§ 3
Mappings and Operations

Functions, also called mappings, build a very important type of relations. Let us recall that a relation from A into B is a subset of $A \times B$. Under special circumctances, a relation will be called a function or a mapping. These two terms will be used interchangably.

3.1 Definition: Let A and B be nonempty sets. A relation f from A into B is called a *function from A into B*, or a *mapping from A into B* if every element of A is the first component of a single ordered pair in $f \subseteq A \times B$.

This definition embraces two conditions. First, every element a of A will appear as the first component of at least one ordered pair $(a,*)$ in f, that is, the first components of the ordered pairs in f should make up the whole A. No element of A can be left out. There should be no element of A which is not the first component of any pair in f. Second, for any $a \in A$, there can be only one ordered pair in f whose first component is a. In other words, if (a,b) and (a,b') are both in f, these pairs should be identical, which means $b = b'$. A relation f from A into B is a mapping if and only if every element of A is the first component of one and only one ordered pair in f.

If f is a mapping from A into B, then A is called the *domain* of f, and B is called the *range* of f. A function f from A into B must be thought of as a rule or mechanism by which elements of A are assigned to certain elements of B. The first condition, that every element of A is the first component of at least one ordered pair in f, is a formal way of expressing that elements of A, not of any other set, in particular not of any proper subset of A, are the objects that are assigned (to some elements of B). The second condition, that every element of A is the first component of at most one ordered pair in f, is a formal way of expressing that no element of A is assigned to two, three or more elements of B.

We introduce some notation. We write $f\colon A \to B$ to mean that f is a mapping from A into B. Occasionally, we write $A \xrightarrow{f} B$. The reader probably expects that we write $f(a) = b$ in place of $(a,b) \in f$. This is the symbolism that the reader is accustomed to, and reminds us of a mapping rule that assigns b to a. However, we will rarely write $f(a) = b$. We prefer to write $(a)f = b$ or $af = b$, with the function symbol f on the right side of the element a. This might seem odd, and the reader might wonder about this strange order of elements and functions. It takes some time to get accustomed to this way of writing functions on the right, but the advantages of this notation will far outweigh the little trouble it causes at first. This will be amply clear in the sequel. We remark that not every algebraist conforms to this usage, and an isolated notation will have different meanings according as whether the functions are written on the right or on the left. We will point out these differences as occasions arise.

Suppose f is a mapping from A into B and $a \in A$ and $b \in B$ are such that $af = b$ (in this case, we sometimes write $a \rightarrow b$ or $a \xrightarrow{f} b$ and say that f *maps* a to b). Then b is called the *image of a under f*. We also say a is a *preimage* or an *inverse image of b under f*. Please mark the articles: b is the image of a, since a has one and only one image, but a is a preimage of b, for b may have many preimages.

3.2 Examples:(a) Let A be a nonempty set and let $\iota = \{(a,a)\colon a \in A\} \subseteq A \times A$. Then ι is a function from A into A. In our second notation, this reads $a\iota = a$. This function is called the *identity mapping on A*. When we want to point out the set A, we write ι_A instead of ι.

Now let $A \subseteq B$ and put $\mu = \{(a,a) \in A \times B : a \in A\} \subseteq A \times B$. Then μ is a function from A into B. In our second notation, this reads $a\mu = a$. This function is called the *inclusion mapping from A into B*. Writing $a\mu$ for a is a formal way of recalling $A \subseteq B$ and $a \in B$.

(b) Let $A = \{1,2,3,4,5\}$ and $B = \{a,b,c,d\}$. Consider

$$f = \{(1,b), (2,a), (4,d), (5,d)\}.$$

Then f is not a function from A into B since $3 \in A$ is not the first component of any ordered pair in $f \subseteq A \times B$. Consider

$$g = \{(1,b), (2,a), (3,a), (3,b), (4,c), (5,d)\}.$$

Then g is not a function from A into B since $3 \in A$ is the first component of two distinct ordered pairs in $g \subseteq A \times B$.

