Math 221 Winter 2023, Lecture 15: Enumeration

website: `https://www.cip.ifi.lmu.de/~grinberg/t/23wd`

4. An introduction to enumeration

4.5. Maps (aka functions)

4.5.8. Composition of functions (cont'd)

In Definition 4.5.11 (in Lecture 14), we defined the **composition** of two functions[1] f to g to be the function

$$\begin{aligned} &(\text{domain of } g) \to (\text{target of } f), \\ &\qquad x \mapsto f(g(x)). \end{aligned}$$

This composition is denoted by $f \circ g$.

As we saw, the compositions $f \circ g$ and $g \circ f$ are usually not the same (in fact, in many cases, one of these is defined and the other isn't). In other words, composition of functions does not satisfy commutativity. However, it has a few other nice properties:

Theorem 4.5.1 (associativity of composition)**.** Let X, Y, Z, W be four sets. Let $f : Z \to W$, $g : Y \to Z$ and $h : X \to Y$ be three functions. Then,

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

Proof. Both $(f \circ g) \circ h$ and $f \circ (g \circ h)$ are functions from X to W. Moreover, for each $x \in X$, we have

$$\begin{aligned} (f \circ (g \circ h))(x) &= f((g \circ h)(x)) && \text{(by the definition of } f \circ (g \circ h)) \\ &= f(g(h(x))) && \left(\begin{array}{c}\text{since the definition of } g \circ h \\ \text{yields } (g \circ h)(x) = g(h(x))\end{array}\right) \end{aligned}$$

and

$$\begin{aligned} ((f \circ g) \circ h)(x) &= (f \circ g)(h(x)) && \text{(by the definition of } (f \circ g) \circ h) \\ &= f(g(h(x))) && \text{(by the definition of } f \circ g), \end{aligned}$$

so that

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

Since this holds for each $x \in X$, we conclude that $f \circ (g \circ h) = (f \circ g) \circ h$ (because two functions u and v from X to W are equal if and only if the equality $u(x) = v(x)$ holds for each $x \in X$). This proves the theorem. □

[1] Recall: "Function" and "map" mean the same thing.

Intuitively, the claim of Theorem 4.5.1 is pretty obvious: It is just saying that if you can do three things (applying h, applying g and applying f) in succession, then it does not matter whether you view it as "first doing h followed by g, and then doing f" or as "first doing h, and then doing g followed by f".

Thanks to Theorem 4.5.1, we can write compositions of several functions without parentheses: i.e., instead of writing $f \circ (g \circ h)$ or $(f \circ g) \circ h$, we can just write $f \circ g \circ h$.

The following property of composition of functions is even easier. We recall that id_P means the identity map on a given set P; this is the map from P to P that sends each element $p \in P$ to itself.

Theorem 4.5.2. Let $f : X \to Y$ be a function. Then,

$$f \circ \mathrm{id}_X = \mathrm{id}_Y \circ f = f.$$

Proof. For each $x \in X$, we have

$$\begin{aligned} (f \circ \mathrm{id}_X)(x) &= f(\mathrm{id}_X(x)) \\ &= f(x) \qquad (\text{since the definition of } \mathrm{id}_X \text{ yields } \mathrm{id}_X(x) = x). \end{aligned}$$

This shows that $f \circ \mathrm{id}_X = f$ (since both $f \circ \mathrm{id}_X$ and f are functions from X to Y). A similar computation yields $\mathrm{id}_Y \circ f = f$. Thus, the theorem follows. □

Thanks to Theorem 4.5.2, we can remove identity maps from compositions: e.g., the composition $f \circ g \circ \mathrm{id}_P \circ h$ (where P is the target of h and the domain of g) can be simplified to $f \circ g \circ h$.

4.5.9. Jectivities (injectivity, surjectivity and bijectivity)

Now we introduce some important properties of functions, which have to do with how often they attain certain values. There are three of these properties, and I refer to them as the "jectivity properties", as they are called injectivity, surjectivity and bijectivity.

Definition 4.5.3. Let $f : X \to Y$ be a function. Then:

(a) We say that f is **injective** (aka **one-to-one**, aka an **injection**) if

for each $y \in Y$, there exists **at most one** $x \in X$ such that $f(x) = y$.

In other words: We say that f is **injective** if there are no two distinct elements $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$.

In other words: We say that f is **injective** if any two elements $x_1, x_2 \in X$ satisfying $f(x_1) = f(x_2)$ must also satisfy $x_1 = x_2$.

(b) We say that f is **surjective** (aka **onto**, aka a **surjection**) if

for each $y \in Y$, there exists **at least one** $x \in X$ such that $f(x) = y$.

In other words: We say that f is **surjective** if every element of Y is an output value of f.

(c) We say that f is **bijective** (aka a **one-to-one correspondence**, aka a **bijection**) if

$$\text{for each } y \in Y, \text{ there exists } \textbf{exactly one } x \in X \text{ such that } f(x) = y.$$

Thus, f is bijective if and only if f is both injective and surjective.

Here are some examples:

- The function

$$\begin{aligned} f : \mathbb{N} &\to \mathbb{N}, \\ k &\mapsto k^2 \end{aligned}$$

  is injective (because no two distinct nonnegative integers x_1, x_2 satisfy $x_1^2 = x_2^2$) but not surjective (because, e.g., the nonnegative integer $2 \in \mathbb{N}$ is not the square of any nonnegative integer). Thus, it is not bijective.

- Let $S = \{0, 1, 4, 9, 16, \ldots\}$ be the set of all perfect squares (i.e., all squares of integers). Then, the function

$$\begin{aligned} g : \mathbb{N} &\to S, \\ k &\mapsto k^2 \end{aligned}$$

  is injective (for the same reason as the f in the previous example) and also surjective (since every perfect square can be written as k^2 for some $k \in \mathbb{N}$). Thus, it is bijective.

