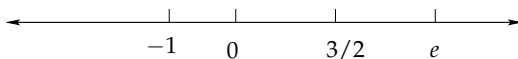


Chapter 15

Appendix

15.1 Sets

In the text we often refer to the **real numbers**. This set is denoted by $\mathbb{R}$, and we understand it to contain “all the numbers.” $\mathbb{R}$ can be visualized as the “continuous” real line:



The term “real” sounds odd but it means “not imaginary.”

$\mathbb{R}$ is an example of a **set**. A set is a collection of distinct objects viewed as a whole. In the case of $\mathbb{R}$, the objects are numbers. Other examples of sets are the set of all rectangles in the plane, or the set of all monkeys in Japan.

If A is a set, then the statement $x \in A$ means that x is contained in (alternatively, is an element of) A . If B is another set, then $A \subset B$ means that any element of A is also an element of B , and we say that A is a **subset** of B . The statement $A = B$ means that A and B contain the same elements (or, equivalently, $A \subset B$ and $B \subset A$). For example, if $\mathbb{I}$ is the irrational numbers,¹ then $\mathbb{I} \subset \mathbb{R}$. Also, $0 \in \mathbb{R}$, $\pi \in \mathbb{R}$, $-3 \in \mathbb{R}$, $e \in \mathbb{R}$, and so on.

Commonly used subsets of $\mathbb{R}$ include the intervals. For arbitrary a and b in $\mathbb{R}$, the

1. The **irrationals** are those numbers such as π and $\sqrt{2}$ that cannot be expressed as fractions of whole numbers.

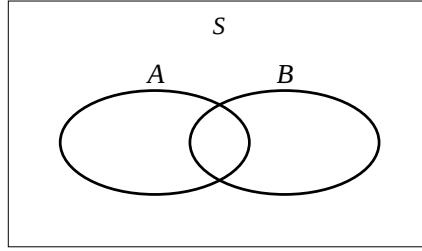


Figure 15.1 Sets A and B in S

open interval (a, b) is defined as

$$(a, b) := \{x \in \mathbb{R} : a < x < b\}$$

while the **closed interval** $[a, b]$ is defined as

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$$

We also use half open intervals such as $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$, half lines such as $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$, and so on.

Let S be a set and let A and B be two subsets of S , as illustrated in figure 15.1. The **union** of A and B is the set of elements of S that are in A or B or both:

$$A \cup B := \{x \in S : x \in A \text{ or } x \in B\}$$

Here and elsewhere, “or” is used in the mathematical sense. It means “and/or”. The **intersection** of A and B is the set of all elements of S that are in both A and B :

$$A \cap B := \{x \in S : x \in A \text{ and } x \in B\}$$

The set $A \setminus B$ is all points in A that are not points in B :

$$A \setminus B := \{x \in S : x \in A \text{ and } x \notin B\}$$

The **complement** of A is the set of elements of S that are not contained in A :

$$A^c := S \setminus A := \{x \in S : x \notin A\}$$

Figure 15.2 illustrate these definitions.

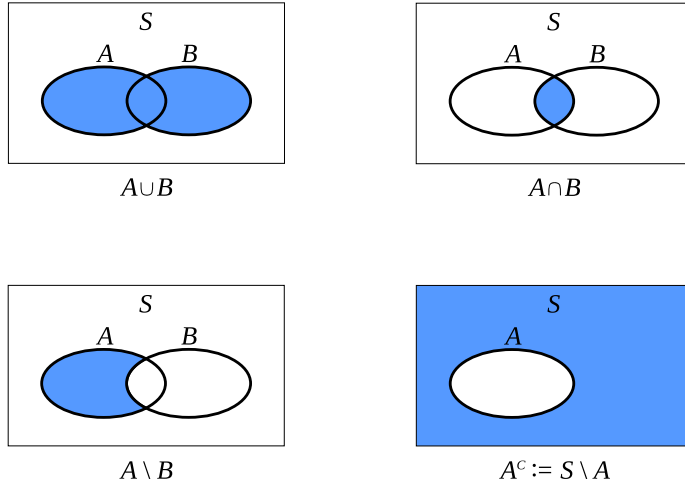


Figure 15.2 Union, intersection, and complement

For example, since $\mathbb{R}$ consists of the irrationals $\mathbb{I}$ and the rationals $\mathbb{Q}$, we have

$$\mathbb{Q} \subset \mathbb{R}, \mathbb{I} \subset \mathbb{R}, \mathbb{Q} \cup \mathbb{I} = \mathbb{R}, \mathbb{Q}^c = \mathbb{I}, \text{ etc.}$$

Also

$$\mathbb{N} := \{1, 2, 3, \dots\} \subset \mathbb{Q} \subset \mathbb{R}$$

The **empty set** is the set containing no elements. It is denoted by $\emptyset$. If the intersection of A and B equals $\emptyset$, then A and B are said to be **disjoint**.

