More Joy of Sets

In this handout we continue our summary of basic set theory begun in *The Joy of Sets*, with a special emphasis on FUNCTIONS.

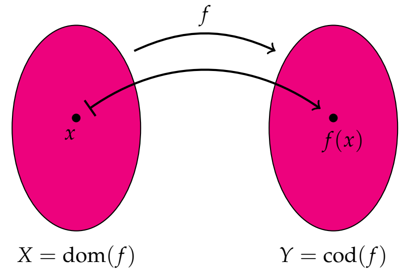


Figure 1: The function $f : X \to Y$.

Functions

If X and Y are sets, a *function from* X *to* Y is a rule[1] that assigns to each element x in the set X a unique element y in the set Y. A good name for a function is f. (You can probably guess why). If f is a function from X to Y and $x \in X$, the unique element $y \in Y$ that f associates to x is called the *value* of f at x, usually written[2] $f(x)$. To indicate that f is a function from X to Y, we write $f : X \to Y$. In math, the words *map* or *mapping* are synonymous[3] with *function*.

If $f : X \to Y$ is a function from X to Y, the set X is called the *domain* or[3] *source* of f, and the set Y is called the *codomain* or[3] *target space* of f. Sometimes it is useful to have notation for this, so we might write $\mathrm{dom}(f)$ for the domain of the function f and $\mathrm{cod}(f)$ for its codomain.

It often helps to picture functions using "blobs and arrows" as in Figure 1. If you picture $\mathrm{dom}(f)$ as one blob (on the left) and $\mathrm{cod}(f)$ as another blob (on the right), then you can represent f using arrows that transform inputs in $\mathrm{dom}(f)$ into outputs in $\mathrm{cod}(f)$.

Functions are often defined using rules that specify how to convert an input x into an output $y = f(x)$. When variables are used in this manner to define a function via a rule, the input variable (often, but not always, x) is called the *independent variable*, and the output variable (often, but not always, y) is called the *dependent variable*.

For any function $f : X \to Y$, the *image*[4] of f, written $\mathrm{im}(f)$, is the set

$$\mathrm{im}(f) \;:=\; \{f(x) \,:\, x \in X\}$$

of all values that f takes on (see Figure 2). More generally, if $f : X \to Y$ is any function, then for subsets $A \subseteq X$ and $B \subseteq Y$ we define the *direct image* or[3] *forward image* of A under f to be the set

$$f[A] \;:=\; \{f(a) \in Y \,:\, a \in A\} \;\subseteq\; \mathrm{cod}(f),$$

and we define the *inverse image* or[3] *preimage* of B under f to be the set

$$f^{-1}[B] \;:=\; \{x \in X \,:\, f(x) \in B\} \;\subseteq\; \mathrm{dom}(f).$$

These operations have friendly properties that are fun to prove.[5]

Example. For any set X, the *identity function* $\mathrm{id}_X : X \to X$ is defined by the rule $\mathrm{id}_X(x) = x$ for all $x \in X$. Identity functions may seem kind of boring, but you will encounter them frequently and find them to be quite useful.

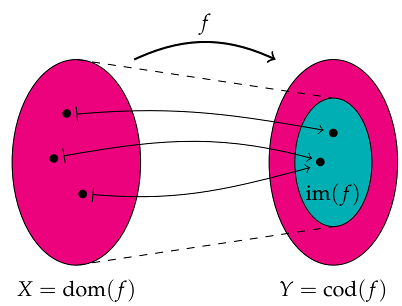


Figure 2: $\mathrm{im}(f) = f[X] \subseteq \mathrm{cod}(f)$.

[1] If you are worried about what exactly a "rule" is or suspect that this definition is not entirely rigorous, have patience! We will remedy this below.

[2] Thanks, Euler! (For those who read left-to-right, it would have been better[6] to write $(x)f$ instead of $f(x)$. Oh well.)

"$f : X \to Y$" is read "f maps X to Y." Note that the arrow "$\to$" goes between the domain X and codomain Y; for individual elements in X and Y, we use the arrow "$\mapsto$" and write "$x \mapsto f(x)$."

[3] Variety is the spice of life.

[4] Some folks use *range* to mean image, but others use it to mean codomain, so we avoid the term altogether.