  Take note: The functions f and g differ only in their choice of target! Other than that, they are indistinguishable (both have domain $\mathbb{N}$, and send each element of this domain to its square). But of course, this little difference matters for the surjectivity, since the surjectivity depends crucially on the target. No wonder that g is surjective while f is not.

- The function

$$\begin{aligned} h : \mathbb{N} &\to \mathbb{N}, \\ k &\mapsto k//2 \end{aligned}$$

  (recall that $k//2$ is the quotient of the division of k by 2) is not injective (for example, the two distinct elements $0, 1 \in \mathbb{N}$ satisfy $h(0) = h(1)$, because both $h(0) = 0//2$ and $h(1) = 1//2$ are 0), but is surjective (because for each $y \in \mathbb{N}$, there exists an $x \in \mathbb{N}$ such that $h(x) = y$, namely for example $x = 2y$). Hence, it is not bijective.

- Let $E = \{0, 2, 4, 6, \ldots\}$ be the set of all even nonnegative integers. The function

$$\begin{aligned} h_{\text{even}} : E &\to \mathbb{N}, \\ k &\mapsto k//2 \end{aligned}$$

(note that $k//2 = k/2$ here, since k is even) is both injective and surjective, thus bijective.

- Let $O = \{1, 3, 5, 7, \ldots\}$ be the set of all odd nonnegative integers. The function

$$\begin{aligned} h_{\text{odd}} : O &\to \mathbb{N}, \\ k &\mapsto k//2 \end{aligned}$$

is also injective and surjective, thus bijective.

The following criterion for injectivity, surjectivity and bijectivity is just a restatement of Definition 4.5.3, but it can be quite useful for checking these properties:

Remark 4.5.4. Consider a function $f : X \to Y$ given by a table of all its values (possibly an infinite table if X is infinite). Imagine that all possible inputs $x \in X$ appear in the top row, and the corresponding outputs $f(x)$ appear in the bottom row, so the table looks as follows:

x	a	b	c	d	$\cdots$
$f(x)$	$f(a)$	$f(b)$	$f(c)$	$f(d)$	$\cdots$

.

Then:

(a) The function f is injective if and only if the bottom row of this table has no two equal entries.

(b) The function f is surjective if and only if every element of Y appears in the bottom row.

(c) The function f is bijective if and only if every element of Y appears exactly once in the bottom row.

For example:

- The function

$$\begin{aligned} f : \{1, 2, 3\} &\to \{7, 8, 9\}, \\ k &\mapsto k + 6 \end{aligned}$$

is bijective, as you can see from its table of values:

k	1	2	3
$f(k)$	7	8	9

(by noticing that every element of $\{7,8,9\}$ appears exactly once in the bottom row of this table). Of course, this can also be shown logically (by arguing that f is injective and surjective because adding 6 can be undone by subtracting 6).

- The function[2]

$$\begin{aligned} f : \{4,6,7\} &\to \{0,1,2\}, \\ k &\mapsto k\%3 \end{aligned}$$

is neither injective nor surjective. Indeed, its table of values

k	4	6	7
$f(k)$	1	0	1

has the element 1 appear twice in the bottom row (so f is not injective) and does not have the element 2 in its bottom row (so f is not surjective).

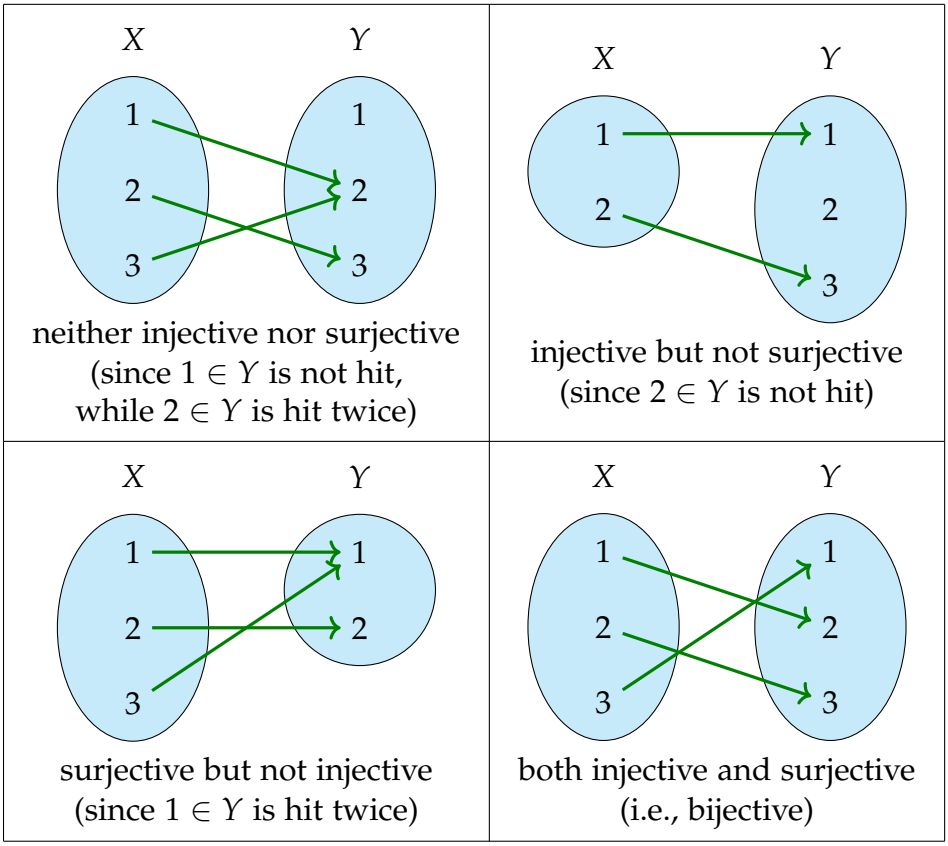
Here is yet another way to restate Definition 4.5.3:

Remark 4.5.5. If you visualize a function $f : X \to Y$ as a blobs-and-arrows picture (as explained in §4.5.3), then

- the function f is injective if and only if no two arrows hit the same Y-node;
- the function f is surjective if and only if every node in the Y-blob gets hit by at least one arrow;
- the function f is bijective if and only if every node in the Y-blob gets hit by exactly one arrow.

