

1.4 Functions, Composition, Injections and Surjections

Section 1.3 made “the pairs a description picks out” into a set: a relation from A to B is a subset of $A \times B$. Among its examples, the graph of squaring relates each real number to exactly one number, and the unit circle relates 0 to two numbers and 2 to none. This section gives the first kind its name, **function**, and builds the rest on it: we restrict functions, use them to make families, sequences and the set Y^X precise, compose them, and ask the two questions that the rest of the chapter turns on: does a function send different inputs to different outputs, and does it reach every point of its target?

1.4.1 What is a function?

Everyone has used functions since school, and the everyday definition is “a rule that assigns to each input one output”. Look at the word “rule”. The rules $x \mapsto x^2$ and $x \mapsto (x+1)^2 - 2x - 1$ are different: one squares, the other adds, squares, subtracts, and subtracts again. Yet on every real number they give the same answer, and nobody would call them different functions. Other functions have no rule at all that one could write down. A table of measured temperatures, one per day, assigns to each day one number, and no formula produces it.

So the rule is not what matters. What matters is **which output goes with which input**: the set of input–output pairs. That set is a relation, and Section 1.3 said exactly when two relations are the same, namely when they contain the same pairs. So we forget the rule and keep the pairs. One more piece of data is needed, the set in which the outputs are **declared** to lie; the reason is given after the definition.

A function is its graph, with a declared target.

Definition 1.4.1. Function

Let X and Y be sets. A **function from X to Y** , written $f: X \rightarrow Y$, is given by a relation $G \subseteq X \times Y$, the **graph** of f , such that

for every $x \in X$ there is exactly one $y \in Y$ with $(x, y) \in G$,

together with the set Y , the **codomain** of f . Formally, f is the pair (G, Y) . The set X is the **domain** of f . For $x \in X$, the unique y with $(x, y) \in G$ is written $f(x)$, the **value** of f at x . We write $(x, y) \in f$ for $(x, y) \in G$, and $x \mapsto f(x)$ for the whole function when its domain and codomain are clear.

In symbols, the condition on the graph is

$$\forall x \in X \quad \exists! y \in Y \quad (x, y) \in G,$$

where, as with the bounded quantifiers of Logic and Proof, Definition 5.1.1, $\exists! y \in Y \psi$ abbreviates $\exists! y (y \in Y \wedge \psi)$.

The whole definition is the quantifier pattern $\forall\exists!$. By Logic and Proof, Definition 5.8.1, “ $\exists! y$ ” is an existence claim and a uniqueness claim, and by Logic and Proof, Proposition 5.8.1 the uniqueness claim may be stated without reference to the witness; so the condition on the graph is two clauses:

- **(Fn1)** (existence) for every $x \in X$ there is **at least one** $y \in Y$ with $(x, y) \in G$: every input has an output;
- **(Fn2)** (uniqueness) for every $x \in X$ and all $y, y' \in Y$, if $(x, y) \in G$ and $(x, y') \in G$, then $y = y'$: no input has two outputs.

In words: the graph is a set of input–output pairs, with inputs from X and outputs from Y ; (Fn1) says no input is left without an output, and (Fn2) says no input is given two. The codomain Y is where outputs are allowed to lie. It need not consist of outputs: $x \mapsto x^2$ can be declared as a function from $\mathbb{R}$ to $\mathbb{R}$, although no output is negative. The set of outputs that actually occur is the **range** of f ,

$$\{y \in Y : \exists x \in X \ f(x) = y\},$$

which Section 1.6 will call the image of X and write $f(X)$. The domain needs no slot of its own in the pair (G, Y) : by (Fn1) every $x \in X$ occurs as a first coordinate in G , and since $G \subseteq X \times Y$, nothing else does. So the graph determines the domain.

The notation $f(x)$ is well defined. Writing “**the** y with $(x, y) \in f$ ” is legitimate only if there is one, and only one. Clause (Fn1) supplies one, and clause (Fn2) says that any two are equal; together they license the article “the”. This gives the equivalence that every argument with functions uses:

$$(x, y) \in f \iff y = f(x) \quad (x \in X, y \in Y). \quad (**)$$

For $(\Leftarrow)$: $(x, f(x)) \in f$ by the choice of $f(x)$, and if $y = f(x)$ then $(x, y) \in f$ by substituting equals ($=$ E, Logic and Proof, Definition 5.7.1). For $(\Rightarrow)$: both (x, y) and $(x, f(x))$ lie in f , so $y = f(x)$ by (Fn2). Two consequences follow at once. Equal inputs give equal outputs: if $x_1 = x_2$, then $f(x_1) = f(x_2)$. And every value lies in the codomain: $f(x) \in Y$ for every $x \in X$.

Functions given by formulas. Most functions in this book are introduced by a formula: “let $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ ”. This means the function whose graph is $\{(x, y) \in \mathbb{R}^2 : y = x^2\}$. Clause (Fn2) holds for any such graph: if $y = x^2$ and $y' = x^2$, then $y = y'$. So for a formula only (Fn1) needs checking, and it says: for each x in the domain, the formula **makes sense** and its value **lies in the codomain**. For x^2 on $\mathbb{R}$ both are clear. For “ $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 1/x$ ” the formula makes no sense at $x = 0$, and (Fn1) fails there; for “ $f: \mathbb{R} \rightarrow [0, \infty)$, $f(x) = x^2$ ” the value must be checked to lie in $[0, \infty)$, which it does, since squares are nonnegative.

A function whose domain is a product $A \times B$ is a **function of two variables**. Its inputs are pairs, and we write $f(a, b)$ rather than $f((a, b))$. Addition $(m, n) \mapsto m + n$ from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ is an example.

Example. First functions

Check (Fn1) and (Fn2), and find the range, for each of the following.

- $f: \{1, 2, 3\} \rightarrow \{a, b\}$ with graph $\{(1, a), (2, b), (3, a)\}$.
- The **identity function** $\text{id}_X: X \rightarrow X$, $\text{id}_X(x) = x$, for a set X .
- A **constant function** $X \rightarrow Y$, $x \mapsto c$, for a fixed $c \in Y$.
- The **indicator function** $\mathbf{1}_A: X \rightarrow \{0, 1\}$ of a subset $A \subseteq X$:

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

Solution.