[5] If $f : X \to Y$ is a function, then for all $A, B \subseteq X$ and $C, D \subseteq Y$ we have:
(i) $f[f^{-1}[C]] \subseteq C$
(ii) $f^{-1}[f[A]] \supseteq A$
(iii) $f[A \cup B] = f[A] \cup f[B]$
(iv) $f[A \cap B] \subseteq f[A] \cap f[B]$
(v) $f[A \setminus B] \supseteq f[A] \setminus f[B]$
(vi) $f^{-1}[C \cup D] = f^{-1}[C] \cup f^{-1}[D]$
(vii) $f^{-1}[C \cap D] = f^{-1}[C] \cap f^{-1}[D]$
(viii) $f^{-1}[C \setminus D] = f^{-1}[C] \setminus f^{-1}[D]$

Example. The squaring function $f : \mathbb{R} \to \mathbb{R}$ is defined by the rule $f(x) = x^2$ for all $x \in \mathbb{R}$.

Example. The *power set* of a set X is the collection of all subsets of X. Viewed as a function $\mathcal{P} : V \to V$ on the universe V of all sets, $\mathcal{P}$ is defined by the rule $\mathcal{P}(X) = \{Y : Y \subseteq X\}$.

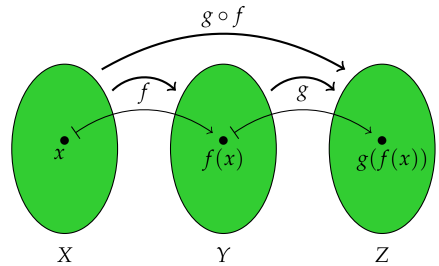


Figure 3: The function $g \circ f : X \to Z$.

Functions can be iterated with each other to produce new functions in a process called *composition* (see Figure 3). Specifically, if X, Y, and Z are sets and $f : X \to Y$ and $g : Y \to Z$ are functions, the *composite function* $g \circ f : X \to Z$ is defined[6] by $(g \circ f)(x) = g(f(x))$ for all $x \in X$. Composition of functions is *associative*; that is, for any sets W, X, Y, and Z and functions $f : W \to X$, $g : X \to Y$, and $h : Y \to Z$, we have $h \circ (g \circ f) = (h \circ g) \circ f$.

[6] Note that composition is read backwards: "$g \circ f$" means *first* apply f, *then* apply g. If we wrote $(x)f$, then we could compose functions the same way we read: from left to right. (Try it!)

Definition. If $f : X \to Y$ is a function, then an *inverse* of f is a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. The function $f : X \to Y$ is said to be *invertible* if it has an inverse.

If f is invertible, then its inverse is unique and is denoted f^{-1}. Fortunately, there is a handy way of checking* whether a function is invertible without having to know much about its inverse.

The terms *injective* and *surjective* have synonyms[3] that you might have heard of: namely, *one-to-one* and *onto*, respectively.

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

- *injective* if for all $x, x' \in X$, $x \neq x'$ implies $f(x) \neq f(x')$;
- *surjective* if for all $y \in Y$ there is $x \in X$ such that $y = f(x)$;
- *bijective* if f is both injective and surjective.

A function $f : \mathbb{R} \to \mathbb{R}$ is injective if and only if every horizontal line meets the graph of f *at most* once, and surjective if and only if every horizontal line meets the graph of f *at least* once.

Can you explain (see Figure 2!) how to think of injectivity and surjectivity in terms of the "blobs and arrows" picture?

Try proving that $f : X \to Y$ is injective if and only if there is $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and surjective if and only if there is $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$.

***Theorem.** *For any function f, f is invertible if and only if f is bijective.*

Note that for two functions to be equal to each other they must have the same domain and codomain. We can obtain new functions from a given function $f : X \to Y$ by changing $\mathrm{dom}(f)$ or $\mathrm{cod}(f)$.

While you're at it, also prove this: for any functions $f : X \to Y$ and $g : Y \to Z$,
(i) If f and g are injective, so is $g \circ f$;
(ii) If f and g are surjective, so is $g \circ f$;
(iii) If f and g are bijective, so is $g \circ f$;
(iv) If $g \circ f$ is injective, then so is f;
(v) If $g \circ f$ is surjective, then so is g.

