set is *finite* if it is equipollent with an initial segment $\{0,\dots, n-1\}$ of $\mathbb{N}$; otherwise it is *infinite*. For a finite set $X$, $|X|$ is the number of the elements of $X$: the natural number $n$ such that $X$ is equipollent with $\{0,\ldots, n-1\}$ (see §3e for the infinite case). The set of all *finite* subsets of a set $Y$ we denote by $\wp Y$.

Sometimes we want to consider a diagonal $\Delta_X$ as a mapping. If the codomain is $Y$ ($\supseteq X$) we write $1_Y^X$, and speak of the *canonical embedding* of $X$ into $Y$. If $Y = X$, we may instead use the notation $1_X$, and speak of *the identity*, or *the identical mapping*, *on* $X$.

**Proposition 2**. Let $f$: $X \longrightarrow Y$ and $g$: $Y \longrightarrow X$ be mappings. If $g \circ f = 1_X$ and $f \circ g = 1_Y$, then $f$ is bijective and $g = f^{-1}$.

**Proof**. The identity $1_Y$ is surjective, so from $f \circ g = 1_Y$ we get that $f$ is surjective. On the other hand, if $f(x_1) = f(x_2)$, then

$$x_1 = 1_X(x_1) = (g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2) = 1_X(x_2) = x_2,$$

so $f$ is injective. Thus $f$ is bijective. For all $x \in X$ and $y \in Y$, if $f(x) = y$, then $g(y) = g(f(x)) = x$, and likewise $g(y) = x$ implies $f(x) = y$; so $g = f^{-1}$. ⊠

Let $f$ be an operation, and $U$ a class. The *restriction* $f\lceil U$ of $f$ to $U$ is the operation defined by $f\lceil U = \{\langle f(u), u\rangle \in f \mid u \in U\}$.[4] For any item $a$ that is not a class, we put $f\lceil a = \emptyset$. If $g$ is a restriction of $f$, we call $f$ an *extension* of $g$.

**h2 Sequences and families**

Sequences (arrays) of $n$ elements of a class $X$ are mappings from $\{0,\ldots, n-1\}$ to $X$; we write $X^n$ instead of $X^{\{0,\ldots,n-1\}}$. We shall often identify $X^2$ with $X \times X$. Usually sequences $\boldsymbol{x} \in X^n$ are denoted by pseudo-enumerations between angled brackets, as $\langle x_0,\ldots, x_{n-1}\rangle$. The indices in the enumeration may derive from another sequence, for example, $\langle x_4, x_9, x_5\rangle$ consists of elements of a sequence $\boldsymbol{x}$ of length no less than 10; or we may simply begin with 1, when we feel like it. We identify the sequence $\langle x_0, x_1\rangle$ with the ordered pair, $\langle x\rangle$, usually, with $x$, and the empty sequence $\varepsilon$ is the set $\emptyset$. The class of all finite sequences of elements of $X$ we denote by $X^{<\omega}$. The *concatenation* $\boldsymbol{x} * \boldsymbol{y}$ of a sequence $\boldsymbol{x} = \langle x_0,\ldots, x_{n-1}\rangle$ of length $n$ and a sequence $\boldsymbol{y} = \langle y_0,\ldots, y_{m-1}\rangle$ of length $m$ may be defined formally as

$$\boldsymbol{x} \cup \{\langle y_i, n+i\rangle \mid i < m\}.$$

We say that $\boldsymbol{x}$ is an *initial segment* of $\boldsymbol{x} * \boldsymbol{y}$; a *proper initial segment* if $\boldsymbol{y} \neq \varepsilon$.