The next fact lists some well-known rules for set-theoretic operations.

Fact 15.1.1 Let A and B be subsets of S . The following statements are true:

- (i) $A \cup B = B \cup A$ and $A \cap B = B \cap A$.
- (ii) $(A \cup B)^c = B^c \cap A^c$ and $(A \cap B)^c = B^c \cup A^c$.
- (iii) $A \setminus B = A \cap B^c$.
- (iv) $(A^c)^c = A$.

If $\{A_\alpha\}_{\alpha \in \Lambda}$ is an infinite collection of subsets of S , the union $\bigcup_{\alpha \in \Lambda} A_\alpha$ is defined as all $x \in S$ such that $x \in A_\alpha$ for some $\alpha \in \Lambda$. The intersection $\bigcap_{\alpha \in \Lambda} A_\alpha$ is all $x \in S$ that lie in A_α for all $\alpha \in \Lambda$.

Example 15.1.1 If $A_n = (-n, n)$ for each $n \in \mathbb{N}$, then $\bigcup_{n \in \mathbb{N}} A_n = \mathbb{R}$.

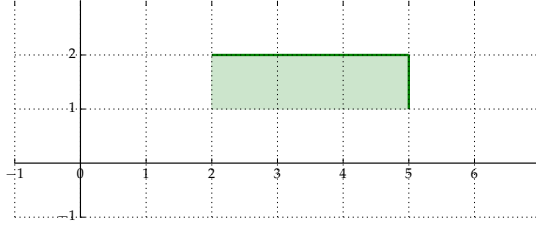


Figure 15.3 Rectangle $(2, 5] \times (1, 2]$ in $\mathbb{R}^2$

15.1.1 Cartesian Products

Collections with natural order are frequently written as **tuples**. These are finite sequences of terms, usually denoted using notation such as (a_1, a_2) or (x_1, x_2, x_3) . Collections of tuples arise naturally in Cartesian products. The **Cartesian product** of set $A_1, \dots, A_N$ is the set

$$A_1 \times \cdots \times A_N := \{(a_1, \dots, a_N) : a_n \in A_n \text{ for } n = 1, \dots, N\}$$

The **vector space** $\mathbb{R}^N$ is one example of a Cartesian product, defined as

$$\mathbb{R}^N = \mathbb{R} \times \cdots \times \mathbb{R} \quad (N \text{ times})$$

Thus $\mathbb{R}^N$ is all tuples $(x_1, \dots, x_N)$ with $x_n \in \mathbb{R}$ for each n .

A useful class of subsets of $\mathbb{R}^N$ is those that are formed as Cartesian products of intervals. These are called **rectangles**. An example is

$$I = \times_{n=1}^N (a_n, b_n] := \{(x_1, \dots, x_N) \in \mathbb{R}^N : a_n < x_n \leq b_n \text{ for } n = 1, \dots, N\}$$

where $a_1, b_1, \dots, a_N, b_N$ are given numbers. The rectangle $(2, 5] \times (1, 2]$ is shown in figure 15.3.

15.2 Functions

There are two fundamental primitives in mathematics: sets and functions.² A **function** f from set A to set B is a rule that associates to each element of A a unique element

2. Actually functions can be represented as a special type of set containing ordered pairs, and hence in pure mathematics cannot claim to be as foundational as sets. However, for our purposes, it will be fine to think of a function as a primitive in its own right.

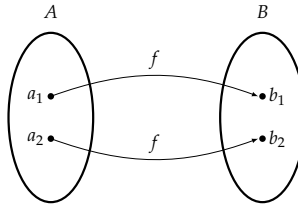


Figure 15.4 A function from A to B

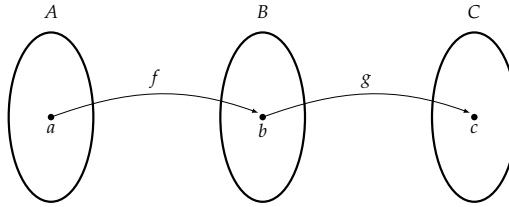


Figure 15.5 The composition $g \circ f$

of B . The symbols $f: A \rightarrow B$ means that f is a function from A to B . The set A is called the **domain** of f , while B is called the **codomain**. Figure 15.4 illustrates.