(c) Let A and B be two nonempty sets and let $b \in B$ be a fixed element of B. Then f, defined by

$$af = b \text{ for all } a \in A, \text{ i.e., } f = \{(a,b) \in A \times B\colon\ a \in A\}$$

is a mapping from A into B. This is sometimes called the *constant function b.*

(d) For any $(a,b) \in \mathbb{R} \times \mathbb{R}$, put $(a,b)s = a + b$. Then s is a function from $\mathbb{R} \times \mathbb{R}$ into $\mathbb{R}$. This s may be called the *sum* function. It is an example of a binary operation. We will examine binary operations later in this paragraph.

(e) Let $A = \{u,x,y,z\}$ and $B = \{1,2,3\}$, and put

$$uf = 1,\ xf = 2,\ yf = 2,\ zf = 1.$$

Then f is a function from A into B.

(f) Let A be a nonempty set and let $\boldsymbol{S}$ be the set of all subsets of A. For any $a \in A$, put $af = \{a\} \in \boldsymbol{S}$. Then f is a function from A into $\boldsymbol{S}$.

(g) Put $xf = x^2$ for all $x \in \mathbb{R}$. Then f is a function from $\mathbb{R}$ into $\mathbb{R}$.

(h) Consider $f = \{(x,y) \in \mathbb{R} \times \mathbb{R}\colon x^2 = y^2\}$. Then f is not a function from $\mathbb{R}$ into $\mathbb{R}$, since 1, for example, is the first component of two distinct ordered pairs $(1,1)$ and $(1,-1)$ in f. On the other hand, if $\mathbb{R}^+$ denotes the set of positive real numbers, then $g = \{(x,y) \in \mathbb{R}^+ \times \mathbb{R}^+\colon x^2 = y^2\}$ is a function from $\mathbb{R}^+$ into $\mathbb{R}^+$. In fact, g is the identity function on $\mathbb{R}^+$.

(i) Let $f\colon A \longrightarrow B$ be a mapping from A into B and let A_1 be a nonempty subset of A. For any $a \in A_1$, we put $ag = af$. Then g is a mapping from A_1 into B. In terms of ordered pairs, we have

$$g = f \cap (A_1 \times B).$$

g is called the *restriction of f to* A_1. We usually write f_{A_1} or $f_{|A_1}$ to denote the restriction of f to A_1. If g is a restriction of f to a subset of the domain of f, then f is called an *extension of g*.

(j) Let A be a nonempty set and let B be a fixed subset of A. For any a in A, we put

$$\chi_B = \begin{cases} 0 \text{ if } a \notin B \\ 1 \text{ if } a \in B. \end{cases}$$

Then χ_B is a function from A into $\{0,1\}$. It is called the *characteristic function of* B. Here we wrote the function on the left.

(k) For any $x \in \mathbb{R}$, we put

$$f(x) = \begin{cases} 0 \text{ if } x \text{ is irrational} \\ 1 \text{ if } x \text{ is rational.} \end{cases}$$

Then f is a function from $\mathbb{R}$ into $\mathbb{R}$. In fact, f is the characteristic function of the set of rational numbers. The image of some x is not known. For instance, it is not known whether Euler's constant γ is rational or not. Nevertheless, f is a genuine function. This example is due to L. Dirichlet (1805-1859).

(l) Let A be a nonempty set and let $\sim$ be an equivalence relation on A. Let $A/\sim$ be the set of equivalence classes under $\sim$. Then

$$\begin{aligned} \nu\colon A &\to A/\sim \\ a &\to [a] \end{aligned}$$

is a mapping from A into $A/\sim$. It is called the *natural mapping* or the *canonical mapping from A into* $A/\sim$.

3.3 Definition: Let $f\colon A \to B$ and $f_1\colon A_1 \to B$ be two functions. f and f_1 are called *equal* if $A = A_1$ and $af = af_1$ for all $a \in A = A_1$.

