Functions – I

How much does it cost to mail this package?

That depends upon how much the package weighs.

Given two sets A and B, one particularly important class of relations between A and B for which the B variable depends in some specific way on the A variable. This is the basis for the mathematical concept of a ***function***. In the situation of the displayed text, the relationship is

cost = (rate per unit) × (weight in units, rounded up to the next integer).

Archaeological discoveries indicate that some version of the function concept was already recognized in prehistoric times, and there are many ways of describing functions formally. Here are three standard methods for doing so:

1. The use of ***tables*** to ***list the values*** of functions in terms of their dependent variables.
2. The use of ***formulas*** to ***express the values*** of functions in terms of their dependent variables.
3. The use of ***graphs*** to ***visualize the behavior*** of functions.

Each of these methods is quite old. There are tables and formulas in the writings of ancient civilizations from approximately 4000 years ago; however, the formulas are expressed in words rather than symbols. The idea of using coordinates to display functions dates back to the 14th century (at least). Both tables and graphs can be described in terms of ordered pairs, where the first coordinate represents the independent variable and the second represents the dependent variable, and this is the basis for the mathematical definition of a function.

Definition. A ***function*** is an ordered pair $f = ((A, B), \Gamma)$ where A and B are sets and Γ is a subset of $A \times B$ with the following property:

$$[!!]\ \textit{For each } a \in A \textit{ there is a } \underline{\textbf{unique}} \textit{ element } b \in B \textit{ such that } (a, b) \in \Gamma.$$

The sets A and B are respectively called the ***domain*** and ***codomain*** of f, and Γ is called the ***graph*** of f. Frequently we write $f : A \rightarrow B$ to denote a function with domain A and codomain B, and as usual we write

$$b = f(a) \ \textit{ if and only if the ordered pair } (a, b) \textit{ lies in the graph of } f.$$

By $[!!]$, for every $a \in A$ there is a unique $b \in B$ such that $b = f(a)$.

Frequently a function is simply defined to be the subset Γ described above, but in our definition the source set A (formally, this is the ***domain*** of the function) and the target set **B** (formally, this is the ***codomain*** of the function) are included explicitly as part of the structure. The domain is generally redundant. However, in some mathematical contexts if $f : A \to B$ is a function and B is a subset of C, then from our perspective it is *absolutely necessary* to distinguish between the function from A to B with graph Γ and the analogous function from A to C whose graph is also equal to Γ. One can also take this in the reverse direction; if $f : A \to B$ is a function such that its graph Γ lies in $A \times D$ for some subset $D \subset B$, then it is often either convenient or even mandatory to view the graph as also defining a related function $f : A \to D$.

The need to specify codomains is often important in computer science; for example, in computer programs one must often declare whether the numerical values of certain functions should be integer variables or real (floating point) variables.

Example of a function. If $\mathbb{R}$ is the set of real numbers, then the function f given by the standard formula $f(x) = x^2$ is given formally by $((\mathbb{R}, \mathbb{R}), P)$ where P denotes the set of all ordered pairs of real numbers (x, y) such that $y = x^2$ (a parabola). Similar considerations apply for most of the functions that arise in differential and integral calculus. On the other hand, the set of all ordered pairs (x, y) such that $x = y^2$ does **not** define a function on the real line because for some values of x there are either zero or two values of y satisfying the given equation. Even if one restricts to the nonnegative real numbers, there are two values of y corresponding to each positive value of x.

When are two functions equal? In set theory, one of the first issues is to state the standard criterion for two sets to be equal. There is also a standard criterion for two functions to be equal.

Proposition (Equal functions). Let $f : A \to B$ *and* $g : A \to B$ *be functions. Then* $f = g$ *if and only if* $f(x) = g(x)$ *for every* $x \in A$.

PROOF. If $f = g$ then their graphs are equal to the same set, which we shall call G. By definition of a function, for each $x \in A$ there is a unique $b \in B$ such that $(x, b) \in G$, and it follows that b must be equal to both $f(x)$ and $g(x)$. Conversely, if $f(x) = g(x)$ for every $x \in A$, then for each choice of x we know that the graphs of f and g both contain the ordered pair (x, b) where $b = f(x) = g(x)$. Since for each x the graphs of f and g each contain exactly one point whose first coordinate is x, it follows that these graphs are equal. By the definition of a function, this implies $f = g$.■

Composite functions

We shall begin with a few ways of constructing new functions out of given data which arise in precalculus and calculus courses.