[2]Recall that $k\%3$ denotes the remainder of the division of k by 3.

This can be illustrated by the following four examples:



4.5.10. Inverses

Bijective maps have a special power: They can be **inverted**. Here is what this means:

Definition 4.5.6. Let $f : X \to Y$ be a function. An **inverse** of f means a function $g : Y \to X$ such that

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

In other words, an **inverse** of f means a function $g : Y \to X$ such that

$$\begin{aligned} f(g(y)) &= y \qquad \text{for each } y \in Y, \qquad \text{and} \\ g(f(x)) &= x \qquad \text{for each } x \in X. \end{aligned}$$

Roughly speaking, an inverse of f thus means a map that both undoes f and is undone by f.

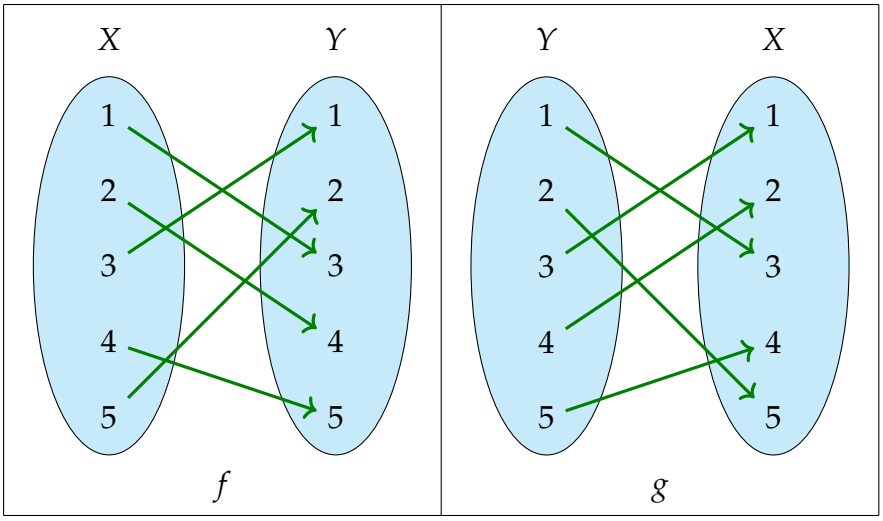
Not every function has an inverse. We shall soon see which ones do and which ones don't; we will also prove that an inverse of f is unique if it exists. For now, however, let us explore a few examples:

- Let $f : \{1,2,3\} \to \{7,8,9\}$ be the "add 6" function – i.e., the function that sends each $x \in \{1,2,3\}$ to $x+6 \in \{7,8,9\}$. Then, f has an inverse: the "subtract 6" function (i.e., the function from $\{7,8,9\}$ to $\{1,2,3\}$ that sends each y to $y-6$). Indeed, if we denote the "subtract 6" function by g, then we have

$$f(g(y)) = f(y-6) = (y-6)+6 = y \qquad \text{for each } y \in \{7,8,9\}, \qquad \text{and}$$
$$g(f(x)) = g(x+6) = (x+6)-6 = x \qquad \text{for each } x \in \{1,2,3\}.$$

- Let $f : \{1,2,3,4,5\} \to \{1,2,3,4,5\}$ be the function that sends $1,2,3,4,5$ to $3,4,1,5,2$, respectively. Then, f has an inverse: namely, the function g that sends $1,2,3,4,5$ to $3,5,1,2,4$, respectively. We can check that $f(g(y)) = y$ for each $y \in \{1,2,3,4,5\}$. For example, for $y = 3$, this is because $f(g(3)) = f(1) = 3$. Similarly we can check that $g(f(x)) = x$ for each $x \in \{1,2,3,4,5\}$.

  This is best seen by drawing the blobs-and-arrows diagrams of f and g side by side:



  As you see, there is a "dual" relationship between these two diagrams: Whenever the diagram of f has an arrow from some $x \in X$ to some $y \in Y$, the diagram of g has an arrow from y to x. In other words, the diagram of g can be obtained from the diagram of f by swapping the X-blob with the Y-blob and reversing the direction of each arrow. This rule applies not just to our specific two maps f and g, but to any map f that has an inverse. Thus, if you have drawn a blobs-and-arrows diagram of a function f, it is fairly easy to construct its inverse (as long as such an inverse exists).

  This rule can also be restated in terms of tables of values: If you have a table of all values of a function $f : X \to Y$, then you can get an inverse of f

by swapping the two rows of this table. For instance, if $f : \{1,2,3,4,5\} \to \{1,2,3,4,5\}$ is the function we just showed, then f has the table of values

k	1	2	3	4	5
$f(k)$	3	5	1	2	4

,

and thus you can get its inverse g by swapping the two rows:

k	3	5	1	2	4
$g(k)$	1	2	3	4	5

.

