2.1 Sets and Mappings

Mappings, also called *functions*, are basic to mathematics and programming. Like a function in a program, a mapping in math takes an argument of one *type* and maps it to (returns) an object of a particular type. In a program, we say "type"; in math, we would identify the set. When we have an object that is a member of a set, we use the $\in$ symbol. For example,

$$a \in \mathbf{S},$$

can be read "a is a member of set $\mathbf{S}$." Given any two sets $\mathbf{A}$ and $\mathbf{B}$, we can create a third set by taking the *Cartesian product* of the two sets, denoted $\mathbf{A} \times \mathbf{B}$. This set $\mathbf{A} \times \mathbf{B}$ is composed of all possible ordered pairs (a, b) where $a \in \mathbf{A}$ and $b \in \mathbf{B}$. As a shorthand, we use the notation $\mathbf{A}^2$ to denote $\mathbf{A} \times \mathbf{A}$. We can extend the Cartesian product to create a set of all possible ordered triples from three sets and so on for arbitrarily long ordered tuples from arbitrarily many sets.

Common sets of interest include

- $\mathbb{R}$—the real numbers;
- $\mathbb{R}^+$—the nonnegative real numbers (includes zero);
- $\mathbb{R}^2$—the ordered pairs in the real 2D plane;
- $\mathbb{R}^n$—the points in n-dimensional Cartesian space;
- $\mathbb{Z}$—the integers;
- S^2—the set of 3D points (points in $\mathbb{R}^3$) on the unit sphere.

Note that although S^2 is composed of points embedded in three-dimensional space, it is on a surface that can be parameterized with two variables, so it can be thought of as a 2D set. Notation for mappings uses the arrow and a colon, for example,

$$f : \mathbb{R} \mapsto \mathbb{Z},$$

which you can read as "There is a function called f that takes a real number as input and maps it to an integer." Here, the set that comes before the arrow is called the *domain* of the function, and the set on the right-hand side is called the *target*. Computer programmers might be more comfortable with the following equivalent language: "There is a function called f which has one real argument and returns an integer." In other words, the set notation above is equivalent to the common programming notation:

$$\text{integer } f(\text{real}) \quad \leftarrow \text{equivalent} \rightarrow \quad f : \mathbb{R} \mapsto \mathbb{Z}.$$

So the colon-arrow notation can be thought of as a programming syntax. It's that simple.

The point $f(a)$ is called the *image* of a, and the image of a set A (a subset of the domain) is the subset of the target that contains the images of all points in A. The image of the whole domain is called the *range* of the function.

2.1.1 Inverse Mappings

If we have a function $f : \mathbf{A} \mapsto \mathbf{B}$, there may exist an *inverse function* $f^{-1} : \mathbf{B} \mapsto \mathbf{A}$, which is defined by the rule $f^{-1}(b) = a$ where $b = f(a)$. This definition only works if every $b \in \mathbf{B}$ is an image of some point under f (i.e., the range equals the target) and if there is only one such point (i.e., there is only one a for which $f(a) = b$). Such mappings or functions are called *bijections*. A bijection maps every $a \in \mathbf{A}$ to a unique $b \in \mathbf{B}$, and for every $b \in \mathbf{B}$, there is exactly one $a \in \mathbf{A}$ such that $f(a) = b$ (Figure 2.1). A bijection between a group of riders and horses indicates that everybody rides a single horse, and every horse is ridden. The two functions would be *rider (horse)* and *horse (rider)*. These are inverse functions of each other. Functions that are not bijections have no inverse (Figure 2.2).

An example of a bijection is $f : \mathbb{R} \mapsto \mathbb{R}$, with $f(x) = x^3$. The inverse function is $f^{-1}(x) = \sqrt[3]{x}$. This example shows that the standard notation can be somewhat awkward because x is used as a dummy variable in both f and f^{-1}. It is sometimes more intuitive to use different dummy variables, with $y = f(x)$ and $x = f^{-1}(y)$. This yields the more intuitive $y = x^3$ and $x = \sqrt[3]{y}$. An example of a function that does not have an inverse is $sqr : \mathbb{R} \mapsto \mathbb{R}$, where $sqr(x) = x^2$. This is true for two reasons: first $x^2 = (-x)^2$, and second no members of the domain map to the negative portions of the target. Note that we can define an inverse if we restrict the domain and range to $\mathbb{R}^+$. Then, $\sqrt{x}$ is a valid inverse.

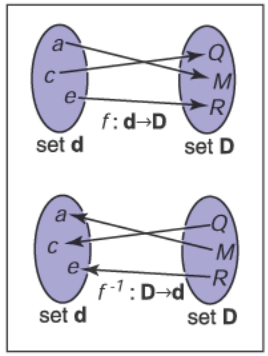

Figure 2.1. A bijection f and the inverse function f^{-1}. Note that f^{-1} is also a bijection.