One basic construction is to form the ***composite*** by taking a function of a function. For example, the composite of $\sin x$ and $2x + 1$ is the function $\sin(2x + 1)$, and the composite of the functions $1 + x^3$ and e^x is equal to $1 + e^{3x}$. More generally, if f and g are suitable functions, then one can form the composite $g(f(x))$ by first applying f to x and then applying g to the resulting value $f(x)$. This definition requires that the value x must be in the domain of f and $f(x)$ must be in the domain of g. To illustrate how this compatibility property might not hold, we note that over the real numbers one cannot form the composite function $\mathrm{sqrt}(\sin x - 2)$ because the expression inside the radical sign is always negative and in elementary calculus one can only define square roots for nonnegative numbers.

Formally, we proceed as follows:

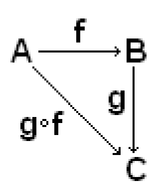
<u>Definition.</u> If $f: A \to B$ and $g: B \to C$ are functions, then the ***composite function***

$$g \circ f : A \to C$$

is defined by $g \circ f(x) = g(f(x))$. Frequently we shorten $g \circ f$ to gf.

<u>Example.</u> Suppose that $f(x) = 7x - 4$ and $g(x) = 3x + 2$. Then direct calculation shows that $g \circ f(x) = 21x - 10$.

Pictorially one often represents a composite by a so – called ***commutative diagram***, the idea being that if one follows the arrows from one object to another, the end result is independent of the path taken.



During the past century the use of commutative diagrams has become extremely widespread in the mathematical sciences and in some closely related areas (*e.g.*, some branches of theoretical physics).

Composition of functions is associative but not commutative. We shall establish the first by proving a proposition and the second by furnishing examples.

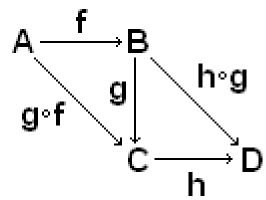
Associativity of composition. *Suppose that* $f: A \to B$, $g: B \to C$, *and* $h: C \to D$ *are functions. Then we have the associativity identity* $h \circ (g \circ f) = (h \circ g) \circ f$.

PROOF. This follows directly from the definition of functional composition. If $x \in A$ is arbitrary, then we have the following chain of equtions:

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

By the proposition on equality of functions, it follows that the two composites in the statement of the result must be equal.■

The proof may be illustrated by the following commutative diagram



in which the two triangles ΔABC and ΔBDC commute; it follows from associativity that the parallelogram $\square ABDC$ also commutes.

Failure of commutativity. One fundamental reason why composition is not commutative (*i.e.*, $g \circ f \neq f \circ g$ in general) is that the existence of one of the composites $g \circ f$ or $f \circ g$ does not guarantee the existence of the other. For example, this happens whenever we have $f: A \to B$ and $g: B \to C$ where A, B and C are all distinct. In order to define both composites we need to have $A = C$, and if $B \neq A$ there is still no way $g \circ f$ and $f \circ g$ can be equal because they still have different domains and codomains. Thus the only remaining situations in which one can ask whether the composites in both orders are equal are those where $A = B = C$. The example below shows that commutativity fails even in such a restricted setting.

Example. Let $A = B = \mathbb{R}$, let $f(x) = x + 3$, and let $g(x) = x^2$. Then the composite $g \circ f(x)$ is equal to $(x + 3)^2$, but the reverse composite $f \circ g(x)$ is equal to $x^2 + 3$, so that $g \circ f$ and $f \circ g$ are completely different functions. In particular, their values at $x = 0$ are unequal.

Images and inverse images

In working change of variables problem in calculus, it is usually necessary to find the image or the inverse image of a set under some function. Here are the formal definitions:

Definition. Let $f: A \to B$ be a function, and let $C \subset A$. Then *the image of* C *under (the mapping)* f is the set

$$f[C] \;=\; \{\, y \in B \mid y = f(x) \text{ for some } x \in C \,\}$$

Similarly, if $D \subset B$ then *the inverse image of* D *under (the mapping)* f is the set

$$f^{-1}[D] \;=\; \{\, x \in A \mid f(x) \in D \,\}.$$

The set $f[A]$, which is the image of the entire domain under f, is often called the ***image*** or ***range*** of the function.