So, in order that two functions f and f_1 be equal, their domains must be equal and the images of any element in this common domain under the mappings f and f_1 must be equal, too. In particular, if $f\colon A \to B$ is a function and $B \subseteq C$, then the function g, defined by $ag = af$ for all $a \in A$, is equal to f. The ranges do not play any role in the definition of equality. (In some branches of mathematics, for example in topology, two functions with different ranges are sometimes considered distinct, even if their domains and functional values coincide.)

In the definition of a mapping $f\colon A \longrightarrow B$, we required that every element of A be the first component of at least one ordered pair in f and also that every element of A be the first component of at most one ordered pair in f. There was no analogous requirement for the elements of B. If we impose similar conditions on the elements of B, we get special types of functions, which we now introduce.

3.4 Definition: Let $f\colon A \longrightarrow B$ be a mapping. If every element of B is the second component of at least one ordered pair in f, then f is called a mapping from A *onto* B.

The reader must be careful about the usage of the prepositions "into" and "onto", for they are used with different meanings. That f is a function from A onto B means that every element of B is the image of some element of A. For an arbitrary mapping $f\colon A \longrightarrow B$, an element of B has perhaps no preimage at all, but if f is a mapping from A onto B, then each element of B has at least one preimage in A.

The range should be specified whenever the term "onto" is used. A function is not "onto" by itself, it is only onto a specific set. We shall frequently treat the word "onto" as an adjective, but it will be always clear from the context which range set is meant.

3.5 Examples: (a) The mapping $f\colon \mathbb{R} \longrightarrow \mathbb{R}$, given by $f(x) = x^2$ for all $x \in \mathbb{R}$, is not onto, since $-1 \in \mathbb{R}$, for instance, has no preimage under f.

(b) Let $\mathbb{R}^+$ denote the set of all positive real numbers. Then the mapping $f\colon \mathbb{R} \rightarrow \mathbb{R}^+$, given by $f(x) = x^2$ for all $x \in \mathbb{R}$, is onto.

(c) The mapping $g\colon \{1,2,3,4,5\} \longrightarrow \{a,b,c\}$, given by

$$1g = a,\ 2g = a,\ 3g = a,\ 4g = b,\ 5g = c$$

is onto.

(d) Let A be any nonempty set. Then $\iota_A : A \longrightarrow A$ is onto, for any $a \in A$ has a preimage a in A under ι_A since $a\iota_A = a$.

3.6 Definition: Let $f: A \longrightarrow B$ be a mapping. If every element of B is the second component of at most one ordered pair in f, then f is called a *one-to-one* mapping from A into B.

A function $f: A \longrightarrow B$ is therefore one-to-one if an arbitrary element of B has either no preimage in A or exactly one preimage: any two preimages of $b \in B$ (if b has a preimage at all) must be equal. So the necessary and sufficient condition for a mapping $f: A \longrightarrow B$ to be one-to-one is

$$af = b \text{ and } a_1 f = b \quad \Longrightarrow \quad a = a_1 \qquad (a, a_1 \in A,\ b \in B)$$

or, more shortly

$$af = a_1 f \quad \Longrightarrow \quad a = a_1 \qquad (a, a_1 \in A),$$

whose contrapositive reads

$$a \neq a_1 \quad \Longrightarrow \quad af \neq a_1 f \qquad (a, a_1 \in A).$$

A one-to-one mapping is a mapping by which different elements in the domain are matched with different elements in the range. Being a one-to-one function is the negation of being a "many-to-one" function, by which many elements in the domain are matched with one element in the range.

3.7 Examples: (a) $\{(x,y): x^2 = y\} \subseteq \mathbb{R} \times \mathbb{R}$ is not a one-to-one function from $\mathbb{R}$ into $\mathbb{R}$, for two distinct elements x and $-x$ (if $x \neq 0$) have the same image.