- (a) The graph is a subset of $\{1, 2, 3\} \times \{a, b\}$. Each of 1, 2, 3 occurs as a first coordinate (Fn1), and each occurs exactly once, so no input has two outputs (Fn2). So f is a function, with $f(1) = a$, $f(2) = b$, $f(3) = a$. Its range is $\{a, b\}$, the whole codomain. A function on a small finite set is often given this way, by listing its values; no formula is needed.
- (b) The graph is the **diagonal** $\{(x, x) : x \in X\}$. For each x , the pair (x, x) is in it (Fn1); and if (x, y) is in it, then $(x, y) = (z, z)$ for some $z \in X$, and the characteristic property (Definition 1.3.1) gives $x = z = y$ (Fn2). The range is X .
- (c) The graph is $X \times \{c\}$. For each $x \in X$, (x, c) is in it (Fn1); if (x, y) is in it, then $y \in \{c\}$ by the membership test (*) for pairs in a product (Definition 1.3.2), so $y = c$ (Fn2). If X has an element, the range is $\{c\}$. If $X = \emptyset$, the range is $\emptyset$.
- (d) The graph is $(A \times \{1\}) \cup ((X \setminus A) \times \{0\})$. For (Fn1), let $x \in X$. By excluded middle (Logic and Proof, Proposition 2.5.3), $x \in A$ or $x \notin A$; in the first case $(x, 1)$ is in the graph, in the second $x \in X \setminus A$ and $(x, 0)$ is. For (Fn2), the two cases cannot both hold, so x never receives both 0 and 1, and each case gives only one pair with first coordinate x . The range is $\{0, 1\}$ if A has an element and $X \setminus A$ has an element; it is $\{0\}$ if $A = \emptyset$ and X has an element, $\{1\}$ if $A = X$ and X has an element, and $\emptyset$ if $X = \emptyset$. This is the first function of the chapter whose existence clause needs excluded middle: the definition by cases is a function only because every x falls into one of the cases.

The next example collects relations that fail to be functions. Each failure is one clause, and it pays to say which.

Example. Relations that are not functions

- (a) Is the unit circle $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ the graph of a function $\mathbb{R} \rightarrow \mathbb{R}$?
- (b) “Send a fraction p/q to its numerator p .” Is this a function from $\mathbb{Q}$ to $\mathbb{Z}$? As a relation it is

$$N := \{(r, p) \in \mathbb{Q} \times \mathbb{Z} : \exists q \in \mathbb{Z}_{>0} \ r = p/q\}.$$

Solution.

- (a) No, and both clauses fail, at different points. Clause (Fn2) fails at $x = 0$: both $(0, 1)$ and $(0, -1)$ lie in S , and $1 \neq -1$. Clause (Fn1) fails at $x = 2$: a pair $(2, y) \in S$ would need $y^2 = 1 - 4 = -3$, and no real square is negative. Restricting the inputs to $[-1, 1]$ repairs (Fn1) but not (Fn2); keeping only the upper half, $y \geq 0$, repairs (Fn2) as well, and gives the function $[-1, 1] \rightarrow \mathbb{R}$, $x \mapsto \sqrt{1 - x^2}$.
- (b) No. Clause (Fn1) holds: every rational number is p/q for some $p \in \mathbb{Z}$ and $q \in \mathbb{Z}_{>0}$. But (Fn2) fails, because a rational number has many ways of being written as a fraction. Since $\frac{1}{2} = \frac{2}{4}$, both $(\frac{1}{2}, 1)$ and $(\frac{1}{2}, 2)$ lie in N , and $1 \neq 2$. The “rule” gives an output that depends on **how the input is written**, not on the input. A function may not do that, since its value is attached to the input itself. This is why definitions of the form “write x as ..., then send it to ...” must be checked for **well-definedness**, a check that returns in Chapter 2 for functions defined on classes. Here the repair is to fix the way of writing: let q_0 be the least $q \in \mathbb{Z}_{>0}$ with $rq \in \mathbb{Z}$ (such q exist, by the first

sentence, so there is a least one by well-ordering, proved in Chapter 2; a least element is unique), and send r to rq_0 . This is a function.

Why this definition. Each clause has a job, and dropping one lets in the failures of [Example \(Relations that are not functions\)](#). Without (Fn1), a “function” could be undefined at some inputs, like $1/x$ at 0; then $f(x)$ would not name anything, and statements “for every $x \in X$ ” about $f(x)$ would lose their meaning. Without (Fn2), $f(x)$ would not name **one** thing, and $f(\frac{1}{2})$ in part (b) would be both 1 and 2. The graph replaces the rule because two rules can describe one function, and sets are compared by their elements.

Why is the codomain part of the data? A relation is a set, so if a function were only its graph, two functions with the same graph would be equal (Definition 1.1.1), whatever target was declared. The graph of $x \mapsto x^2$ as a function $[0, \infty) \rightarrow \mathbb{R}$ is the same set as its graph as a function $[0, \infty) \rightarrow [0, \infty)$. But these should be different functions, because the second reaches every point of its target and the first does not: this is the property called surjectivity below, and an example later in this section shows it changing with the codomain. Recording the codomain in the pair (G, Y) keeps the two apart. Many books take the same convention, some without saying so.

Warning. A formula is not yet a function. “The function $1/x$ ” names no function until a domain and a codomain are given, and the domain must be chosen so that (Fn1) holds: $x \mapsto 1/x$ is a function $\mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$, and there is no function $\mathbb{R} \rightarrow \mathbb{R}$ given by this formula. Likewise $\sqrt{x}$ makes sense on $[0, \infty)$ and not on $\mathbb{R}$. When you meet “let $f(x) = \dots$ ”, ask first where x lives and where $f(x)$ is meant to lie.

Quick check. Which of these relations on $\{1, 2, 3\}$ are graphs of functions $\{1, 2, 3\} \rightarrow \{1, 2, 3\}$: $\{(1, 2), (2, 2), (3, 2)\}$, $\{(1, 1), (2, 2)\}$, $\{(1, 3), (2, 1), (3, 2), (2, 3)\}$? For each that is not, name the clause that fails.

Solution. The first is a function: the constant function with value 2. The second fails (Fn1) at 3, which has no output. The third fails (Fn2) at 2, which has the two outputs 1 and 3.

1.4.2 When are two functions equal?

A function is a pair (G, Y) , so two functions are equal when their graphs are equal and their codomains are equal. In practice we compare functions by their values, and the next result says that this is the same test.

Proposition 1.4.1. Equality of Functions

Let $f: X \rightarrow Y$ and $g: X' \rightarrow Y'$ be functions. Then $f = g$ if and only if $X = X'$, $Y = Y'$, and $f(x) = g(x)$ for every $x \in X$.

Idea. Write $f = (G, Y)$ and $g = (G', Y')$. The characteristic property of pairs splits $f = g$ into “same graph” and “same codomain”. Two graphs are equal when they have the same

elements, and by (**) an element of a graph is a pair $(x, f(x))$; so “same graph” is “same inputs, same outputs”.

Proof. Write $f = (G, Y)$ and $g = (G', Y')$, as in [Definition 1.4.1](#).

($\Rightarrow$) Suppose $f = g$. By the characteristic property of pairs (Definition 1.3.1), $G = G'$ and $Y = Y'$. The domain is determined by the graph, as the set of first coordinates of its elements, so $X = X'$. Let $x \in X$. By (**), $(x, f(x)) \in G = G'$, so $(x, f(x)) \in g$, and (**) for g gives $f(x) = g(x)$.