Comment on notation. Writers often use parentheses to denote images and inverse images by $f(C)$ and $f^{-1}(D)$ rather than $f[C]$ and $f^{-1}[D]$. In most cases this should cause no confusion, but there are some exceptional situations where problems can arise, most notably if the set $Y = A$ or B contains an element x such that both $x \in A$ and $x \subset A$. Such sets are easy to manufacture; in particular, given a set x we can always form $A = x \cup \{x\}$, but in practice the replacement of brackets by parentheses is almost never a source of confusion. However, we shall consistently (try to) use square brackets to indicate images and inverse images.

By definition we know that $\{f(x)\} = f[\{x\}]$. One often also sees ***abuses of notation*** in which an inverse image of a one point set $f^{-1}[\{x\}]$ is simply written in the abbreviated form $f^{-1}(y)$. In such cases it is important to recognize that the latter is a ***subset*** of the domain and ***not an element*** of the latter (the subset may be empty or contain more than one element).

Example. Suppose that $A = B = \mathbb{R}$, $f(x) = x^2$, and C is the closed interval $[2, 3]$. Then $f[C]$ is equal to the closed interval $[4, 9]$, and if C is the closed interval $[-1, 1]$ then $f[C]$ is equal to the closed interval $[0, 1]$. Similarly, if D is the closed interval $[16, 25]$ then $f^{-1}[D]$ equals the union of the two intervals $[-5, -4]$ and $[4, 5]$, while if D is the closed interval $[-9, 4]$ then $f^{-1}[D]$ equals the closed interval $[-2, 2]$. Note that the latter is also the inverse image of $[L, 4]$ where L is an arbitrary nonpositive number because no real number whose square is negative.

In more advanced courses it is sometimes necessary to know something about the behavior of images and inverse images under set – theoretic operations such as union, intersection and taking relative complements. The main identities are summarized in the following result:

Theorem. *If $f: A \to B$ is a function, then the image and inverse image constructions for f have the following properties:*

1. *If $\mathcal{V}$ is a family of subsets of A, then $f[\cup_{C \in \mathcal{V}} C] \;=\; \cup_{C \in \mathcal{V}} f[C]$.*

2. *If $\mathcal{V}$ is also a nonempty family of subsets of A, then we have then $f[\cap_{C \in \mathcal{V}} C] \subset \cap_{C \in \mathcal{V}} f[C]$ and the containment is proper in some cases.*
3. *If C is a subset of A, then $C \subset f^{-1}[f[C]]$.*
4. *If $\mathcal{W}$ is a family of subsets of B, then we have $f^{-1}[\cup_{D \in \mathcal{W}} D] = \cup_{D \in \mathcal{W}} f^{-1}[D]$.*
5. *If $\mathcal{W}$ is also a nonempty family of subsets of B, then $f^{-1}[\cap_{D \in \mathcal{W}} D] = \cap_{D \in \mathcal{W}} f^{-1}[D]$.*
6. *If D is a subset of B, then $f[f^{-1}[D]] \subset D$.*
7. *If D is a subset of B, then $f^{-1}[B - D] = A - f^{-1}[D]$.*

PROOF. Each statement requires separate consideration.

***Verification of* (1):** Suppose that $y \in f[\cup_{C \in \mathcal{V}} C]$. Then $y = f(x)$ for some element x belonging to $\cup_{C \in \mathcal{V}} f[C]$; for the sake of definiteness say that $y \in C_0$. It follows that $y \in f[C_0]$, and since the latter is contained in $\cup_{C \in \mathcal{V}} f[C]$ it follows that the original element y belongs to $\cup_{C \in \mathcal{V}} f[C]$. Conversely, if $y \in \cup_{C \in \mathcal{V}} f[C]$ and we choose C_0 so that $y \in C_0$, then $y = f(x)$ for $x \in C_0$ and the inclusion $C_0 \subset \cup_{C \in \mathcal{V}} C$ combine to imply that $y \in f[\cup_{C \in \mathcal{V}} C]$. Hence the two sets in the statement are equal.