We often write sequences of elements that we think of as symbols in the form of *words*, without brackets, commas or blanks. For example, instead of

$$\langle w, o, r, d\rangle$$

we would write *word*. Concatenation of words is expressed by mere juxtaposition: if $v$ denotes the word *by*, and $w$ denotes *pass*, then $v * w$, that is, the word *bypass*, is denoted by $vw$. A thing $x$ *occurs* in a word $a_0a_1\ldots a_{n-1}$ if there is some relevant index $i$, that is, $i < n$, such that $a_i = x$. An *occurrence* of $x$ in $a_0a_1\ldots a_{n-1}$ is an $i < n$ such that $a_i = x$. For example, $s$ has two occurrences, 4 and 5, in *bypass*. We shall often identify the occurrences of $x$ in the $v$-segment of $vw$ with the corresponding occurrences in $v$, and those in the $w$-segment

[4] The latter notation is short for $\{x \in f \mid \exists u \in U.\, x = \langle f(u), u\rangle\}$.

with the corresponding occurrences in $w$. That is, even though officially the occurrences of $s$ in *pass* are 2 and 3, and those in *bypass* 4 and 5, we shall attempt to get by with saying that the occurrences of $s$ in *bypass* are precisely those in *pass*.

We generalize our notation for sequences to infinite sets of indices, and to sets without an obvious ordering, as follows. Let $A$ be a class, and $I$ a set. A *family* of elements of $A$ *indexed by* $I$ is a mapping $\boldsymbol{a}: I \longrightarrow A$, usually written as $\langle a_i \,|\, i \in I\rangle$, or $\langle a_i\rangle_{i\in I}$, or even, if $I$ is sufficiently fixed, $\langle a_i\rangle_i$. The range $\{a_i \,|\, i \in I\}$ of $\boldsymbol{a}$ may be denoted by $\{a_i\}_{i\in I}$, or $\{a_i\}_i$. For example, as a $\{0,\ldots, n-1\}$-indexed family, $\langle x_0,\ldots, x_{n-1}\rangle$ would be written $\langle x_i \,|\, i < n\rangle$ or $\langle x_i\rangle_{i<n}$; its range is $\{x_i \,|\, i < n\}$ $(= \{x_i\}_{i<n})$.

We define the union and intersection of a family of sets $\langle X_i \,|\, i \in I\rangle$ as the union (intersection) of its range; we write

$$\bigcup(X_i \,|\, i \in I), \quad \bigcup_{i\in I} X_i, \text{ or } \textstyle\bigcup_i X_i,$$

and likewise for intersection. Thus, for example, $X^{<\omega} = \bigcup_{n\in\mathbb{N}} X^n$.

The generalisation from pairs to families corresponds with a generalisation of cartesian products. The cartesian product $\prod(X_i \,|\, i \in I)$ of a family of sets $\langle X_i \,|\, i \in I\rangle$ is the set of all families $\langle x_i \,|\, i \in I\rangle$ with the property that for each $i \in I$, $x_i \in X_i$. So these families *choose* an element of $X_i$ for every index $i$; hence they are called *choice functions*.

Alternative notations for the product are

$$\prod_{i\in I} X_i \quad \text{and} \quad \textstyle\prod_i X_i.$$

Instead of $\prod(X_i \,|\, i < n)$ and the like, we also write

$$X_0 \times\ldots\times X_{n-1};$$

$\prod(X_i \,|\, i < 1)$ we identify with $X_0$, and $\prod(X_i \,|\, i < 0) = \prod\emptyset = \{\emptyset\}$. In the special case that every $X_i$ is the same set $X$, we speak of a cartesian *power*, and write $X^I$: indeed, in this case the cartesian product is simply the set of all mappings from $I$ into $X$.

Suppose $j \in I$. The *projection* $\pi_j$ of $\prod(X_i \,|\, i \in I)$ onto the $j$-factor $X_j$ is defined by

$$\pi_j(x_i \,|\, i \in I) = x_j.$$

Observe that this is the same as saying that for every choice function $f$, $\pi_j(f) = f(j)$. For arbitrary pairs we define projections $\pi$ and $\pi'$ by

$$\pi(x, y) = x \;\text{ and }\; \pi'(x, y) = y.$$

In some particular cases, special notations are in use. For example, the set $\mathbb{C}$ of *complex numbers* may be defined as $\mathbb{R} \times \mathbb{R}$; the first component $\pi(z)$ of a complex number $z$ is called its *real* part, and denoted by $\mathrm{Re}(z)$, and the second component, $\pi'(z)$, is its *imaginary* part $\mathrm{Im}(z)$.