($\Leftarrow$) Suppose $X = X'$, $Y = Y'$, and $f(x) = g(x)$ for every $x \in X$. We show $G \subseteq G'$. Let $p \in G$. Since $G \subseteq X \times Y$, $p = (x, y)$ with $x \in X$ and $y \in Y$. By (**), $y = f(x)$, so $y = g(x)$ by the hypothesis, and then $p = (x, y) \in G'$ by (**) for g . The same argument with f and g exchanged gives $G' \subseteq G$, so $G = G'$ by Proposition 1.1.1. Since also $Y = Y'$, the characteristic property gives $(G, Y) = (G', Y')$, that is, $f = g$. ■

So to show two functions equal, check three things: the same domain, the same codomain, and the same value at an arbitrary point. To show them different, one witness suffices: a point where the values differ, or a difference in domain or codomain.

Example. Equal and unequal functions

Decide which of these functions are equal.

- (a) $f, g: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = (x + 1)^2 - 2x - 1$ and $g(x) = x^2$.
- (b) $h, k: \mathbb{R} \rightarrow \mathbb{R}$, $h(x) = |x|$ and $k(x) = \sqrt{x^2}$.
- (c) $s_1: [0, \infty) \rightarrow \mathbb{R}$ and $s_2: [0, \infty) \rightarrow [0, \infty)$, both given by $x \mapsto x^2$.

Solution.

- (a) Equal. The domains are both $\mathbb{R}$ and the codomains both $\mathbb{R}$. For every real x , $(x + 1)^2 - 2x - 1 = x^2 + 2x + 1 - 2x - 1 = x^2$. By [Proposition 1.4.1](#), $f = g$: two rules, one function.
- (b) Equal. Same domain and codomain, and for every real x , $\sqrt{x^2} = |x|$: the number $|x|$ is nonnegative and its square is x^2 , and $\sqrt{x^2}$ is by definition the nonnegative number with square x^2 , which is unique by Logic and Proof, Theorem 2.4.5.
- (c) Not equal. They have the same domain, the same values, and even the same graph; but their codomains $\mathbb{R}$ and $[0, \infty)$ are different sets (-1 is in one and not the other), so by [Proposition 1.4.1](#) $s_1 \neq s_2$.

Restriction. A function can be cut down to a smaller domain. For $f: X \rightarrow Y$ and $A \subseteq X$, the **restriction** of f to A is the function

$$f|_A: A \rightarrow Y, \quad f|_A(a) := f(a),$$

whose graph is $G \cap (A \times Y)$: the pairs of f whose input lies in A . The squaring function $\mathbb{R} \rightarrow \mathbb{R}$

restricted to $[0, \infty)$ is s_1 of part (c). The bar in $f|_A$ is one of the uses of a vertical bar that the book's colon-for-set-builder convention keeps free.

1.4.3 Families, sequences, and the set of all functions

Section 1.2 used families of sets $(A_i)_{i \in I}$ and described them informally, as “a set A_i for each index $i \in I$ ”. We can now say exactly what a family is, once its members are known to lie in some set. A **family of subsets of X indexed by I** is a function $I \rightarrow \mathcal{P}(X)$, $i \mapsto A_i$, with the power set $\mathcal{P}(X)$ as the codomain the definition asks for. More generally, if S is a set whose elements are sets, a family of members of S indexed by I is a function $I \rightarrow S$. A codomain must be named, because there is no set of all sets (Theorem 1.1.4). The value of a family at i is written A_i rather than $A(i)$, and the family is written $(A_i)_{i \in I}$, in parentheses because it is indexed. A family is not the same thing as its set of values $\{A_i : i \in I\}$: the family $(A_i)_{i \in \{1,2\}}$ with $A_1 = A_2 = \{0\}$ has two indices and one value.

In the same way, a **sequence** of real numbers $(a_n)_{n \geq 1}$ is a function $\mathbb{Z}_{>0} \rightarrow \mathbb{R}$, $n \mapsto a_n$, with the value at n written a_n . The sequence $(1/n)$ is the function $n \mapsto 1/n$, and its set of values is $\{1/n : n \geq 1\}$. Sequences are the main object of *Analysis on the Real Line*.

Finally, all functions from X to Y form a set, written

$$Y^X := \{f \in \mathcal{P}(X \times Y) \times \{Y\} : f \text{ is a function from } X \text{ to } Y\}.$$

It is a set by the working rules of Section 1.1: every function $X \rightarrow Y$ is a pair (G, Y) with $G \in \mathcal{P}(X \times Y)$, so Y^X is the subset of $\mathcal{P}(X \times Y) \times \{Y\}$ cut out by the condition of Definition 1.4.1, and this product is a set by (W2), (W4), and (W5), so Y^X is a set by specification (W1). The letters are in this order because, for finite sets, Y^X has m^n elements when Y has m and X has n , as the next example suggests; counting is Chapter 3's business. The set $\{0, 1\}^{\mathbb{N}}$ of all sequences of zeros and ones, indexed from 0, plays a leading role in Chapter 3.

Example. All functions between two small sets

List the elements of $\{0, 1\}^{\{1,2\}}$, and say which are indicator functions of subsets of $\{1, 2\}$.

Solution. A function $f: \{1, 2\} \rightarrow \{0, 1\}$ is determined by its two values $f(1)$ and $f(2)$ (Proposition 1.4.1), and each may be 0 or 1. So there are four:

$f(1)$	$f(2)$	graph	indicator of
0	0	$\{(1, 0), (2, 0)\}$	$\emptyset$
0	1	$\{(1, 0), (2, 1)\}$	$\{2\}$
1	0	$\{(1, 1), (2, 0)\}$	$\{1\}$
1	1	$\{(1, 1), (2, 1)\}$	$\{1, 2\}$

Every one of them is an indicator function: $f = \mathbf{1}_A$ for $A = \{x \in \{1, 2\} : f(x) = 1\}$. The four subsets of $\{1, 2\}$ and the four functions correspond one to one.

The correspondence at the end of the example, between subsets of X and functions $X \rightarrow \{0, 1\}$, holds for every set X : send f to $\{x \in X : f(x) = 1\}$ and A to $\mathbf{1}_A$. So counting the subsets of a finite set and counting the functions from it to $\{0, 1\}$ are one problem, which Chapter 3 solves.

The empty set gives the degenerate cases, and they are worth settling now, since they are the test cases of many later statements.

Proposition 1.4.2. Functions and the Empty Set

- (a) For every set Y there is exactly one function from $\emptyset$ to Y : the pair $(\emptyset, Y)$, called the **empty function**.
- (b) If X is nonempty (it has an element), there is no function from X to $\emptyset$.

Proof.