(b) Let $\mathbb{R}^+$ denote the set of all positive real numbers. Then the mapping $\{(x,y): x^2 = y\} \subseteq \mathbb{R}^+ \times \mathbb{R}^+$ is a one-to-one function from $\mathbb{R}^+$ into $\mathbb{R}^+$.

(c) The mapping $g: \{1,2,3\} \longrightarrow \{a,b,c,d\}$, given by

$$1g = b,\ 2g = d,\ 3g = a$$

is one-to-one.

(d) Let A be a nonempty set. Then $\iota_A\colon A \to A$ is one-to-one, for if $a\iota_A = b\iota_A$, then $a = b$ from the definition of ι_A.

Suppose we have two functions $f\colon A \to B$ and $g\colon B \to C$. For any $a \in A$, we find $af = b \in B$ and then apply g to this element $af = b$ of B. We get an element $c = bg$ of C. In this way, the element a of A is assigned to an element c of C. Here $af = b$ is uniquely determined by f (since f is a mapping) and $bg = c$ is uniquely determined by g (since g is a mapping). So c is uniquely determined: we have a a mapping from A into C.

3.8 Definition: Let $f\colon A \to B$ and $g\colon B \to C$ be two functions. Then

$$h = \{(a,(af)g) \in A \times C : a \in A\} \subseteq A \times C,$$

which is a function from A into C, is called the *composition of f with g*, or the *product of f by g*.

We write $h = f \circ g$ or more simply $h = fg$. Thus $a(fg)$ is defined as $(af)g$.

In order to compose two functions f and g, we must make sure that the range of the first function f is a subset of the domain of the second function g. Otherwise, their composition is not defined. Note the order of the functions f and g. We apply f first, then g; and we write first f, then g in the composition notation fg. One of the advantages of writing the functions on the right becomes evident here. If we had written the functions on the left, then fg would have meant: first apply g, then f [as in the calculus, where $(f \circ g)(x) = f(g(x))$] and we would have been reading backwards. Notice also that the domain of fg is the domain of f.

3.9 Examples: (a) Let $f\colon A \to B$ be a mapping. Then it is easily seen that $f\iota_B = f$ and $\iota_A f = f$. Indeed, the domains of $f\iota_B$, f, $\iota_A f$ are all equal to A and

$$a(f\iota_B) = (af)\iota_B = af \quad \text{and} \quad a(\iota_A f) = (a\iota_A)f = af$$

for all $a \in A$. In particular, if $g\colon A \to A$ is a mapping, then $g\iota_A = g = \iota_A g$.

(b) Let $f\colon \{1,2,3,4\} \to \{a,b,c,d\}$ and $g\colon \{a,b,c,d\} \to \{5,x,U,\xi,\eta\}$ be given by

$1f = a \qquad ag = U$

$$\begin{array}{ll} 2f = c & bg = x \\ 3f = d & cg = \eta \\ 4f = b & dg = 5 \end{array}$$

Then we have

$$\begin{array}{llll} 1(fg) = & (1f)g & = ag & = U \\ 2(fg) = & (2f)g & = cg & = \eta \\ 3(fg) = & (3f)g & = dg & = 5 \\ 4(fg) = & (4f)g & = bg & = x. \end{array}$$

(c) Given $f\colon \{1,2,3\} \longrightarrow \{a,b\}$ and $g\colon \{a,b\} \longrightarrow \{x,y,z\}$, where

$$\begin{array}{lll} f\colon 1 \longrightarrow a & \text{and} & g\colon a \longrightarrow y \\ 2 \longrightarrow a & & b \longrightarrow z \\ 3 \longrightarrow b & & c \longrightarrow z, \end{array}$$

we have

$$\begin{array}{l} fg\colon 1 \longrightarrow y \\ 2 \rightarrow y \\ 3 \longrightarrow z. \end{array}$$

Notice that gf is not defined.