- Let $f : \{1,2,3,4\} \to \{1,2,3,4\}$ be the function that sends $1,2,3,4$ to $1,2,3,3$, respectively. Then, f has no inverse. Indeed, if g was an inverse of f, then we would have

$$\begin{aligned} 3 &= g(f(3)) && (\text{since } g(f(x)) = x \text{ for each } x \in \{1,2,3,4\}) \\ &= g(f(4)) && (\text{since } f(3) = 3 = f(4)) \\ &= 4 && (\text{since } g(f(x)) = x \text{ for each } x \in \{1,2,3,4\}), \end{aligned}$$

  which is absurd.

  The same argument shows that more generally, if a function $f : X \to Y$ is to have an inverse, then f should be injective, because two distinct elements x_1 and x_2 of X satisfying $f(x_1) = f(x_2)$ would create a contradiction via $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$.

- Let $f : \{1,2,3\} \to \{1,2,3,4\}$ be the function that sends $1,2,3$ to $1,2,3$, respectively. Then, f has no inverse. Indeed, if g was an inverse of f, then we would have $f(g(4)) = 4$, but this is absurd, since 4 is not an output of f.

  The same argument shows that more generally, if a function $f : X \to Y$ is to have an inverse, then f should be surjective, because each $y \in Y$ will satisfy $y = f(g(y))$ and thus be an output value of f.

Combining the morals of the last two examples, we conclude that if a function $f : X \to Y$ is to have an inverse, then f should be both injective and surjective, i.e., should be bijective. In other words, only bijective maps have a chance at having inverses. This turns out to be sufficient as well: If a map is bijective, then it has an inverse. Let us summarize this as a theorem:

Theorem 4.5.7. Let $f : X \to Y$ be a map between two sets X and Y. Then, f has an inverse if and only if f is bijective.

Proof. We must prove the logical equivalence

$$(f \text{ has an inverse}) \iff (f \text{ is bijective}). \tag{1}$$

Let us prove the $\Longrightarrow$ and $\Longleftarrow$ directions separately:

$\Longrightarrow$: Assume that f has an inverse. We must show that f is bijective.[3]

We assumed that f has an inverse. Let g be this inverse.

Let us show that f is injective. Let $x_1, x_2 \in X$ satisfy $f(x_1) = f(x_2)$. We must prove that $x_1 = x_2$. Set $y = f(x_1)$; then, $y = f(x_2)$ as well (since $f(x_1) = f(x_2)$). Since g is an inverse of f, we have $x_1 = g(f(x_1)) = g(y)$ (since $f(x_1) = y$) and $x_2 = g(f(x_2)) = g(y)$ (since $f(x_2) = y$). Thus, $x_1 = g(y) = x_2$. This completes our proof that f is injective.

Let us show that f is surjective. Let $y \in Y$. Then, $y = f(g(y))$ (since g is an inverse of f). Therefore, there exists an $x \in X$ such that $y = f(x)$ (namely, $x = g(y)$). So we have proved for each $y \in Y$ that there exists an $x \in X$ such that $y = f(x)$. In other words, f is surjective.

So f is both injective and surjective, thus bijective. This proves the "$\Longrightarrow$" direction of our equivalence (1).

Let us now prove the "$\Longleftarrow$" direction:

$\Longleftarrow$: Assume that f is bijective. We must show that f has an inverse.

Since f is bijective, for each $y \in Y$, there exists a **unique** $x \in X$ such that $f(x) = y$. Thus, we can define a map

$$g : Y \to X,$$

which sends each $y \in Y$ to this unique x. It is easy to see that g is an inverse of f. Thus, f has an inverse. This proves the "$\Longleftarrow$" direction of our equivalence (1). Thus, the proof of (1) is complete, i.e., Theorem 4.5.7 is proved. □

Theorem 4.5.7 says that bijective maps are the same as invertible maps (i.e., maps that have an inverse). This is a fundamental result that is used all over mathematics.

As we promised, let us now show that an inverse of a map f is unique if it exists:

Theorem 4.5.8. Let $f : X \to Y$ be a function. Then, f has at most one inverse.

Proof. What does "at most one inverse" mean? It means that f has no two distinct inverses. In other words, it means that any two inverses of f are identical.

So let us prove this. Let g_1 and g_2 be two inverses of f. We must show that $g_1 = g_2$.

Since g_1 is an inverse of f, we have $g_1 \circ f = \mathrm{id}_X$ and $f \circ g_1 = \mathrm{id}_Y$.

Since g_2 is an inverse of f, we have $g_2 \circ f = \mathrm{id}_X$ and $f \circ g_2 = \mathrm{id}_Y$.

By associativity of composition (Theorem 4.5.1), the two maps $(g_1 \circ f) \circ g_2$ and $g_1 \circ (f \circ g_2)$ are equal. Thus, we can denote both of these maps by $g_1 \circ f \circ g_2$.

[3] We have already done this in the above examples, but we repeat it for the sake of completeness.

Comparing

$$g_1 \circ \underbrace{f \circ g_2}_{=\mathrm{id}_Y} = g_1 \circ \mathrm{id}_Y = g_1 \qquad \text{with}$$
$$\underbrace{g_1 \circ f}_{=\mathrm{id}_X} \circ g_2 = \mathrm{id}_X \circ g_2 = g_2,$$

we find $g_1 = g_2$, qed. □

Definition 4.5.9. Let $f : X \to Y$ be a map that has an inverse. Then, this inverse (which is unique by Theorem 4.5.8) is called f^{-1}.

Thus, if $f : X \to Y$ is a map that has an inverse (i.e., by Theorem 4.5.7, a bijective map), then we have

$$f^{-1} \circ f = \mathrm{id}_X \qquad \text{and} \qquad f \circ f^{-1} = \mathrm{id}_Y,$$

that is,

$$f^{-1}(f(x)) = x \qquad \text{for each } x \in X, \qquad \text{and} \tag{2}$$
$$f\left(f^{-1}(y)\right) = y \qquad \text{for each } y \in Y. \tag{3}$$

These equalities should explain why the notation f^{-1} was chosen for the inverse of f.