- (a) *Existence.* The empty relation satisfies $\emptyset \subseteq \emptyset \times Y$, since $\emptyset$ is a subset of every set (Section 1.1). The condition of Definition 1.4.1 reads “for every x , if $x \in \emptyset$, then there is exactly one $y \in Y$ with $(x, y) \in \emptyset$ ”. For each x the hypothesis $x \in \emptyset$ is false, so the implication holds (Logic and Proof, Proposition 2.1.2). So $(\emptyset, Y)$ is a function from $\emptyset$ to Y .

Uniqueness. Let (G, Y) be any function from $\emptyset$ to Y . Its graph satisfies $G \subseteq \emptyset \times Y = \emptyset$, by the mirror-image $\emptyset$ law of Proposition 1.3.1. So an element of G would be an element of $\emptyset$; G has no elements, and $G = \emptyset$ by Proposition 1.1.3. Hence $(G, Y) = (\emptyset, Y)$.

- (b) Let $x_0 \in X$, and suppose $f: X \rightarrow \emptyset$ were a function. By (Fn1) there would be $y \in \emptyset$ with $(x_0, y) \in f$; but $\emptyset$ has no elements. So there is no such function $[-I]$. ■

In the notation Y^X : the set $Y^\emptyset$ has exactly one element for every Y , including $Y = \emptyset$; and $\emptyset^X = \emptyset$ whenever X is nonempty. The empty function $\emptyset \rightarrow Y$ returns in Section 1.5 as the witness that a hypothesis there cannot be dropped.

1.4.4 Composition

Two functions can be applied one after the other when the outputs of the first are inputs of the second. Temperature in degrees Celsius, as a function of the day, followed by the conversion to degrees Fahrenheit, gives temperature in Fahrenheit as a function of the day.

Definition 1.4.2. Composition

Let $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ be functions, the codomain of f being the domain of g . Their **composite** is the function

$$g \circ f: X \rightarrow Z, \quad (g \circ f)(x) := g(f(x)),$$

read “ g after f ”.

This is a function by the remark on formulas: for $x \in X$, the value $f(x)$ lies in Y , the domain of g , so $g(f(x))$ makes sense, and it lies in Z . Its graph is $\{(x, z) \in X \times Z : z = g(f(x))\}$. The order of the letters is the order of the formula $g(f(x))$, which is the **reverse** of the order in

which the functions act: f acts first.

Example. The order of composition matters

Let $f, g: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x + 1$ and $g(x) = 2x$. Compute $g \circ f$ and $f \circ g$, and decide whether they are equal.

Solution. By [Definition 1.4.2](#), $(g \circ f)(x) = g(x + 1) = 2x + 2$ and $(f \circ g)(x) = f(2x) = 2x + 1$. At $x = 0$ the values are 2 and 1, so by [Proposition 1.4.1](#) $g \circ f \neq f \circ g$. In words: “add one, then double” is not “double, then add one”. (In fact the two composites differ at every x , since $2x + 2 \neq 2x + 1$; one point already suffices.)

Composition can also fail to be defined in one order only. For $f: \{1, 2\} \rightarrow \mathbb{R}$ and $g: \mathbb{R} \rightarrow \mathbb{R}$, the composite $g \circ f: \{1, 2\} \rightarrow \mathbb{R}$ is defined, and $f \circ g$ is not: g ’s outputs are real numbers, and f accepts only 1 and 2.

Composition is not commutative, but it is associative, and the identity functions do nothing.

Proposition 1.4.3. Composition Is Associative

Let $f: X \rightarrow Y$, $g: Y \rightarrow Z$, and $h: Z \rightarrow W$ be functions. Then

$$h \circ (g \circ f) = (h \circ g) \circ f,$$

and $f \circ \text{id}_X = f = \text{id}_Y \circ f$.

Proof. Both sides of the first equation are functions from X to W , by [Definition 1.4.2](#). For $x \in X$, applying [Definition 1.4.2](#) twice on each side gives

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x).$$

So the two functions agree at every point, and they are equal by [Proposition 1.4.1](#). Likewise $f \circ \text{id}_X$ and $\text{id}_Y \circ f$ are functions from X to Y with values $f(\text{id}_X(x)) = f(x)$ and $\text{id}_Y(f(x)) = f(x)$, so both equal f . ■

Because of associativity, we write $h \circ g \circ f$ without brackets: every way of bracketing gives the same function, $x \mapsto h(g(f(x)))$.

Quick check. For $f, g: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ and $g(x) = x - 3$, compute $(g \circ f)(2)$ and $(f \circ g)(2)$.

Solution. $(g \circ f)(2) = g(4) = 1$ and $(f \circ g)(2) = f(-1) = 1$. The values agree at this point, but the functions are different: at $x = 0$, $(g \circ f)(0) = -3$ and $(f \circ g)(0) = 9$. Agreement at one point proves nothing about equality; disagreement at one point proves inequality.

1.4.5 Injective, surjective, bijective

Two questions can be asked of any function $f: X \rightarrow Y$, and much of mathematics turns on them. First: can we recover the input from the output, or can two different inputs give the same output? Second: is every point of the codomain an output? The first question is about how f treats X ; the second is about how it covers Y .

1.4.5.1 Injective functions

An injective function never sends two different inputs to the same output.

Definition 1.4.3. Injective

A function $f: X \rightarrow Y$ is **injective** (an **injection**, or **one-to-one**) if, **for all** $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$. In symbols:

$$\forall x_1 \in X \quad \forall x_2 \in X \quad (f(x_1) = f(x_2) \rightarrow x_1 = x_2).$$

In words: if two inputs give the same output, they were the same input. By Logic and Proof, Proposition 2.3.2, the definition gives the **contrapositive form**

$$\forall x_1 \in X \quad \forall x_2 \in X \quad (x_1 \neq x_2 \rightarrow f(x_1) \neq f(x_2)) :$$

different inputs give different outputs. The two forms are equivalent; the way back, from the contrapositive form to the definition, is Logic and Proof, Proposition 2.5.2, which uses RAA. A proof of injectivity usually uses the form of the definition, because an equation $f(x_1) = f(x_2)$ is something to compute with, while an inequation $x_1 \neq x_2$ is not.

The negation is found by the laws of *Logic and Proof*, Chapter 6 for negating bounded quantifiers and implications: “ f is **not** injective” is

$$\exists x_1 \in X \quad \exists x_2 \in X \quad (f(x_1) = f(x_2) \wedge x_1 \neq x_2).$$

So a failure of injectivity is shown by **two witnesses**: two different inputs with the same output.

1.4.5.2 Surjective functions

A surjective function reaches every point of its codomain.