***Verification of* (2):** Suppose that $y \in f[\cap_{C \in \mathcal{V}} C]$. Then $y = f(x)$ for some element x belonging to $\cap_{C \in \mathcal{V}} C$, and therefore $y \in f[C]$ for each $C \in \mathcal{V}$. But this means that y belongs to $\cap_{C \in \mathcal{V}} f[C]$, and this proves the containment assertion. To see that this containment may be proper, consider the function x^2 from the real numbers to themselves, and let B and C denote the closed intervals $[-1, 1]$ and $[0, 1]$ respectively. Then $f[B \cap C] = \{0\}$ but $f[B] \cap f[C] = [0, 1]$.

***Verification of* (3):** If $x \in C$ then $f(x) \in f[C]$, and therefore $x \in f^{-1}[f[C]]$, proving the containment assertion.

***Verification of* (4):** Suppose that $x \in f^{-1}[\cup_{D \in \mathcal{W}} D]$. By definition we then know that $f(x) \in \cup_{D \in \mathcal{W}} D$, and for the sake of definiteness let us say that $f(x) \in D_0$. It follows that $x \in f^{-1}[D_0]$, and since the latter is contained in $\cup_{D \in \mathcal{W}} f^{-1}[D]$ we conclude that $f^{-1}[\cup_{D \in \mathcal{W}} D] \subset \cup_{D \in \mathcal{W}} f^{-1}[D]$.

Conversely, suppose that we have $x \in \cup_{D \in \mathcal{W}} f^{-1}[D]$. Once again, for the sake of definiteness choose D_0 so that $x \in f^{-1}[D_0]$. We then have that $f(x) \in D_0$, where the latter is contained in $\cup_{D \in \mathcal{W}} f^{-1}[D]$, so that $f(x)$ belongs to the set $\cup_{D \in \mathcal{W}} D$. This implies that $x \in f^{-1}[\cup_{D \in \mathcal{W}} D]$. Therefore we have shown that each of the sets under consideration is contained in the other and hence they must be equal.

Verification of **(5):** Suppose that $x \in f^{-1}[\cap_{D \in \mathcal{W}} D]$. Then $f(x) = y$ for some element y belonging to $\cap_{D \in \mathcal{W}} D$, so that $y \in D$ for each $D \in \mathcal{W}$. Thus we have $x \in f^{-1}[D]$ for each $D \in \mathcal{W}$, which means that x belongs to $\cap_{D \in \mathcal{W}} f^{-1}[D]$, and this proves one containment direction.

Conversely, suppose $x \in \cap_{D \in \mathcal{W}} f^{-1}[D]$. Then by definition we know that $f(x) \in D$ for every $D \in \mathcal{W}$, so that we must also have $f(x) \in \cap_{D \in \mathcal{W}} D$. But this means that $x \in f^{-1}[\cap_{D \in \mathcal{W}} D]$, proving containment in the other direction; it follows that the two sets under consideration must be equal.

Verification of **(6):** If $y \in f\left[f^{-1}[D]\right]$, then $y = f(x)$ for some $x \in f^{-1}[D]$, and by definition of the latter we know that $f(x) \in D$; since $y = f(x)$ this means that y must belong to D, proving the containment assertion.

Verification of **(7):** Suppose first that $x \in f^{-1}[B - D]$. By definition we know that $f(x) \in B - D$, and in particular it follows that $f(x) \notin D$, so that $x \notin f^{-1}[D]$. The latter in turn implies that $x \in A - f^{-1}[D]$, and thus we have established the containment of $f^{-1}[B - D]$ in $A - f^{-1}[D]$. Conversely, if $x \in A - f^{-1}[D]$, then $x \notin f^{-1}[D]$ implies $f(x) \notin D$, so that $f(x) \in B - D$ and therefore we have $x \in f^{-1}[B - D]$. This yields containment in the other direction.■

Notes. In the next lecture, we shall prove that equality holds for parts **(3)** and **(6)** if the function f satisfies an additional condition (there are separate ones for each part). Likewise, there are results for comparing $f[A - C]$ to $B - f[C]$ in ***some*** cases (see the exercises).