(d) Given $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ and $g\colon \mathbb{R} \longrightarrow \mathbb{R}$, we have
$x \longrightarrow \sin x$ $\qquad$ $x \longrightarrow x^2$

$$x(fg) = (xf)g = (\sin x)g = (\sin x)^2 = \sin^2 x,$$

$$x(gf) = (xg)f = (x^2)f = \sin(x^2).$$

(e) Given $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ and $g\colon \mathbb{R} \longrightarrow \mathbb{R}$, we have
$x \longrightarrow x^2\text{-}1$ $\qquad$ $x \longrightarrow x^2+1$

$$x(fg) = (xf)g = (x^2-1)g = (x^2-1)^2 + 1 = x^4 - 2x^2 + 2$$

$$x\,(gf) = (xg)f = (x^2+1)f = (x^2+1)^2 - 1 = x^4 + 2x^2.$$

Given two functions $f\colon A \longrightarrow B$ and $g\colon B \longrightarrow C$, we might be tempted to ask whether $fg = gf$. Example 3.9(b) and Example 3.9(c) tell us that this question is meaningless, for, although fg is defined in these examples, gf is not even defined, let alone is equal to fg. Example 3.9(d) and Example 3.9(e) show that the two functions fg and gf, even if they both exist, are not necessarily equal. We have $fg \neq gf$ in general: the composition of mappings is not commutative.

However, it is associative.

3.10 Theorem: *Let* $f\colon A \longrightarrow B$, $g\colon B \longrightarrow C$, $h\colon C \to D$ *be three functions. Then* $(fg)h = f(gh)$.

Proof: We must prove that the domains of $(fg)h$ and $f(gh)$ are equal and that an arbitrary element in the common domain is assigned to the same element by $(fg)h$ and by $f(gh)$.

The domain of $(fg)h$ is the domain of fg, which is the domain of f, which is A. The domain of $f(gh)$ is the domain of f, which is A. So the domains of $(fg)h$ and $f(gh)$ coincide.

Now let a be an arbitrary element of A. Then

$$\begin{aligned} a\big((fg)h\big) &= \big(a(fg)\big)h && \text{(by the definition of } (fg)h\text{; forget that } fg \text{ is a composition itself)}\\ &= \big((af)g\big)h && \text{(recall now that } fg \text{ is a composition, applied to an element } a)\\ &= (af)(gh) && \text{(definition of } gh\text{, applied to an element } af)\\ &= a\big(f(gh)\big) && \text{(definition of } f(gh)), \end{aligned}$$

which yields $(fg)h = f(gh)$. □

Onto mappings and one-to-one mappings behave very nicely when they are composed.

3.11 Theorem: *Let* $f\colon A \longrightarrow B$, $g\colon B \longrightarrow C$ *be two functions and let* $fg\colon A \longrightarrow C$ *be their composition.*
(1) *If* f *is onto and* g *is onto, then* fg *is onto.*
(2) *If* f *is one-to-one and* g *is one-to-one, then* fg *is one-to-one.*

Proof: (1) Suppose f and g are onto. For any $c \in C$, we must find a preimage of c under fg.The only thing we know about C is that C is the range of g. Now g is onto, so c has a preimage in B under g. Let $b \in B$ be such that $bg = c$. Since $b \in B$ and B is the range of f, and f is onto, b has a preimage $a \in A$ under f, so that $af = b$. Then we get $a(fg) = (af)g = bg = c$.

So a is a preimage of c under fg. This proves that fg is onto. (Summary: a preimage of a preimage is a preimage that works.)