Definition 1.4.4. Surjective

A function $f: X \rightarrow Y$ is **surjective** (a **surjection**, or **onto**) if, **for every** $y \in Y$, **there exists** $x \in X$ with $f(x) = y$. In symbols:

$$\forall y \in Y \quad \exists x \in X \quad f(x) = y.$$

In words: every point of the codomain is an output; equivalently, the range of f is all of Y . An x with $f(x) = y$ is called a **preimage** of y . The definition does not ask for the preimage to be unique. Negating as before, “ f is **not** surjective” is

$$\exists y \in Y \quad \forall x \in X \quad f(x) \neq y,$$

so a failure of surjectivity is shown by **one witness**: a point of the codomain that no input reaches.

Surjectivity depends on the codomain, which is why Definition 1.4.1 records it: s_1 and s_2 of Example (Equal and unequal functions) have the same graph, and we shall see that s_2 is surjective and s_1 is not.

Warning. Surjectivity cannot be checked by trying some y . The function $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2 + x$, reaches $0 = f(0)$, $2 = f(1)$, $6 = f(2)$, and $12 = f(3)$, and a reader who tries these may believe it reaches everything. It does not reach -1 : $x^2 + x = -1$ would mean $x^2 + x + 1 = 0$, but $x^2 + x + 1 = (x + \frac{1}{2})^2 + \frac{3}{4} > 0$ for every real x . Concluding “for every y ” from a few instances is entry (N1) of the catalogue in *Logic and Proof*, Chapter 8.

1.4.5.3 Bijective functions

Definition 1.4.5. Bijective

A function is **bijective** (a **bijection**) if it is both injective and surjective.

A bijection is a function in which every point of the codomain is reached **exactly once**: for every $y \in Y$ there is exactly one $x \in X$ with $f(x) = y$. This is the $\exists!$ of Definition 1.4.1 again, now read from the codomain side, and its two halves are the two properties. Existence of such an x , for every y , is surjectivity. Uniqueness is injectivity: if $f(x_1) = y = f(x_2)$, injectivity gives $x_1 = x_2$; and conversely, if every y has at most one preimage, then $f(x_1) = f(x_2)$ makes x_1 and x_2 two preimages of the same point $y := f(x_1)$, so $x_1 = x_2$.

1.4.5.4 Worked examples

Proofs of these properties have a fixed shape, read off the definitions.

- To prove f **injective**: “Let $x_1, x_2 \in X$ with $f(x_1) = f(x_2)$.” Compute until you reach $x_1 = x_2$.
- To prove f **surjective**: “Let $y \in Y$.” Then produce an x , **say what it is**, check that it lies in X , and check that $f(x) = y$. The x is usually found by solving $f(x) = y$ for x on scrap paper; the proof then states the solution first and checks it, as every existence proof does.
- To prove f **not injective**: name $x_1 \neq x_2$ with $f(x_1) = f(x_2)$. **Not surjective**: name $y \in Y$ and show that $f(x) = y$ has no solution $x \in X$.

Example. Deciding injective and surjective

For each function below, decide with proof whether it is injective and whether it is surjective.

- $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 3x - 5$.
- $g: \mathbb{Z} \rightarrow \mathbb{Z}$, $g(n) = 3n$.
- $h: \{1, 2, 3\} \rightarrow \{a, b\}$, with $h(1) = a$, $h(2) = b$, $h(3) = a$.
- $k: \{1, 2\} \rightarrow \{a, b, c\}$, with $k(1) = a$, $k(2) = c$.

Solution.

- Injective.* Let $x_1, x_2 \in \mathbb{R}$ with $3x_1 - 5 = 3x_2 - 5$. Adding 5 and dividing by 3 gives $x_1 = x_2$.
Surjective. On scrap paper, $3x - 5 = y$ gives $x = (y + 5)/3$. Now the proof: let $y \in \mathbb{R}$, and put $x := (y + 5)/3$. Then $x \in \mathbb{R}$, and $f(x) = 3 \cdot \frac{y+5}{3} - 5 = y$. So f is bijective.
- Injective.* Let $n_1, n_2 \in \mathbb{Z}$ with $3n_1 = 3n_2$. Dividing by 3 gives $n_1 = n_2$.
Not surjective. The witness is $1 \in \mathbb{Z}$. If $n \geq 1$, then $3n \geq 3$; if $n \leq 0$, then $3n \leq 0$. So $3n \neq 1$ for every integer n , and 1 has no preimage. (Solving $3n = 1$ gives $n = \frac{1}{3}$, which

is not in the domain $\mathbb{Z}$: the same formula on $\mathbb{Q}$ would be surjective.)

- (c) *Not injective*: $h(1) = a = h(3)$ and $1 \neq 3$. *Surjective*: the codomain has two points, and $a = h(1)$, $b = h(2)$. For a finite codomain, checking every point is a proof: the instances are all the cases.
- (d) *Injective*: the two inputs have the different outputs $a \neq c$, and there are no other pairs of different inputs to check. *Not surjective*: b is not an output, since the only outputs are $k(1) = a$ and $k(2) = c$.

Parts (c) and (d) show two more things. A function from a three-element set to a two-element set cannot be injective, and one from a two-element set to a three-element set cannot be surjective: there are too many inputs in the first case and too few in the second. Chapter 3 proves facts of this kind in general, starting from the pigeonhole principle. And the finite pictures are worth drawing: two columns of dots, with an arrow from each input to its output. Injective means no dot on the right receives two arrows; surjective means every dot on the right receives at least one.

The key example of the section is one formula read as three different functions.

Example. Squaring, three ways

For each of the three functions, settle both questions: injective or not, surjective or not.

- (a) $q_1: \mathbb{R} \rightarrow \mathbb{R}$, $q_1(x) = x^2$.
- (b) $q_2: [0, \infty) \rightarrow \mathbb{R}$, $q_2(x) = x^2$.
- (c) $q_3: [0, \infty) \rightarrow [0, \infty)$, $q_3(x) = x^2$.

Solution.

- (a) *Not injective*: $q_1(1) = 1 = q_1(-1)$ and $1 \neq -1$. *Not surjective*: the witness is $-1 \in \mathbb{R}$, since $x^2 \geq 0 > -1$ for every real x . So q_1 is neither.
- (b) *Injective*. Let $x_1, x_2 \geq 0$ with $x_1^2 = x_2^2$. Then $x_1^2 \leq x_2^2$, so $x_1 \leq x_2$ by Logic and Proof, Theorem 2.4.5, which applies since both numbers are nonnegative; and likewise $x_2 \leq x_1$. Hence $x_1 = x_2$. *Not surjective*: the witness -1 still works, for the same reason.
- (c) *Injective*, by the same proof as in (b): injectivity concerns only the domain and the values, and these are the same as for q_2 . *Surjective*. Let $y \in [0, \infty)$. A nonnegative real number has a nonnegative square root, so put $x := \sqrt{y}$. Then $x \in [0, \infty)$ and $q_3(x) = (\sqrt{y})^2 = y$. So q_3 is bijective.