**h3 The Axiom of Choice**

Two of the axioms of set theory that we discussed so far — Extensionality and Regularity — express essential characteristics of classes; the rest is of the form 'construction $\Gamma$, applied under conditions $\Upsilon$, results in a set'.

Thus they are *constructive* in the sense that they all support a claim that something exists with a recipe for building it. The constructive axioms imply that the cartesian product $\prod_i X_i$ of a family $\langle X_i \mid i \in I \rangle$ of sets exists, in the universe of sets. They also imply that a *finite* product of nonvoid sets has elements. They do not tell us much, however, about the infinite case. For this we need a new assumption, the *Axiom of*

***Choice***: Every family of nonvoid sets has a choice function.

Equivalently: any cartesian product of nonvoid sets is nonvoid.

Russell illustrated the non-constructive character of this axiom as follows. Intuitively a function is given by a rule that specifies for every element of its domain the corresponding value. So if we have pairs of shoes, numbered from 0 to infinity, then 'take the left shoe' specifies a function that chooses an element from every unordered pair of shoes. For infinitely many pairs of *socks* there is no such rule. The choice of socks must be left to the chooser's initiative.

We finish with a simple application of the Choice Axiom.

**Proposition**. Let $\langle X_i \mid i \in I \rangle$ be a family of sets, and $j \in I$. The projection $\pi_j$ is surjective if and only if either $X_j$ is void, or $X_i$ is nonvoid for all $i \in I$.

**Proof**. ($\Rightarrow$) If any $X_i$ is void, then so is $\prod_i X_i$, hence $\pi_j$ is void, and if $X_j$ is nonvoid, $\pi_j$ is not surjective.
($\Leftarrow$) If $x \in X_j$, define $\langle Y_i \mid i \in I \rangle$ by

$$Y_j = \{x\};\ Y_i = X_i \text{ if } i \neq j.$$

If $X_i$ is nonvoid for all $i \neq j$, then by the Axiom of Choice $\prod_i Y_i$ has an element $\boldsymbol{y}$. Since $Y_i \subseteq X_i$ for all $i \in I$, $\boldsymbol{y} \in \prod_i X_i$ as well, and $\pi_j(\boldsymbol{y}) = \boldsymbol{y}(j) = x$. ⊠

**h4 Operations and relations of arbitrary finite arity**

An operation $f$ whose domain consists of sequences of a fixed finite length $n$ is called an *n-ary operation*. For small $n$ actual words are current: nullary for 0-ary, unary for 1-ary, and so on (binary, ternary, quaternary,…). We write $f(x_0,\ldots,x_{n-1})$ instead of $f(\langle x_0,\ldots,x_{n-1}\rangle)$. If $f \subseteq X \times X^n$, we call $f$ an $n$-ary operation *over* or *on* $X$. Such an operation is *total* if $\mathrm{Dom} f = X^n$. A nullary operation is usually called a *constant*. Officially, a nullary operation $f$ is either void or a singleton $\{\langle c, \varepsilon \rangle\}$; but it is commonly identified with the object $c = f(\varepsilon)$, if such an object exists.

A similar generalization of the relation concept defines $n$-ary relations to be classes of sequences of length $n$, and $n$-ary relations over (or *on*) a class $X$ as subclasses of $X^n$. Binary relations are identified with the relations of §e. As with binary relations, instead of $\langle x_1,\ldots, x_n \rangle \in R$ we also write $R(x_1,\ldots, x_n)$ or $Rx_1\ldots x_n$.

If $R$ is an $n$-ary relation and $B$ is a class, then we sometimes abbreviate $R \cap B^n$ to $R_B$. We call $R_B$ the *limitation* of $R$ to $B$. Similarly an $n$-ary operation $f$ has a limitation $f_B = f \cap (B \times B^n)$.