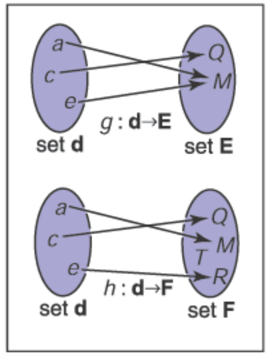

Figure 2.2. The function g does not have an inverse because two elements of **d** map to the same element of **E**. The function h has no inverse because element T of **F** has no element of **d** mapped to it.

2.1.2 Intervals

Often, we would like to specify that a function deals with real numbers that are restricted in value. One such constraint is to specify an *interval*. An example of an interval is the real numbers between zero and one, not including zero or one. We denote this $(0, 1)$. Because it does not include its endpoints, this is referred to as an *open interval*. The corresponding *closed interval*, which does contain its endpoints, is denoted with square brackets: $[0, 1]$. This notation can be mixed; i.e., $[0, 1)$ includes zero but not one. When writing an interval $[a, b]$, we assume that $a \leq b$. The three common ways to represent an interval are shown in Figure 2.3. The Cartesian products of intervals are often used. For example, to indicate that a point $\mathbf{x}$ is in the unit cube in 3D, we say $\mathbf{x} \in [0, 1]^3$.

Intervals are particularly useful in conjunction with set operations: *intersection*, *union*, and *difference*. For example, the intersection of two intervals is the set of points they have in common. The symbol $\cap$ is used for intersection. For example, $[3, 5) \cap [4, 6] = [4, 5)$. For unions, the symbol $\cup$ is used to denote points in either interval. For example, $[3, 5) \cup [4, 6] = [3, 6]$. Unlike the first two operators, the difference operator produces different results depending on argument order.

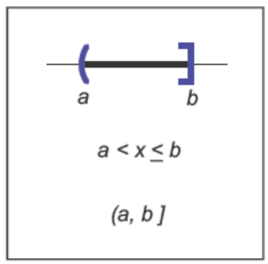

Figure 2.3. Three equivalent ways to denote the interval from a to b that includes b but not a.

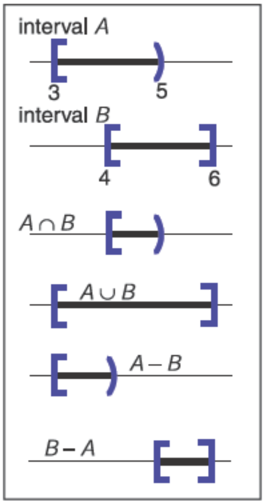

Figure 2.4. Interval operations on [3,5) and [4,6].

The minus sign is used for the difference operator, which returns the points in the left interval that are not also in the right. For example, $[3, 5) - [4, 6] = [3, 4)$ and $[4, 6] - [3, 5) = [5, 6]$. These operations are particularly easy to visualize using interval diagrams (Figure 2.4).

2.1.3 Logarithms

Although not as prevalent today as they were before calculators, *logarithms* are often useful in problems where equations with exponential terms arise. By definition, every logarithm has a *base* a. The "log base a" of x is written $\log_a x$ and is defined as "the exponent to which a must be raised to get x," i.e.,

$$y = \log_a x \quad \Leftrightarrow \quad a^y = x.$$

Note that the logarithm base a and the function that raises a to a power are inverses of each other. This basic definition has several consequences:

$$\begin{aligned} a^{\log_a(x)} &= x; \\ \log_a(a^x) &= x; \\ \log_a(xy) &= \log_a x + \log_a y; \\ \log_a(x/y) &= \log_a x - \log_a y; \\ \log_a x &= \log_a b \ \log_b x. \end{aligned}$$

When we apply calculus to logarithms, the special number $e = 2.718\ldots$ often turns up. The logarithm with base e is called the *natural logarithm*. We adopt the common shorthand ln to denote it:

$$\ln x \equiv \log_e x.$$

Note that the "$\equiv$" symbol can be read "is equivalent by definition." Like π, the special number e arises in a remarkable number of contexts. Many fields use a particular base in addition to e for manipulations and omit the base in their notation, i.e., $\log x$. For example, astronomers often use base 10 and theoretical computer scientists often use base 2. Because computer graphics borrows technology from many fields, we will avoid this shorthand.

The derivatives of logarithms and exponents illuminate why the natural logarithm is "natural":

$$\begin{aligned} \frac{d}{dx} \log_a x &= \frac{1}{x \ln a}; \\ \frac{d}{dx} a^x &= a^x \ln a. \end{aligned}$$

The constant multipliers above are unity only for $a = e$.