Definition. If $f : X \to Y$ is a function and if $A \subseteq X$, the *restriction* of f to A is the function $g : A \to Y$ defined by the rule $g(x) = f(x)$ for all $x \in A$. The restriction of f to A is often denoted $f \restriction A$ or[3] $\mathrm{res}_A f$.

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be the squaring function. Then f is neither injective nor surjective, but $f \restriction [0, \infty)$ is injective, and the function $g : [0, \infty) \to [0, \infty)$ defined by $g(x) = x^2$ is bijective (thus invertible).

For any function $f : X \to Y$, the function $g : X \to \mathrm{im}(f)$ defined by $g(x) = f(x)$ for all $x \in X$ is surjective, which shows that any function can be converted into a surjective one simply by shrinking its codomain.

Lists

Recall from *The Joy of Sets* that sets do not care about order or repetition; for instance, $\{N, A, S, A\} = \{N, S, A\} = \{S, A, N, S\}$. If we want to distinguish between NASA, the NSA, and a useful bit of Latin, we will need to use finite ordered *lists* rather than sets.

As in our notation for sets, we can name a list by writing out its elements separated by commas, but in order to distinguish lists from sets we will enclose the elements between parentheses rather than between braces. The crucial difference between lists and finite sets is that order and repetition *do* matter for lists. So, for instance,

$$(N, A, S, A) \neq (N, A, S) \quad \text{and} \quad (N, A, S) \neq (N, S, A).$$

The *length* of a list is the number of elements in it. It is often convenient to index the elements of a list of length n using the natural numbers from 1 to n. That is, we might write

$$L = (x_1, \ldots, x_n) \quad \text{or} \quad L = (x_k \,:\, 1 \leq k \leq n)$$

if L is a list of length n whose kth element is x_k. Two lists are *equal* if they have the same length and the same elements, in the same order.

For us, a *list* is by definition a finite ordered set. Of course, infinite sets can be ordered as well, and an infinite ordered set that is ordered like $\mathbb{N}$ is called a *sequence*.

Thus (N, A, S, A) is a list, while $\{N, A, S, A\}$ is a set.

In linear algebra, bases are sets but finite ordered bases are lists. And sometimes we really *do* need to use ordered bases, such as when we define coordinate vectors.

Lists of length n are often called *n-tuples*, particularly when their elements are numbers.

Repetition is allowed in lists: $(1, 1, 1) \neq (1, 1)$, since these lists do not even have the same length.

Cartesian Products

Of special importance are lists of length two, which are called *ordered pairs*. In the past you have probably used ordered pairs (a, b) of real numbers to represent points in the Cartesian plane. More generally, for any sets X and Y, the *Cartesian product* of X and Y is the set

$$X \times Y \;:=\; \{(x, y) \,:\, x \in X \text{ and } y \in Y\}$$

consisting of all ordered pairs whose first element belongs to X and whose second element belongs to Y.

More generally still, we can form the *Cartesian product* of any finite list of sets $(X_1, \ldots, X_n)$, namely

$$X_1 \times \cdots \times X_n \;:=\; \{(x_1, \ldots, x_n) \,:\, x_k \in X_k \text{ for each } 1 \leq k \leq n\}.$$

As you might guess, we can also use exponential shorthand for repeated products: e.g., $X \times X = X^2$, $Y \times Y \times Y = Y^3$, etc. Thus

$$\mathbb{R}^2 \;=\; \mathbb{R} \times \mathbb{R} \;=\; \{(a, b) \,:\, a \in \mathbb{R} \text{ and } b \in \mathbb{R}\},$$

and, in general, $\mathbb{R}^n$ is the set of all n-tuples of real numbers.

René Descartes 1596–1650

You are faimilar with the *summation* symbol, which is the capital Greek letter sigma: $\sum$. The corresponding symbol for products is a capital pi: $\prod$. So we might write $X_1 \times \cdots \times X_n = \prod_{k=1}^{n} X_k$.

In linear algebra, we often refer to the n-tuples in $\mathbb{R}^n$ as *vectors*. This is because the Cartesian product $\mathbb{R}^n$ becomes a vector space once we introduce the addition and scalar multiplication operations on it, so it is natural to think of its elements as vectors. There is no contradiction in $\mathbb{R}^n$ being both a Cartesian product and a vector space, or in $\vec{x} \in \mathbb{R}^n$ being both an n-tuple and a vector. It is a bit like the fact that you are both a human being and a student at UM.