Here are some further examples of inverses:

- Let $E = \{0, 2, 4, 6, \ldots\}$ be the set of all even nonnegative integers. Consider the function

$$f : E \to \mathbb{N},$$
$$k \mapsto k/2.$$

  Then, f has an inverse. This inverse is the function

$$f^{-1} : \mathbb{N} \to E,$$
$$k \mapsto 2k.$$

- Let $\mathbb{R}_{\geq 0} = \{\text{all nonnegative real numbers}\}$. Then, the function

$$f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0},$$
$$x \mapsto x^2$$

  has an inverse. This inverse is the function

$$f^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0},$$
$$x \mapsto \sqrt{x}.$$

- In contrast, the function

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^2 \end{aligned}$$

has no inverse. In fact, this function is not injective (since $f(1) = f(-1)$) and not surjective (since -1 is not a square of a real number), so it is certainly not bijective, and thus not invertible.

- The function

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^3 \end{aligned}$$

has an inverse. This inverse is the function

$$\begin{aligned} f^{-1} : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto \sqrt[3]{x}. \end{aligned}$$

Here are some more general properties of inverses:

Proposition 4.5.10. Let X be any set. Then, the identity map $\mathrm{id}_X : X \to X$ is bijective, and is its own inverse.

Proof. The map id_X is an inverse of itself (since $\mathrm{id}_X \circ \mathrm{id}_X = \mathrm{id}_X$ and $\mathrm{id}_X \circ \mathrm{id}_X = \mathrm{id}_X$). Hence, it has an inverse, and thus is bijective (by Theorem 4.5.7). □

Theorem 4.5.11. Let $f : X \to Y$ be a map that has an inverse $f^{-1} : Y \to X$. Then, f^{-1} has an inverse, namely f.

Proof. Since f^{-1} is an inverse of f, we have $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$. But the same two equalities can be read as saying that f is an inverse of f^{-1}. □

Theorem 4.5.12 (socks-and-shoes formula)**.** Let X, Y and Z be three sets. Let $g : X \to Y$ and $f : Y \to Z$ be two bijective functions. Then, the composition $f \circ g : X \to Z$ is bijective as well, and its inverse is

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

Proof. This is obvious from the blobs-and-arrows picture; but let us check this rigorously.

For any $x \in X$, we have

$$\left(g^{-1} \circ f^{-1}\right)((f \circ g)(x)) = g^{-1}\left(\underbrace{f^{-1}(f(g(x)))}_{=g(x)}\right) = g^{-1}(g(x)) = x.$$

For any $z \in Z$, we have

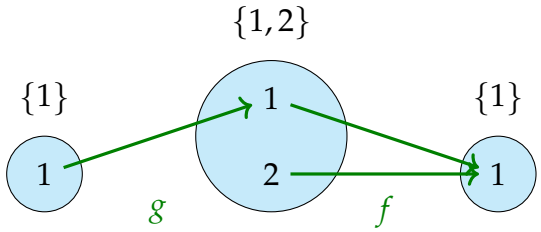
$$(f \circ g)\left(\left(g^{-1} \circ f^{-1}\right)(z)\right) = f\left(\underbrace{g\left(g^{-1}\left(f^{-1}(z)\right)\right)}_{=f^{-1}(z)}\right) = f\left(f^{-1}(z)\right) = z.$$

Thus, $g^{-1} \circ f^{-1}$ is an inverse of $f \circ g$. Hence, $f \circ g$ has an inverse, and thus is bijective (by Theorem 4.5.7). □

Remark 4.5.13. Note that $g^{-1} \circ f^{-1}$ is not the same as $f^{-1} \circ g^{-1}$. Indeed, $f^{-1} \circ g^{-1}$ might not even exist in Theorem 4.5.12.

A surprising feature of the socks-and-shoes formula $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ is that the order in which the inverses f^{-1} and g^{-1} appear on the right hand side is different from the order in which f and g appear on the left hand side. However, this is completely natural: If you want to undo two things you have done in some order, then you should undo them in the opposite order! For example, if you have put on your socks and then your shoes in the morning, then you need to first take off the shoes and then the socks when you go to bed. (The formula owes its moniker to this metaphor.)

Remark 4.5.14. Part of Theorem 4.5.12 says that a composition of two bijective functions is bijective. However, a composition $f \circ g$ of two non-bijective functions f and g can sometimes also be bijective. Here is an example:



4.5.11. Some solved exercises

Here are a few solved exercises on jectivities.

Exercise 1. Let X, Y and Z be three sets, and $f : Y \to Z$ and $g : X \to Y$ be two maps. Which of the following are true?

(a) If f and g are injective, then $f \circ g$ is injective.

(b) If $f \circ g$ is injective, then f is injective.

(c) If $f \circ g$ is injective, then g is injective.

(d) If f and g are surjective, then $f \circ g$ is surjective.

(e) If $f \circ g$ is surjective, then f is surjective.

(f) If $f \circ g$ is surjective, then g is surjective.

(g) If f and g are bijective, then $f \circ g$ is bijective.

(h) If $f \circ g$ is bijective, then f is bijective.

(i) If $f \circ g$ is bijective, then g is bijective.

Solution. We shall use the following definitions of "injective", "surjective" and "bijective"[4]:

- A map $h : U \to V$ is **injective** if and only if it has the following property: For any $u_1, u_2 \in U$ satisfying $h(u_1) = h(u_2)$, we have $u_1 = u_2$.
- A map $h : U \to V$ is **surjective** if and only if it has the following property: For any $v \in V$, there exists some $u \in U$ such that $h(u) = v$.
- A map $h : U \to V$ is **bijective** if and only if h is both injective and surjective.