The same rule gives three different functions. Shrinking the domain from $\mathbb{R}$ to $[0, \infty)$ removed the two witnesses ± 1 against injectivity. Shrinking the codomain to $[0, \infty)$ removed the witness -1 against surjectivity. The graphs of q_2 and q_3 are the same set; only the declared codomain differs, and surjectivity changes with it.

Warning. Injectivity is not the defining property of a function read backwards. Every function satisfies “if $x_1 = x_2$, then $f(x_1) = f(x_2)$ ”: equal inputs give equal outputs, which is part of $f(x)$ being well defined. Injectivity is the **converse**, “if $f(x_1) = f(x_2)$, then $x_1 = x_2$ ”. A “proof” of injectivity that starts from $x_1 = x_2$ and reaches $f(x_1) = f(x_2)$ has proved the

wrong implication, entry (N5) of *Logic and Proof*, Chapter 8’s catalogue, and it “proves” q_1 injective, which is false.

Quick check. Is $n \mapsto n + 1$ surjective as a function $\mathbb{N} \rightarrow \mathbb{N}$? As a function $\mathbb{Z} \rightarrow \mathbb{Z}$?

Solution. On $\mathbb{N}$: no. The witness is 0, since $n + 1 \geq 1$ for every $n \in \mathbb{N}$, so $n + 1 = 0$ has no solution in $\mathbb{N}$. On $\mathbb{Z}$: yes. For $m \in \mathbb{Z}$, the integer $m - 1$ satisfies $(m - 1) + 1 = m$. Both functions are injective, since $n_1 + 1 = n_2 + 1$ gives $n_1 = n_2$.

1.4.6 Composites of injections and surjections

Both properties survive composition. This is the first theorem about functions whose proof uses the definitions and nothing else, and it shows the two proof shapes of the last subsection working together.

Theorem 1.4.4. Composites of Injections and Surjections

Let $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ be functions.

- (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.

Idea. For (a), start from equal outputs of $g \circ f$. They are equal outputs of g , at the points $f(x_1)$ and $f(x_2)$; peel off g by its injectivity, then f by its. For (b), work backwards along the chain $X \rightarrow Y \rightarrow Z$: to reach $z \in Z$, first find a point of Y that g sends to z , then a point of X that f sends there. Part (c) is (a) and (b) together.

Proof.

- (a) Let $x_1, x_2 \in X$ with $(g \circ f)(x_1) = (g \circ f)(x_2)$, that is, $g(f(x_1)) = g(f(x_2))$ by [Definition 1.4.2](#). The points $f(x_1)$ and $f(x_2)$ lie in Y , and g is injective, so $f(x_1) = f(x_2)$ [$\forall E$, $\rightarrow E$: injectivity of g at $f(x_1), f(x_2)$]. Since f is injective, $x_1 = x_2$ [$\forall E$, $\rightarrow E$: injectivity of f at x_1, x_2]. So $g \circ f$ is injective.
- (b) Let $z \in Z$. Since g is surjective, there is $y \in Y$ with $g(y) = z$ [$\exists E$: the name y is fresh]. Since f is surjective, there is $x \in X$ with $f(x) = y$ [$\exists E$: x is fresh]. Then

$$(g \circ f)(x) = g(f(x)) = g(y) = z,$$

so z has the preimage x under $g \circ f$ [$\exists I$, witness x]. Since $z \in Z$ was arbitrary, $g \circ f$ is surjective.

- (c) If f and g are bijective, they are injective and surjective, so $g \circ f$ is injective by (a) and surjective by (b), that is, bijective.



Part (a) is short enough to see whole at level 0 of *Logic and Proof*, Section 8.3, a full derivation. Here is its derivation, with the quantifiers bounded and handled by the bounded rules (Logic and Proof, Proposition 6.1.1), and $(g \circ f)(u)$ already unwound to $g(f(u))$ by [Definition 1.4.2](#). Premises 1 and 2 are the injectivity of f and of g ; premise 3 says that f maps X into Y , which the prose used in the clause “the points $f(x_1)$ and $f(x_2)$ lie in Y ”.

1	$\forall u \in X \forall v \in X (f(u) = f(v) \rightarrow u = v)$	
2	$\forall u \in Y \forall v \in Y (g(u) = g(v) \rightarrow u = v)$	
3	$\forall u \in X f(u) \in Y$	
4	$x_1 \mid x_1 \in X$	
5	$x_2 \mid x_2 \in X$	
6	$g(f(x_1)) = g(f(x_2))$	
7	$f(x_1) \in Y$	$\forall E$ 3, 4
8	$f(x_2) \in Y$	$\forall E$ 3, 5
9	$\forall v \in Y (g(f(x_1)) = g(v) \rightarrow f(x_1) = v)$	$\forall E$ 2, 7
10	$g(f(x_1)) = g(f(x_2)) \rightarrow f(x_1) = f(x_2)$	$\forall E$ 9, 8
11	$f(x_1) = f(x_2)$	$\rightarrow E$ 10, 6
12	$\forall v \in X (f(x_1) = f(v) \rightarrow x_1 = v)$	$\forall E$ 1, 4
13	$f(x_1) = f(x_2) \rightarrow x_1 = x_2$	$\forall E$ 12, 5
14	$x_1 = x_2$	$\rightarrow E$ 13, 11
15	$g(f(x_1)) = g(f(x_2)) \rightarrow x_1 = x_2$	$\rightarrow I$ 6–14
16	$\forall v \in X (g(f(x_1)) = g(f(v)) \rightarrow x_1 = v)$	$\forall I$ 5–15
17	$\forall u \in X \forall v \in X (g(f(u)) = g(f(v)) \rightarrow u = v)$	$\forall I$ 4–16

Compare it with the prose of (a). “Let $x_1, x_2 \in X$ ” opens the two flagged boxes of lines 4 and 5, and “with $g(f(x_1)) = g(f(x_2))$ ” is the assumption of line 6. The first tagged sentence is lines 7–11, and the second is lines 12–14. The closing “so $g \circ f$ is injective” stands for lines 15–17, the closing lines of $\rightarrow I$ and of the two $\forall I$ ’s; by the convention of *Logic and Proof*, Chapter 8 on silent steps, the prose may state their result without naming them. The two tags in the prose mark the only steps where a hypothesis is used, and that is what the proof is about.

Each hypothesis is needed. In (a), dropping the injectivity of g breaks the conclusion: with $f = \text{id}_{\mathbb{R}}$ and $g = q_1$, squaring on $\mathbb{R}$, the composite $g \circ f = q_1$ is not injective, although f is. Dropping the injectivity of f instead, with $f = q_1$ and $g = \text{id}_{\mathbb{R}}$, gives the same composite. In (b), with $f = \text{id}_{\mathbb{R}}$ and $g = q_1$, the composite q_1 is not surjective, although f is; with $f = q_1$ and $g = \text{id}_{\mathbb{R}}$, again $g \circ f = q_1$ is not surjective, although g is.