(2) Now suppose f and g are one-to-one. We must prove $a = a_1$ whenever $a(fg) = a_1(fg)$, for all $a, a_1 \in A$. Indeed, if

$$\begin{array}{lrcll} & a(fg) & = & a_1(fg), & \\ \text{then} & (af)g & = & (a_1f)g, & \\ & af & = & a_1f & \text{(since } g \text{ is one-to-one),} \\ & a & = & a_1 & \text{(since } f \text{ is one-to-one).} \end{array}$$

This proves that fg is one-to-one. □

The converse of Theorem 3.11 is wrong. If $f\colon A \to B$ and $g\colon B \to C$ are two functions and if $fg\colon A \to C$ is onto, it does not always follow that both f and g are onto. Also, if $f\colon A \to B$ and $g\colon B \to C$ are two functions and if $fg\colon A \to C$ is one-to-one, it does not always follow that both f and g are one-to-one. This can be read off from the functions displayed below.

$$\begin{array}{ccccc} \{a,b,c\} & \xrightarrow{f} & \{x,y,z\} & \xrightarrow{g} & \{1,2\} \\ a & & x & & 1 \\ b & & y & & 2 \\ c & & z & & \end{array} \qquad \begin{array}{ccccc} \{a,b\} & \xrightarrow{f_1} & \{x,y,z\} & \xrightarrow{g_1} & \{1,2,3\} \\ a & & x & & 1 \\ b & & y & & 2 \\ & & z & & 3 \end{array}$$

Here fg is onto, but f is not onto; and f_1g_1 is one-to-one, but g_1 is not one-to-one.

However, we have a partial result in this direction. Observe that g is onto and f_1 is one-to-one in these examples. This is not a coincidence.

3.12 Lemma: *Let $f\colon A \to B$, $g\colon B \to C$ be two functions and let $fg\colon A \to C$ be their composition.*
(1) *If fg is onto, then g is onto.*
(2) *If fg is one-to-one, then f is one-to-one.*

Proof: (1) Assume fg is onto. For any $c \in C$, we must find a preimage of c in B under g. Now any $c \in C$ has a preimage in A under fg. Let $c = a(fg)$, where $a \in A$. Then $c = (af)g$. So $af \in B$ is a preimage of c in B under g. This proves that g is onto. (Summary: the image of a preimage is a preimage that works.)

(2) Assume fg is one-to-one. We wish to prove that f is one-to-one. Suppose that $af = a_1f$, where $a, a_1 \in A$. Applying g to both sides of this equation, we get $(af)g = (a_1f)g$, therefore $a(fg) = a_1(fg)$. Since fg is one-to-one by hypothesis, we get $a = a_1$. This proves that $af = a_1f$ implies $a = a_1$. Thus f is one-to-one. □

In view of its importance, we record the most important corollary of Theorem 3.11 as a separate theorem.

3.13 Theorem: *Let $f\colon A \to B$, $g\colon B \to C$ be one-to-one and onto. Then the composition $fg\colon A \to C$ is one-to-one and onto.* □

Assume we have a mapping $f\colon A \to B$. We want to define a new mapping $g\colon B \to A$ by inverting the order of the components of the ordered pairs in f. In other words, we want to define g by putting $(b,a) \in g$ if and only if $(a,b) \in f$. This g is a relation from B into A. The question arises: when is g in fact a mapping from B into A?

The necessary and sufficient condition for g to be a mapping is that each element of B be the first component of at least one and at most one ordered pair in g. By the definition of g, this is equivalent to the condition that each element of B be the second component of at least one ordered pair in f (i.e., f be onto) and also of at most one ordered pair in f (i.e., f be one-to-one). Let us observe that the mapping g is then uniquely determined by

$$bg = a \text{ if and only if } af = b.$$

We proved the

3.14 Theorem: *Let $f\colon A \to B$ be a mapping. The following assertions are equivalent.*

(i) *f is one-to-one and onto.*

(ii) *There is a unique mapping $g\colon B \to A$ such that*

$$bg = a \quad \text{if and only if} \quad af = b \qquad (a \in A,\ b \in B)$$ □

3.15 Definition: The mapping g of Theorem 3.14 is called the *inverse mapping of* f, or simply the *inverse of* f. It is denoted by f^{-1}.