The notation $R[Y]$ of §e has a useful generalization to relations. Let $\mathcal{D}$ be a class, $R$ a binary relation on $\mathcal{D}$, and $P$ an $n$-ary relation ($n \geq 1$) on $\mathcal{D}$. Then $R[P]$ is

$$\{\langle y_1,\ldots,y_n\rangle \mid \exists x_1,\ldots,x_n(\langle y_1,x_1\rangle,\ldots,\langle y_n,x_n\rangle \in R \;\&\; \langle x_1,\ldots,x_n\rangle \in P\}.$$

In particular, if $f$ is a function then

$$f[P] = \{\langle f(x_1),\ldots,f(x_n)\rangle \mid \langle x_1,\ldots,x_n\rangle \in P\},$$

and

$$f^{-1}[P] = \{\langle x_1,\ldots,x_n\rangle \mid \langle f(x_1),\ldots,f(x_n)\rangle \in P\}.$$

## Exercises

**§c**

Show that $\{\{x\},\{x,y\}\} = \{\{u\},\{u,v\}\}$ if and only if $x = u$ and $y = v$.

**§d**

Let $K = \{x \mid x \notin x\}$. Prove that $\{K\}$ is not a set.

**§e**

**1**. Prove: if $R$ is a set, then Dom $R$ and Ran $R$ are sets as well.

**2**. Prove, *without using the Union Postulate*: if $X$ is a set of sets, then $\bigcup X$ is a set.

**3**. Let $x$ be a set. Prove that $x \notin x$.

**§f**

Use Kuratowski's definition of ordered pairs to prove that $X \times Y$ is a set if $X$ and $Y$ are sets.

**§g**

**1**. Prove Proposition 1.

**2**. Prove Proposition 2.

**3**. Prove that composition of relations is *associative*: $Q \circ (R \circ S) = (Q \circ R) \circ S$. (Because of this, parentheses in composition-terms may be omitted.)

**4**. Consider each of the following statements about a binary relation $R$, and either prove it or give a counterexample:
(a) If $X = \mathrm{Fld}\,R$, then $\Delta_X \subseteq R \circ R^{-1}$;
(b) If $X = \mathrm{Ran}\,R$, then $\Delta_X \subseteq R \circ R^{-1}$;
(c) $R \subseteq R \circ R^{-1} \circ R$;
(d) $R = R \circ R^{-1} \circ R$.

**5**. If the relations $R$ and $S$ are sets, then why is $R \circ S$ a set?

**§h**

**1**. Let $f$, $g$, $h$ be operations.
(i) Prove: if $\mathrm{Ran}f = \mathrm{Dom}g = \mathrm{Dom}h$ and $g \circ f = h \circ f$, then $g = h$.
(ii) Prove: if $f$ is injective, $\mathrm{Ran}g \cup \mathrm{Ran}h \subseteq \mathrm{Dom}f$, and $f \circ g = f \circ h$, then $g = h$.

**2**. Prove Proposition 1.

**3**. Let $f\colon X \longrightarrow Y$ and $g\colon Y \longrightarrow Z$ be mappings. Prove:
(i) if $g \circ f\colon X \longrightarrow Z$ is surjective, then $g$ is surjective;
(ii) if $g \circ f\colon X \longrightarrow Z$ is injective, then $f$ is injective.
(iii) if $g \circ f\colon X \longrightarrow Z$ is injective, and $f$ is surjective, then $g$ is injective.

**4**. Let $\theta$ be the relation defined on $\mathbb{N} \times \mathbb{N}$ by: $\langle x,y\rangle\theta\langle u,v\rangle$ if and only if $x + v = u + y$ (§g). Prove that the mapping $\langle x,y\rangle/\theta \longmapsto x - y$ is a bijection from $(\mathbb{N} \times \mathbb{N})/\theta$ onto $\mathbb{Z}$.