(a) This is **true**.

[*Proof:* Assume that f and g are injective. We must prove that $f \circ g$ is injective. Let $u_1, u_2 \in X$ satisfy $(f \circ g)(u_1) = (f \circ g)(u_2)$. We shall show that $u_1 = u_2$. Indeed, we have $(f \circ g)(u_1) = f(g(u_1))$ (by the definition of $f \circ g$), so that

$$f(g(u_1)) = (f \circ g)(u_1) = (f \circ g)(u_2) = f(g(u_2))$$

(again by the definition of $f \circ g$). Since f is injective, we thus conclude that $g(u_1) = g(u_2)$ [5]. Since g is injective, we thus conclude that $u_1 = u_2$.

Forget that we fixed u_1, u_2. We thus have shown that for any $u_1, u_2 \in X$ satisfying $(f \circ g)(u_1) = (f \circ g)(u_2)$, we have $u_1 = u_2$. In other words, the map $f \circ g$ is injective (by our definition of "injective"). This completes our proof.]

(b) This is **false**.

[*Counterexample:* For instance, we can set $X = \{1\}$ and $Y = \{1, 2\}$ and $Z = \{1\}$, and let $f : Y \to Z$ be the map that sends both elements of Y to 1, while $g : X \to Y$ is the map sending 1 to 1. Then, $f \circ g$ is injective (in fact, $f \circ g$ is the identity map $\mathrm{id}_{\{1\}}$), but f is not.]

(c) This is **true**.

[4] We gave several equivalent definitions for "injective", "surjective" and "bijective" in Definition 4.5.3; you can just as well use any of them instead.

[5] In some more detail:

We know that f is injective. In other words, for any $v_1, v_2 \in Y$ satisfying $f(v_1) = f(v_2)$, we have $v_1 = v_2$ (by our definition of "injective"). Applying this to $v_1 = g(u_1)$ and $v_2 = g(u_2)$, we obtain $g(u_1) = g(u_2)$ (since $f(g(u_1)) = f(g(u_2))$).

[*Proof:* Assume that $f \circ g$ is injective. We must prove that g is injective.

Let $u_1, u_2 \in X$ satisfy $g(u_1) = g(u_2)$. We shall show that $u_1 = u_2$.

Indeed, we have $(f \circ g)(u_1) = f(g(u_1))$ (by the definition of $f \circ g$) and $(f \circ g)(u_2) = f(g(u_2))$ (likewise). Hence,

$$(f \circ g)(u_1) = f\left(\underbrace{g(u_1)}_{=g(u_2)}\right) = f(g(u_2)) = (f \circ g)(u_2).$$

Since $f \circ g$ is injective, this entails $u_1 = u_2$.

Forget that we fixed u_1, u_2. We thus have shown that for any $u_1, u_2 \in X$ satisfying $g(u_1) = g(u_2)$, we have $u_1 = u_2$. In other words, the map g is injective (by our definition of "injective"). This completes our proof.]

(d) This is **true**.

[*Proof:* Assume that f and g are surjective. We must prove that $f \circ g$ is surjective.

Let $z \in Z$ be arbitrary. We shall show that there exists some $x \in X$ such that $(f \circ g)(x) = z$.

Indeed, recall that f is surjective. Thus, there exists some $y \in Y$ such that $f(y) = z$. Consider this y.

Recall now that g is surjective. Thus, there exists some $w \in X$ such that $g(w) = y$. Consider this w.

We have $(f \circ g)(w) = f\left(\underbrace{g(w)}_{=y}\right) = f(y) = z$. Hence, there exists some $x \in X$ such that $(f \circ g)(x) = z$ (namely, $x = w$).

Forget that we fixed z. We thus have shown that for any $z \in Z$, there exists some $x \in X$ such that $(f \circ g)(x) = z$. In other words, the map $f \circ g$ is surjective (by our definition of "surjective"). This completes our proof.]

(e) This is **true**.

[*Proof:* Assume that $f \circ g$ is surjective. We must prove that f is surjective.

Let $z \in Z$ be arbitrary. We shall show that there exists some $y \in Y$ such that $f(y) = z$.

Indeed, recall that $f \circ g$ is surjective. Thus, there exists some $x \in X$ such that $(f \circ g)(x) = z$. Consider this x.

Now, $f(g(x)) = (f \circ g)(x) = z$. Hence, there exists some $y \in Y$ such that $f(y) = z$ (namely, $y = g(x)$).

Forget that we fixed z. We thus have shown that for any $z \in Z$, there exists some $y \in Y$ such that $f(y) = z$. In other words, the map f is surjective (by our definition of "surjective"). This completes our proof.]

(f) This is **false**.

[*Counterexample:* For instance, we can set $X = \{1\}$ and $Y = \{1, 2\}$ and $Z = \{1\}$, and let $f : Y \to Z$ be the map that sends both elements of Y to 1, while

$g : X \to Y$ is the map sending 1 to 1. Then, $f \circ g$ is surjective (in fact, $f \circ g$ is the identity map $\mathrm{id}_{\{1\}}$), but g is not.]

(g) This is **true**.

[*Proof:* Assume that f and g are bijective. Thus, f and g are both injective and surjective. Hence, $f \circ g$ is injective (by Exercise 1 **(a)**) and surjective (by Exercise 1 **(d)**). Thus, $f \circ g$ is bijective.]

(h) This is **false**.