The converses are more interesting, and they are the exercises of this section. An injective composite $g \circ f$ forces f to be injective and says nothing of the kind about g (exit check C1). A surjective composite forces g to be surjective, and f may miss points (B2). Section 1.5 builds on these facts: a function f with a function g such that $g \circ f = \text{id}_X$ is forced to be injective.

1.4.7 Exercises

1.4.7.1 A. Check your understanding

Exercise. A1

- (a) Define “ $f: X \rightarrow Y$ is injective” in symbols, and write its negation in symbols.
- (b) Is the empty function $\emptyset \rightarrow \{1\}$ injective? Is it surjective? Justify your answers.
- (c) Which of these relations on $\{1, 2, 3\}$ are graphs of functions $\{1, 2, 3\} \rightarrow \{1, 2, 3\}$? For each that is not, name the clause of the definition that fails. $R_1 = \{(1, 2), (2, 2), (3, 1)\}$, $R_2 = \{(1, 1), (2, 3)\}$, $R_3 = \{(1, 1), (1, 2), (2, 2), (3, 3)\}$.
- (d) How many functions $\{1, 2\} \rightarrow \{a, b, c\}$ are there, and how many of them are injective? Answer by listing.

Solution.

- (a) f is injective: $\forall x_1 \in X \forall x_2 \in X (f(x_1) = f(x_2) \rightarrow x_1 = x_2)$ (Definition 1.4.3). Its negation: $\exists x_1 \in X \exists x_2 \in X (f(x_1) = f(x_2) \wedge x_1 \neq x_2)$. The bounds stay as they are, and the negated implication becomes “hypothesis and not conclusion”, by the negation laws of *Logic and Proof*, Chapter 6.
- (b) Injective: yes. The condition “for all $x_1, x_2 \in \emptyset, \dots$ ” holds vacuously, since there are no $x_1 \in \emptyset$ (Logic and Proof, Proposition 2.1.2). Surjective: no. The witness is 1: there is no $x \in \emptyset$ at all, so none with value 1.
- (c) R_1 is a function ($1 \mapsto 2, 2 \mapsto 2, 3 \mapsto 1$). R_2 is not: clause (Fn1) fails at 3, which has no output. R_3 is not: clause (Fn2) fails at 1, which has the two outputs 1 and 2.
- (d) A function $f: \{1, 2\} \rightarrow \{a, b, c\}$ is determined by the pair of values $(f(1), f(2))$ (Proposition 1.4.1), and each value can be any of the three letters. Listing the pairs:

$(a, a), (a, b), (a, c),$
 $(b, a), (b, b), (b, c),$
 $(c, a), (c, b), (c, c).$

There are 9 functions. A function on $\{1, 2\}$ is injective exactly when $f(1) \neq f(2)$, since 1 and 2 are the only two different inputs. Removing the three pairs with equal entries, on the diagonal of the list, leaves 6 injective functions.

1.4.7.2 B. Practice

Exercise. B1: Five functions

Determine which of the following are injective, which are surjective, and which are not functions at all. Justify your answers.

- (a) $\mathbb{Z} \rightarrow \mathbb{Z}, n \mapsto 2n.$
- (b) $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, (m, n) \mapsto m + n.$
- (c) $\mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3.$

- (d) $\mathbb{R} \rightarrow (-1, 1)$, $x \mapsto x/(1 + |x|)$.
 (e) $\mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$, $n \mapsto n + (-1)^n$.

Solution.

- (a) *Injective:* $2n_1 = 2n_2$ gives $n_1 = n_2$. *Not surjective:* the witness is 1, which is odd, while every $2n$ is even, and no integer is both even and odd.
 (b) *Not injective:* $(1, 2) \neq (2, 1)$ (Definition 1.3.1), and both are sent to 3. *Surjective:* for $k \in \mathbb{N}$, the pair $(0, k)$ lies in $\mathbb{N} \times \mathbb{N}$ and is sent to k .
 (c) *Injective.* Let $x_1^3 = x_2^3$. Then

$$0 = x_1^3 - x_2^3 = (x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2), \quad x_1^2 + x_1x_2 + x_2^2 = (x_1 + \tfrac{1}{2}x_2)^2 + \tfrac{3}{4}x_2^2.$$

If $x_1 \neq x_2$, the second factor is positive: if $x_2 \neq 0$, then $\frac{3}{4}x_2^2 > 0$; if $x_2 = 0$, it equals $x_1^2 > 0$, because $x_1 \neq 0$. Then the product is nonzero, a contradiction. So $x_1 = x_2$. *Surjective:* for $y \in \mathbb{R}$, every real number has a real cube root, and $x := \sqrt[3]{y}$ satisfies $x^3 = y$. So the function is bijective.

- (d) *A function:* for every x , $1 + |x| > 0$, and $|x/(1 + |x|)| = |x|/(1 + |x|) < 1$, so the value lies in $(-1, 1)$. *Injective.* Suppose $y = x/(1 + |x|)$; we solve for x . Taking absolute values, $|y| = |x|/(1 + |x|)$, so $|y|(1 + |x|) = |x|$, that is, $|x|(1 - |y|) = |y|$; since $|y| < 1$ (shown above), $1 - |y| > 0$, and $|x| = |y|/(1 - |y|)$, and then $1 + |x| = 1/(1 - |y|)$. Hence

$$x = y(1 + |x|) = \frac{y}{1 - |y|}.$$

So x is determined by y : two inputs with the same output y are both equal to $y/(1 - |y|)$, hence equal. *Surjective.* Let $y \in (-1, 1)$ and put $x := y/(1 - |y|)$, a real number since $1 - |y| > 0$. Then $|x| = |y|/(1 - |y|)$, so $1 + |x| = 1/(1 - |y|)$, and

$$\frac{x}{1 + |x|} = \frac{y}{1 - |y|} \cdot (1 - |y|) = y.$$

So the function is bijective. (For example, $x = 1 \mapsto \frac{1}{2}$ and $y = \frac{1}{2}$ comes back to $x = \frac{1/2}{1/2} = 1$.)

- (e) *Not a function* $\mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$: at $n = 1$ the formula gives $1 + (-1) = 0$, and $0 \notin \mathbb{Z}_{>0}$, so (Fn1) fails at 1. This is the planted failure: the question whether it is injective or surjective does not arise. (On $\mathbb{N} = \{0, 1, 2, \dots\}$ the same formula is a function, and a bijection: it swaps $0 \leftrightarrow 1$, $2 \leftrightarrow 3$, and so on.)

Exercise. B2: Put the proof in order

Let $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ be functions. The following lines contain a proof that **if $g \circ f$ is surjective, then g is surjective**, shuffled, together with two lines that do not belong. Write the six correct lines in an order that makes a proof, and say why each of the other two lines fails.

- (i) Then $g(y) = g(f(x)) = (g \circ f)(x) = z$, by Definition 1.4.2.
 (ii) Let $y \in Y$.
 (iii) Since $g \circ f$ is surjective, there is $x \in X$ with $(g \circ f)(x) = z$.