**5**. Let $X$ be any set.
(a) Prove that $|X^{\emptyset}| = 1$.
(b) Prove that $|\emptyset^{X}| = 0$ if $X$ is nonvoid.

**6**. Let $X$ and $Y$ be finite sets; $|X| = m$ and $|Y| = n$. Prove that $|X^{Y}| = m^{n}$.

**7**. Let $X$, $Y$ and $Z$ be sets, $X$ and $Y$ disjoint. Show that $Z^X \times Z^Y$ and $Z^{X \cup Y}$ are equipollent. Now let $m$ and $n$ be natural numbers; conclude that $Z^m \times Z^n$ and $Z^{m+n}$ are equipollent.

**8**. (a) Prove that $\mathbb{N}$ and $\mathbb{Z}$ are equipollent.
(b) Prove that $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ are equipollent.

**9**. Let $f$: $A \longrightarrow B$ be a mapping, $X \subseteq A$ and $Y \subseteq B$.
(a) Verify that $ff^{-1}[Y] \subseteq Y$.
(b) Show that $ff^{-1}[Y] = Y$ for all $Y \subseteq B$ if and only if $f$ is surjective.
(c) Give an example to show that not necessarily $f^{-1}f\,[X] \subseteq X$.

**10**. Verify that the restriction of a mapping $f$: $X \longrightarrow Y$ to a class $U \subseteq X$ may be constructed by composition with the canonical embedding: $f{\restriction}U = f \circ 1_X^U$.

**11**. Prove that concatenation of sequences is *associative*: if $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X^{<\omega}$, then

$$(\boldsymbol{x} * \boldsymbol{y}) * \boldsymbol{z} = \boldsymbol{x} * (\boldsymbol{y} * \boldsymbol{z}).$$

(Use the formal definition of sequences.)

**12**. Why cannot there be an infinite sequence

$$x_0 \ni x_1 \ni x_2 \ni x_3 \ni \ldots$$

of sets?

**13**. (a) Define, for $n \in \mathbb{N}$, $[n)$ as $\{m \in \mathbb{N} \,|\, n \leq m\}$. Observe that $\langle [n) \,|\, n \in \mathbb{N} \rangle$ is an $\mathbb{N}$-indexed family. Simplify

$$\bigcap_{n \in \mathbb{N}} [n) .$$

(b) For every positive integer $n$, $(-1/n,\, 1/n) \subseteq \mathbb{R}$ is defined as $\{x \in \mathbb{R} \,|\, -1/n < x < 1/n\}$. Observe that $\langle (-1/n,\, 1/n) \,|\, n \in \mathbb{Z}_+ \rangle$ is a $\mathbb{Z}_+$-indexed family. Simplify

$$\bigcap_{n \in \mathbb{Z}_+} (-1/n,\, 1/n) .$$

**14**. Show that the cartesian product $\prod_i X_i$ of a family $\langle X_i \,|\, i \in I \rangle$ of sets is a set.

**15**. (a) Let $X$ be any set. Recall that by definition, $X^1$ consists of all the functions from $\{0\}$ into $X$; and for any such function $\boldsymbol{x}$, $\pi_0(\boldsymbol{x}) = \boldsymbol{x}(0)$. Construct a bijection $\phi$: $X \longrightarrow X^1$ such that for all $x \in X$, $\pi_0(\phi(x)) = x$.
(b) Construct a bijection $\phi$: $X_0 \times X_1 \longrightarrow \prod(X_i \,|\, i = 0,\, 1)$ such that for all $w \in X_0 \times X_1$, $\pi_0(\phi(w)) = \pi(w)$ and $\pi_1(\phi(w)) = \pi'(w)$.
(c) Construct a bijection $\phi$: $\prod(X_i \,|\, 0 \leq i < m) \times \prod(X_j \,|\, m \leq j < n) \longrightarrow \prod(X_k \,|\, 0 \leq k < n)$ such that for all $w \in \prod(X_i \,|\, 0 \leq i < m) \times \prod(X_j \,|\, m \leq j < n)$, and all $i < m$, $m \leq j < n$, $\pi_i(\phi(w)) = \pi_i(\pi(w))$ and $\pi_j(\phi(w)) = \pi_j(\pi'(w))$.
In view of these correspondences, iterated products $X_0 \times \ldots \times X_{n-1}$ are written without parentheses, and usually identified with $\prod(X_i \,|\, 0 \leq i < n)$.