3.16 Theorem: *Let* $f: A \to B$ *be one-to-one and onto, and let* $f^{-1}: B \to A$ *be its inverse. Then* $ff^{-1} = \iota_A$ *and* $f^{-1}f = \iota_B$.

Proof: We must show that the domains and functional values coincide. The domain of ff^{-1} is the domain of f, which is A, and A is the domain of ι_A. Further, for any $a \in A$, we have $a(ff^{-1}) = (af)f^{-1} = a = a\,\iota_A$ by the definition of f^{-1} · This proves $ff^{-1} = \iota_A$.

The domain of $f^{-1}f$ is the domain of f^{-1}, which is B, and B is the domain of ι_B. Further, for any $b \in B$, we have

$$\begin{aligned} b(f^{-1}f) &= (bf^{-1})f = af \quad \text{(where } a \text{ is the unique element of } A \text{ with } af = b\text{)} \\ &= b = b\iota_B. \end{aligned}$$

This proves $f^{-1}f = \iota_B$. □

3.17 Theorem: (1) *Let* $f: A \longrightarrow B$ *be one-to-one and onto. Then* $f^{-1}: B \to A$ *is one-to-one and onto.*

(2) *Let* $f: A \longrightarrow B$ *be a mapping. If there is a mapping* $g: B \to A$ *such that* $fg = \iota_A$ *and* $gf = \iota_B$*, then* f *is one-to-one and onto (and therefore* g *is the the inverse of* f*).*

Proof: (1) We have $f^{-1}f = \iota_B$ by Theorem 3.16. Since ι_B is one-to-one (Example 3.7(d)), f^{-1} is one-to-one by Lemma 3.12(2). Also, we have $ff^{-1} = \iota_A$ by Theorem 3.16. Since ι_A is onto (Example 3.5(d)), f^{-1} is onto by Lemma 3.12(1).

(2) We use the same reasoning. $fg = \iota_A$ is one-to-one, so f is one-to-one, and $gf = \iota_B$ is onto, so f is onto. □

A mapping $f: A \longrightarrow B$ is said to be a *one-to-one correspondence between* A *and* B in case f is one-to-one and onto. If f is a one-to-one correspondence between A and B, then f^{-1} is a one-to-one correspondence between B and A by Theorem 3.17(1).

*

* *

We now introduce binary operations. They constitute a generalization of the four elementary operations addition, subtraction, multiplication and division that everybody learns in the primary school. Consider addition, for example. Given any two numbers a and b, their sum is a uniquely determined number. This is the core of the operation concept: given two objects a and b, associate with them a unique object of the same kind. More precisely, we have the

3.18 Definition: Let S be a nonempty set. A *binary operation on S* is a mapping from $S \times S$ into S.

The important thing about a binary operation λ is that it is defined for all ordered pairs $(a,b) \in S \times S$ and that the result of the operation, $(a,b)\lambda$, is an element of S.

Although a binary operation λ is a mapping, we will not employ the functional notation $(a,b)\lambda$. As in the case of the elementary operations, we write a sign like "+", "-", "$\circ$", "$\oplus$", "$\otimes$" between the elements a and b to denote the image of (a,b) under λ. So the image of (a,b) will be denoted by $a + b$, $a - b$, $a \circ b$, $a \oplus b$, $a \otimes b$ or by a similar symbol.

3.19 Examples: (a) The elementary operations addition, subtraction, multiplication are binary operations on $\mathbb{R}$. Subtraction is not a binary operation on $\mathbb{N}$, since $1 - 2$, for instance, is not an element of $\mathbb{N}$ (although 1 and 2 are).

(b) Let M be a set and let $\boldsymbol{S}$ be the set of all subsets of M. Taking union and taking intersection are binary operations on $\boldsymbol{S}$. The usual notation "$A \cup B$", "$A \cap B$" conforms to the remarks above.