[*Counterexample:* For instance, we can set $X = \{1\}$ and $Y = \{1,2\}$ and $Z = \{1\}$, and let $f : Y \to Z$ be the map that sends both elements of Y to 1, while $g : X \to Y$ is the map sending 1 to 1. Then, $f \circ g$ is bijective (in fact, $f \circ g$ is the identity map $\mathrm{id}_{\{1\}}$), but f is not.]

(i) This is **false**.

[*Counterexample:* For instance, we can set $X = \{1\}$ and $Y = \{1,2\}$ and $Z = \{1\}$, and let $f : Y \to Z$ be the map that sends both elements of Y to 1, while $g : X \to Y$ is the map sending 1 to 1. Then, $f \circ g$ is bijective (in fact, $f \circ g$ is the identity map $\mathrm{id}_{\{1\}}$), but g is not.] □

Exercise 2. Let $f : X \to Y$ be a map that has an inverse $f^{-1} : Y \to X$. Let $x \in X$ and $y \in Y$. Prove that we have the logical equivalence

$$\left(f(x) = y\right) \iff \left(f^{-1}(y) = x\right).$$

Solution. We shall prove the "$\Longrightarrow$" and "$\Longleftarrow$" parts of this equivalence separately:

$\Longrightarrow$: If we have $f(x) = y$, then

$$f^{-1}\left(\underbrace{y}_{=f(x)}\right) = f^{-1}(f(x)) = x \qquad \text{(by (2))}.$$

Thus, the "$\Longrightarrow$" part of the equivalence holds.

$\Longleftarrow$: If we have $f^{-1}(y) = x$, then

$$f\left(\underbrace{x}_{=f^{-1}(y)}\right) = f\left(f^{-1}(y)\right) = y \qquad \text{(by (3))}.$$

Thus, the "$\Longleftarrow$" part of the equivalence holds. □

4.5.12. Isomorphic sets

As an application of inverses, we can define the concept of isomorphic sets:

Definition 4.5.15. Let X and Y be two sets. We say that these two sets X and Y are **isomorphic as sets** (or, for short, **isomorphic**, or **in bijection**, or **in one-to-one correspondence**, or **equinumerous**) if there exists a bijective map from X to Y.

Note that this relation "isomorphic as sets" is symmetric (i.e., if X and Y are isomorphic, then Y and X are isomorphic). This is because if $f : X \to Y$ is a bijective map, then f has an inverse f^{-1} (by Theorem 4.5.7), and this inverse f^{-1} is again bijective (since Theorem 4.5.11 shows that f^{-1} again has an inverse).

Some examples:

- The sets $\{1,2\}$ and $\{1,2,3\}$ are **not** isomorphic. In fact, there is no surjective map $f : \{1,2\} \to \{1,2,3\}$ (since, informally, a map from $\{1,2\}$ to $\{1,2,3\}$ has only two arrows, but two arrows cannot hit all three elements of $\{1,2,3\}$). Thus, there is no bijective map $f : \{1,2\} \to \{1,2,3\}$ either.
- The sets $\{1,2,3\}$ and $\{6,7,8\}$ are isomorphic. In fact, the map

$$\begin{aligned} \{1,2,3\} &\to \{6,7,8\}, \\ k &\mapsto k+5 \end{aligned}$$

  (that is, the "add 5" map) is bijective (and its inverse sends $k \mapsto k-5$).
- The sets $\{1,2,3\}$ and $\{1,3,5\}$ are isomorphic. In fact, the map

$$\begin{aligned} \{1,2,3\} &\to \{1,3,5\}, \\ k &\mapsto 2k-1 \end{aligned}$$

  is a bijection.
- The sets $\mathbb{N}$ and $E := \{\text{all even nonnegative integers}\}$ are isomorphic, since the map

$$\begin{aligned} \mathbb{N} &\to E, \\ n &\mapsto 2n \end{aligned}$$

  is a bijection.
- The sets $\mathbb{N}$ and $O := \{\text{all odd nonnegative integers}\}$ are isomorphic, since the map

$$\begin{aligned} \mathbb{N} &\to O, \\ n &\mapsto 2n+1 \end{aligned}$$

  is a bijection.

- The sets $\mathbb{N}$ and $\mathbb{Z}$ are isomorphic, since there is a bijection from $\mathbb{N}$ to $\mathbb{Z}$ that sends

$$\begin{aligned} &0,1,2,3,4,5,6,7,8,\ldots && \text{to} \\ &0,1,-1,2,-2,3,-3,4,-4,\ldots, && \text{respectively.} \end{aligned}$$

  Explicitly, this map f can be defined by the following formula:

$$f(n) = \begin{cases} -n/2, & \text{if } n \text{ is even;} \\ (n+1)/2, & \text{if } n \text{ is odd} \end{cases} \qquad \text{for each } n \in \mathbb{N}.$$

  (This formula ensures that the values $f(0), f(2), f(4), f(6), \ldots$ cover exactly the integers $0, -1, -2, -3, \ldots$ that are ≤ 0, whereas the values $f(1), f(3), f(5), f(7), \ldots$ cover exactly the positive integers $1, 2, 3, 4, \ldots$.)

  There are, of course, many other bijections from $\mathbb{N}$ to $\mathbb{Z}$.