**16**. A *transversal* of a block system $\mathcal{B}$ is a class $C$ such that
(i) $C \subseteq \bigcup \mathcal{B}$;
(ii) for all $B \in \mathcal{B}$, $|B \cap C| = 1$.
Prove that the Axiom of Choice is equivalent to the statement

Every partition has a transversal.

**17**. A *representation* of a partition $\mathcal{A}$ of a set $X$ is a mapping $f$: $X \longrightarrow X$ such that for all $x, y \in X$, $x \equiv f(x) \pmod{\theta_{\mathcal{A}}}$, and if $x \equiv y \pmod{\theta_{\mathcal{A}}}$, then $f(x) = f(y)$. Prove that the Axiom of Choice is equivalent to the statement

Every partition has a representation.

**18**. Let $f$ be a unary operation on a set $B$, and $A \subseteq B$. Is it necessarily the case that $f{\restriction}A = f_A$?

**19**. Let $X$ be a set. Show that $\mathcal{P}(X) \nsubseteq X$.

**20**. Let $M$ be a set. Show that $\mathcal{P}(\bigcup M) \notin M$.

**21**. Let $M$ be a set, $a$ anything. Show that $\{a, \mathcal{P}(\bigcup\bigcup M)\} \notin M$.

**22**. Let $M$ be a set, $a$ anything. Show: if $\{a, \mathcal{P}(\bigcup\bigcup M)\} = \{a', \mathcal{P}(\bigcup\bigcup M)\}$, then $a = a'$.

**23**. Let $\mathcal{U}$ be the universe, the class of all sets and all primitive elements. Let $M$ be a set. Show that the assignment $a \longmapsto \{a, \mathcal{P}(\bigcup\bigcup M)\}$ is an injective mapping of $\mathcal{U}$ into $\mathcal{U} - M$.

**24** (Pickert-Van der Waerden Lemma). Let $A$ and $M$ be sets, and $i$: $M \hookrightarrow A$ an injective mapping. Construct a set $B$ and a bijection $j$: $B \hookrightarrow\!\!\!\!\!\rightarrow A$ such that $(A \cup M) \cap B = \emptyset$.

Let $C := M \cup (B - j^{-1}i[M])$. Show

$$i \cup j\restriction(B - j^{-1}i[M]) \colon C \hookrightarrow\!\!\!\!\!\rightarrow A.$$

**The natural numbers**

If the natural numbers are not given in advance, we must postulate that there are sufficiently many sets. Let us call a set *inductive* if it contains the void set $\emptyset$, and with every set $x$ that it contains, also contains $x \cup \{x\}$. Then the *Axiom of Infinity* runs:

There exists an inductive set.

We define $\omega$ as the intersection of all inductive sets; it will figure as a substitute for $\mathbb{N}$.

**25**. Show that $\omega$ is inductive.

Since $\omega$ is the least inductive set, we have mathematical induction in the following form:

If a set $X$ is inductive, then $\omega \subseteq X$.

The elements of $\omega$ are $\emptyset$, which we may call 0; $0 \cup \{0\} = \{0\}$, which is called 1; $1 \cup \{1\}$, called 2, and so on. Observe that $0 \in 1, 2, 3,\ldots$; $1 \in 2, 3,\ldots$, and so on: $\in$ on $\omega$ corresponds with $<$ on $\mathbb{N}$. Every $n \in \omega$ is at the same time a subset of $\omega$, consisting of the first $n$ elements of $\omega$.

A set is finite if it is equipollent with an initial segment $\{0,\ldots, n - 1\}$ of $\mathbb{N}$. If instead of $\mathbb{N}$ we use $\omega$, this becomes: a set $x$ is finite if and only if there exists a surjection onto $x$ from an element of $\omega$.