- (iv) Since z was arbitrary, g is surjective.
- (v) Let $y := f(x)$, an element of Y .
- (vi) Since f is surjective, $y = f(x)$ for some $x \in X$.
- (vii) Let $z \in Z$.
- (viii) So z has a preimage under g , namely y .

Solution. The order is (vii), (iii), (v), (i), (viii), (iv):

Let $z \in Z$. Since $g \circ f$ is surjective, there is $x \in X$ with $(g \circ f)(x) = z$.
 Let $y := f(x)$, an element of Y . Then $g(y) = g(f(x)) = (g \circ f)(x) = z$, by [Definition 1.4.2](#). So z has a preimage under g , namely y . Since z was arbitrary, g is surjective.

No other order is admissible. Each line uses the one before it: (iii) needs the z of (vii); (v) needs the x of (iii); (i) needs the y of (v); (viii) is the $\exists I$ that (i) makes possible; and (iv) closes the $\forall I$ opened by (vii).

The distractor (ii) opens with an arbitrary element of the wrong set. The goal “ g is surjective” is $\forall z \in Z \exists y \in Y g(y) = z$, so the proof must take an arbitrary $z \in Z$; the y is the witness to be found, not something given. The distractor (vi) uses a hypothesis the statement does not have: nothing says f is surjective, and the claim holds without it. For example, with $f: \{1\} \rightarrow \{1, 2\}$, $f(1) = 1$, and $g: \{1, 2\} \rightarrow \{1\}$ constant, $g \circ f$ and g are surjective and f is not.

Exercise. B3: Examples on the natural numbers

Give an example of a function $\mathbb{N} \rightarrow \mathbb{N}$ that is surjective and not injective, and one that is injective and not surjective. Check each property.

Solution. Surjective, not injective. Let $s(0) := 0$ and $s(n) := n - 1$ for $n \geq 1$; for $n \geq 1$, $n - 1 \in \mathbb{N}$, so $s: \mathbb{N} \rightarrow \mathbb{N}$ is a function. It is surjective: for $m \in \mathbb{N}$, $m + 1 \geq 1$, so $s(m + 1) = m$. It is not injective: $s(0) = 0 = s(1)$ and $0 \neq 1$.

Injective, not surjective. Let $t(n) := n + 1$. It is injective: $n_1 + 1 = n_2 + 1$ gives $n_1 = n_2$. It is not surjective: 0 is not an output, since $n + 1 \geq 1$.

Many other answers are correct. An answer to the first part must have: every value in $\mathbb{N}$ (so it is a function into $\mathbb{N}$); at least one value taken at two different inputs, with the two inputs named; and every $m \in \mathbb{N}$ reached, with a preimage given as a formula in m , since checking some m is not enough. An answer to the second part must have: values in $\mathbb{N}$; a proof that equal outputs come from equal inputs; and one named natural number that is never reached, with the reason.

1.4.7.3 C. Going deeper

Exercise. C1: Exit check: composites

Let $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ be functions.

- (a) Prove that if $g \circ f$ is injective, then f is injective.
- (b) Give an example in which $g \circ f$ is injective and g is not injective, and check both claims.

Solution.

- (a) Suppose $g \circ f$ is injective. Let $x_1, x_2 \in X$ with $f(x_1) = f(x_2)$; we must show $x_1 = x_2$. Applying g to both sides gives $g(f(x_1)) = g(f(x_2))$, because a function gives equal outputs at equal inputs. By [Definition 1.4.2](#), this says $(g \circ f)(x_1) = (g \circ f)(x_2)$. Since $g \circ f$ is injective, $x_1 = x_2$. So f is injective.
- (b) Let $X = \{1\}$, $Y = \{1, 2\}$, $Z = \{1\}$; let $f(1) = 1$, and let $g: \{1, 2\} \rightarrow \{1\}$ be the constant function $g(1) = g(2) = 1$. Then $g \circ f: \{1\} \rightarrow \{1\}$ is injective, because its domain has only one element: if $x_1, x_2 \in \{1\}$, then $x_1 = 1 = x_2$. And g is not injective: $g(1) = g(2)$ with $1 \neq 2$.

Marking scheme. (10 points) - 2: fixes arbitrary $x_1, x_2 \in X$ with $f(x_1) = f(x_2)$ and states the goal $x_1 = x_2$. - 2: applies g to both sides, with the reason that a function gives equal outputs at equal inputs. - 3: rewrites the result as an equation of values of $g \circ f$ and invokes the injectivity of $g \circ f$. - 3: an example with explicit sets and functions, and both claims checked: $g \circ f$ injective, and two different points with the same value under g .

Fatal errors (0 for the step): assuming g is injective; starting from $x_1 = x_2$ and deriving $f(x_1) = f(x_2)$ (the converse, entry (N5) of *Logic and Proof*, Chapter 8's catalogue); an example in which $g \circ f$ is not defined because the codomain of f is not the domain of g . A proof by contrapositive (from $x_1 \neq x_2$, $g \circ f$ injective gives $g(f(x_1)) \neq g(f(x_2))$, hence $f(x_1) \neq f(x_2)$, since equal values of f would give equal values of g) earns full marks: 2 points for fixing $x_1 \neq x_2$ and the goal, 3 for the use of the injectivity of $g \circ f$, and 2 for the step from $g(f(x_1)) \neq g(f(x_2))$ to $f(x_1) \neq f(x_2)$.

Exercise. C2: When the composite is a bijection

Let $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ be functions.

- (a) Deduce from [Exercise \(C1: Exit check: composites\)](#) and [Exercise \(B2: Put the proof in order\)](#) that if $g \circ f$ is bijective, then f is injective and g is surjective.
- (b) Give an example in which $g \circ f$ is bijective and neither f nor g is bijective.

Solution.

- (a) Suppose $g \circ f$ is bijective. Then $g \circ f$ is injective, so f is injective by part (a) of [Exercise \(C1: Exit check: composites\)](#). And $g \circ f$ is surjective, so g is surjective by the proof put in order in [Exercise \(B2: Put the proof in order\)](#).
- (b) Take the functions of [Exercise \(C1: Exit check: composites\)](#) (b): $f: \{1\} \rightarrow \{1, 2\}$, $f(1) = 1$, and $g: \{1, 2\} \rightarrow \{1\}$, $g(1) = g(2) = 1$. Then $g \circ f: \{1\} \rightarrow \{1\}$ sends 1 to 1, so $g \circ f = \text{id}_{\{1\}}$ by [Proposition 1.4.1](#), which is bijective. But f is not surjective, since 2 is not an output; and g is not injective, since $g(1) = g(2)$. So neither is bijective. Part (a) is sharp: it gives exactly the half of each property that survives, and this example shows that the other halves can fail together.