The Graph of a Function

In calculus, one of the best ways to get a visual representation of a function is to draw its graph. For instance, consider the exponential function $\exp : \mathbb{R} \to \mathbb{R}$ defined by $\exp(x) = e^x$ for all $x \in \mathbb{R}$. Its graph is a certain subset of $\mathbb{R}^2$, namely

$$\text{graph}(\exp) \;=\; \{(x, y) \in \mathbb{R}^2 \,:\, e^x = y\} \;\subseteq\; \mathbb{R}^2.$$

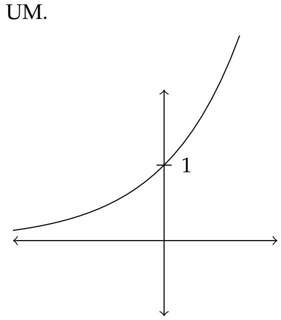


Figure 4: The graph of the exponential function $y = e^x$.

Now that we have defined Catesian products in general, there is nothing to stop us from doing this with *any* function. That is, for any function $f : X \to Y$, we define the *graph* of f to be the set

$$\mathrm{graph}(f) \;:=\; \{(x,y) \in X \times Y \;:\; f(x) = y\} \;\subseteq\; X \times Y.$$

When X and Y are arbitrary sets, we cannot really "draw" the graph of a function $f : X \to Y$ the way we would for a function from $\mathbb{R}$ to $\mathbb{R}$, but the set of points $(x,y) \in X \times Y$ such that $f(x) = y$ still makes good sense as a set.

Rigorous Definition of Function

"In mathematics rigor is not everything, but without it there is nothing." –*Henri Poincaré*

"Everything is vague to a degree you do not realize till you have tried to make it precise." –*Bertrand Russell*

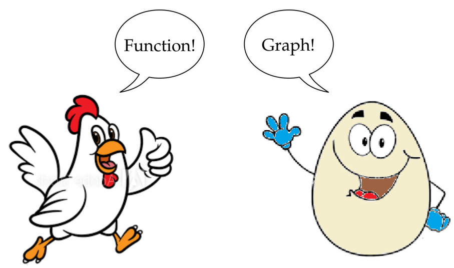
Earlier we defined a function to be a "rule," and informally this can be a useful way to think about functions, but it has some serious drawbacks that make it untenable as an official definition. Chief among them: what is a *rule*? "Rule" is not a precise mathematical notion. Furthermore, consider the functions $f, g : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \sqrt{x^2} \qquad \text{and} \qquad g(x) = \begin{cases} -x & \text{if } x < 0; \\ x & \text{if } x \geq 0. \end{cases}$$

For instance, does the rule "$f(n) =$ the least natural number that cannot be described in fewer than n words" define a function? If so, what is $f(14)$?

The functions f and g are defined by different rules, but we are inclined to say that they are the same function. This is because they have the same domain and codomain and their values agree on every input. In other words, they have the same *graph*.

In fact, the graph of a function encodes all the information we need to know about it, and is already a well-defined mathematical object. So we elect to bypass the idea of a "rule" altogether and just define a function to *be* its graph.



Definition. A *function f from X to Y* is a subset $f \subseteq X \times Y$ with the property that for every element $x \in X$ there is exactly one element $y \in Y$ such that $(x,y) \in f$.

Of course, "$(x,y) \in f$" is a bit of set-theoretic folderol that will never appear again outside this definition, since it just means "$y = f(x)$."

Sets All the Way Down

One of the goals — and one of the great achievements — of set theory is to represent literally every mathematical object as a set. In defining a function to be its graph, which is a set of ordered pairs, we have reduced the notion of function to that of ordered pair. This is enough to satisfy everyone but the set theorists, and now to satisfy them as well we show how to represent ordered pairs as sets. Given elements x and y, define the ordered pair (x,y) to be the set

$$(x,y) \;:=\; \{\{x\}, \{x,y\}\}.$$

Then for all a, b, c, d we have $(a,b) = (c,d)$ if and only if $a = c$ and $b = d$, which is all we ever needed from ordered pairs to begin with. There — now everything[7] is a set!

[7] Thanks, Bourbaki! (Essentially all mathematical objects can be represented as sets. To get a feel for how this is done, take Math 582.)