(c) Let $\boldsymbol{F}$ be the set of all functions from a set A into A. The usual composition of functions is a binary operations on $\boldsymbol{F}$.

(d) Let us write $x \circ y = x + y^2$ and $x \vartriangle y = x^2+x+1$ for real numbers x,y. Then $\circ$ and $\vartriangle$ are binary operations on $\mathbb{R}$. Here y does not enter into

$x \vartriangle y$ in any way, but this does not preclude $\vartriangle$ from being a binary operation.

(e) Let V be the set of all vectors in the three space $\mathbb{R}^3$. Taking dot product of two vectors is not a binary operation on V, since the result is a scalar (real number), not a vector. On the other hand, taking cross product is a binary operation on V, since the result is a uniquely determined vector in V.

(f) For any natural numbers m,n, let $m \bullet n$ denote their (positive) greatest common divisor. Then $\bullet$ is a binary operation on $\mathbb{N}$.

(g) Let S be the set of all students in a classroom. For any students a,b in S, let $a{\cdot}b$ be that student who sits in front of a. Then $\cdot$ is not a binary operation on S, for $a{\cdot}b$ is not defined if a happens to sit in the foremost row. Remember that a binary operation on S has to be defined for all pairs in $S \times S$.

(h) For any ordered pairs (a,b), (c,d) of real numbers, we put

$$\begin{aligned}(a,b) + (c,d) &= (a + c,\ b + d),\\ (a,b).(c,d) &= (ac - bd,\ ad + bc).\end{aligned}$$

Then + and . are binary operations on $\mathbb{R} \times \mathbb{R}$. Notice that one and the same symbol "+" stands for two different binary operations, one on $\mathbb{R}$, and one on $\mathbb{R} \times \mathbb{R}$.

Exercises

1. Let $f: A \longrightarrow B$ be a mapping. Prove that f is one-to-one if and only if there is a mapping $g: B \longrightarrow A$ such that $fg = \iota_A$; prove that f is onto if and only if there is a mapping $h: B \longrightarrow A$ such that $hf = \iota_B$.

2. Let $f: A \longrightarrow B$ be a mapping. For any subset A_1 of A, we put

$$f(A_1) = \{f(a) \in B\colon\ a \in A_1\}$$

and for any subset B_1 of B, we put

$$f^{\leftarrow}(B_1) = \{a \in A : f(a) \in B_1\}.$$

($f(A_1)$ is called the *image of* A_1, and $f^{\leftarrow}(B_1)$ is called the *preimage of* B_1. Most people refer to $f(A)$ as the range of f. Here we wrote the functions on the left.) Prove that

$$\begin{aligned}f(A_1 \cap A_2) &\subseteq f(A_1) \cap f(A_2), & f(A_1 \cup A_2) &= f(A_1) \cup f(A_2)\\ f^{\leftarrow}(B_1 \cap B_2) &= f^{\leftarrow}(B_1) \cap f^{\leftarrow}(B_2), & f^{\leftarrow}(B_1 \cup B_2) &= f^{\leftarrow}(B_1) \cup f^{\leftarrow}(B_2)\end{aligned}$$

$$A_1 \subseteq f^{\leftarrow}(f(A_1)), \qquad f(f^{\leftarrow}(B_1)) \subseteq B_1$$

for any subsets A_1, A_2 of A and for any subsets B_1, B_2 of B.

3. Keep the notation of Ex. 2. Prove that f is one-to-one if and only if $f(A_1 \cap A_2) = f(A_1) \cap f(A_2)$ for any subsets A_1, A_2 of A.

4. Keep the notation of Ex. 2. Assume that f is one-to-one and onto, and let $f^{-1}: B \to A$ be its inverse. Show that

$$f^{\leftarrow}(B_1) = f^{-1}(B_1) \quad \text{and} \quad (f^{-1})^{\leftarrow}(A_1) = f(A_1)$$

for any subsets B_1 and A_1 of B and A, respectively.