- The sets $\mathbb{N}$ and $\mathbb{Q}$ are isomorphic, since there is a bijection from $\mathbb{N}$ to $\mathbb{Q}$ that sends

$$0,1,2,3,4,5,6,7,8,9,\ldots \quad \text{to}$$
$$\underbrace{\frac{-1}{1}, \frac{0}{1}, \frac{1}{1}}_{\substack{\text{all reduced fractions whose numerator and denominator are } \leq 1 \text{ in absolute value (ordered from smallest to largest)}}} , \underbrace{\frac{-2}{1}, \frac{-1}{2}, \frac{1}{2}, \frac{2}{1}}_{\substack{\text{all reduced fractions whose numerator and denominator are } \leq 2 \text{ but not } \leq 1 \text{ in absolute value (ordered from smallest to largest)}}} , \underbrace{\frac{-3}{1}, \frac{-3}{2}, \frac{-2}{3}, \frac{-1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{3}{2}, \frac{3}{1}}_{\substack{\text{all reduced fractions whose numerator and denominator are } \leq 3 \text{ but not } \leq 2 \text{ in absolute value (ordered from smallest to largest)}}}, \ldots,$$

  respectively. (To be precise, we must only allow **fully reduced** fractions – i.e., fractions $\frac{a}{b}$ with $a \in \mathbb{Z}$ and $b \in \{1,2,3,\ldots\}$ and $\gcd(a,b) = 1$ – in order to avoid having the same rational number appear twice.)

- The sets $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ are isomorphic, since there is a bijection f from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$ that sends

$$0,1,2,3,4,5,6,7,8,9,\ldots \quad \text{to}$$
$$\underbrace{(0,0)}_{\substack{\text{all pairs whose entries sum to } 0}} , \underbrace{(0,1), (1,0)}_{\substack{\text{all pairs whose entries sum to } 1 \text{ (ordered by increasing first entry)}}}, \underbrace{(0,2), (1,1), (2,0)}_{\substack{\text{all pairs whose entries sum to } 2 \text{ (ordered by increasing first entry)}}}, \underbrace{(0,3), (1,2), (2,1), (3,0)}_{\substack{\text{all pairs whose entries sum to } 3 \text{ (ordered by increasing first entry)}}}, \ldots,$$

  respectively. The inverse $f^{-1} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ of this bijection can actually be described by an explicit formula:

$$f^{-1}(n,m) = \frac{(n+m)(n+m+1)}{2} + n$$

  (nice and not-too-easy exercise: prove this!). This is the so-called Cantor pairing function.

4.6. Counting, formally

4.6.1. Definition

As you might have noticed, isomorphic sets (at least when they are finite) have the same number of elements – i.e., the same size. We shall now use this to **define** the size of a set!

First, some notations:

Definition 4.6.1. (a) If $n \in \mathbb{N}$, then $[n]$ shall mean the set $\{1, 2, \ldots, n\}$.

For example, $[3] = \{1, 2, 3\}$ and $[7] = \{1, 2, 3, 4, 5, 6, 7\}$ and $[0] = \varnothing$ and $[1] = \{1\}$.

(b) If $a, b \in \mathbb{Z}$, then $[a, b]$ shall mean the set

$$\begin{aligned}\{a, a+1, a+2, \ldots, b\} &= \{\text{all integers } x \text{ satisfying } a \leq x \leq b\} \\ &= \{x \in \mathbb{Z} \mid a \leq x \leq b\}.\end{aligned}$$

If $a > b$, then this is understood to be the empty set.

For example, $[2, 6] = \{2, 3, 4, 5, 6\}$ and $[3, 3] = \{3\}$ and $[4, 2] = \varnothing$.

Now, let us define the size of a finite set:

Definition 4.6.2. Let $n \in \mathbb{N}$. A set S is said to have **size** n if S is isomorphic to $[n]$ (that is, if there exists a bijection from S to $[n]$).

For example:

- The set {"cat", "dog", "rat"} has size 3, since the map

$$\begin{aligned}\{\text{“cat”, “dog”, “rat”}\} &\to [3], \\ \text{“cat”} &\mapsto 1, \\ \text{“dog”} &\mapsto 2, \\ \text{“rat”} &\mapsto 3\end{aligned}$$

  is a bijection.

- The set $\{4, 5, 6, 7\}$ has size 4, since the map

$$\begin{aligned}\{4, 5, 6, 7\} &\to [4], \\ k &\mapsto k - 3\end{aligned}$$

  is a bijection.

- The set $\mathbb{N}$ is infinite, so there is no bijection from $\mathbb{N}$ to $[n]$ for any $n \in \mathbb{N}$. Thus, $\mathbb{N}$ does not have size n for any $n \in \mathbb{N}$.

Here is another equivalent definition of size:

Definition 4.6.3. We define the notion of a "set of size n" recursively as follows:

(a) A set S is said to have **size** 0 if and only if it is empty.

(b) Let n be a positive integer. A set S is said to have **size** n if and only if there exists an $s \in S$ such that $S \setminus \{s\}$ has size $n-1$.

In other words, a set has size n (for $n > 0$) if and only if we can remove a single element from it and obtain a set of size $n-1$. This is a recursive definition, as it reduces the question "what is a set of size n" to the (simpler) question "what is a set of size $n-1$".

The following fact is not obvious, but can be proved:

Theorem 4.6.4. (a) The above two definitions of size (Definition 4.6.2 and Definition 4.6.3) are equivalent.

(b) The size of a finite set is determined uniquely – i.e., a set cannot have two different sizes at the same time.

Now, we are ready to introduce some notations for sizes of sets:

Definition 4.6.5. (a) An **n-element set** (for some $n \in \mathbb{N}$) means a set of size n.

(b) A set is said to be **finite** if it has size n for some $n \in \mathbb{N}$.

(c) If S is a finite set, then $|S|$ shall denote the size of S (which is unique because of Theorem 4.6.4 **(b)**).

(d) We also refer to $|S|$ as the **cardinality** of S, or as the **number** of elements of S.

Thus, our examples above show that

$$|\{\text{“cat”, “dog”, “rat”}\}| = 3 \qquad \text{and} \qquad |\{4,5,6,7\}| = 4.$$