If $f: A \rightarrow B$ and $g: B \rightarrow C$ then the function $h: A \rightarrow C$ defined by $h(a) = g(f(a))$ is called the **composition** of g and f , and written as $g \circ f$.

Example 15.2.1 If $f: \mathbb{R} \rightarrow \mathbb{R}$ is defined by $f(x) = e^x := \exp(x)$ and $g: \mathbb{R} \rightarrow \mathbb{R}$ is defined by $g(x) = x^2$, then

$$(g \circ f)(x) = g(f(x)) = \exp(2x)$$

If a and b are points such that $f(a) = b$, then a is called a **preimage** of b . The set of preimages of b can be empty, a singleton or contain multiple values. The set of points

$$\{b \in B : f(a) = b \text{ for some } a \in A\}$$

is called the **range** of f , and written as $\text{rng } f$. Thus b is in the range of f if it has at least one preimage. Figure 15.6 depicts the range of a given function $f: A \rightarrow B$.

A function $f: A \rightarrow B$ is called **one-to-one** if distinct elements of A are always mapped into distinct elements of B . That is, if

$$a \in A, a' \in A, \text{ and } a \neq a' \implies f(a) \neq f(a')$$

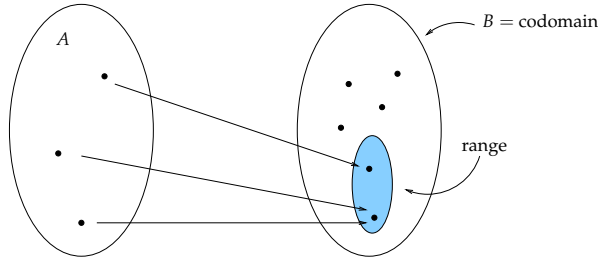


Figure 15.6 Range of $f: A \rightarrow B$

or, equivalently, if $f(a) = f(a') \implies a = a'$. Evidently $f: A \rightarrow B$ is one-to-one if and only if each element of B has at most one preimage under f . If a function $f: A \rightarrow B$ and f is one-to-one, then the equation $f(x) = b$ has at most one solution.

A function $f: A \rightarrow B$ is called **onto** if $\text{rng } f = B$; this is if every $b \in B$ has at least one preimage. If $f: A \rightarrow B$ and f is onto, then the equation $f(x) = b$ has at least one solution.

Finally, $f: A \rightarrow B$ is called a **bijection** if it is both one-to-one and onto. Bijections are sometimes called one-to-one correspondences.

Example 15.2.2 The function $f(x) = 2x$ is a bijection from $\mathbb{R}$ to $\mathbb{R}$.

If f is a bijection, then the equation $f(x) = b$ always has exactly one solution. We can therefore define the **inverse function** to f , denoted by f^{-1} , as the map from $B \rightarrow A$ such that $f^{-1}(b)$ is the unique a with $f(a) = b$.

Example 15.2.3 The function $f: A \rightarrow B$ in figure 15.4 is a bijection, with $f^{-1}: B \rightarrow A$ mapping b_1 to a_1 and b_2 to a_2 .

Example 15.2.4 If $f: \mathbb{R} \rightarrow (0, \infty)$ is defined by $f(x) = \exp(x) := e^x$ and $\phi: (0, \infty) \rightarrow \mathbb{R}$ is defined by $\phi(x) = \log(x)$, then $\phi = f^{-1}$ because, for any $a \in \mathbb{R}$, we have $\phi(f(a)) = \log(\exp(a)) = a$.

Fact 15.2.1 Let $f: A \rightarrow B$ and $g: B \rightarrow C$ be bijections.

- (i) f^{-1} is a bijection and its inverse is f .
- (ii) $f^{-1}(f(a)) = a$ for all $a \in A$.
- (iii) $f(f^{-1}(b)) = b$ for all $b \in B$.
- (iv) $g \circ f$ is a bijection from A to C and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Part (iv) of fact 15.2.1 is illustrated in figure 15.7.

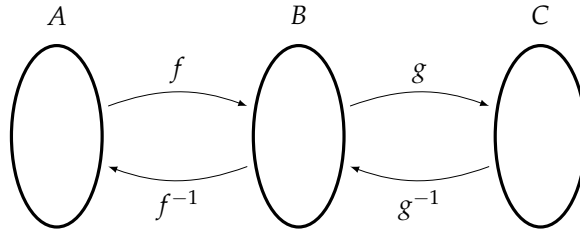


Figure 15.7 The inverse of $g \circ f$ is $f^{-1} \circ g^{-1}$

15.2.1 Preimage of Sets

Given any function f from A to B and any subset E of B , the symbol $f^{-1}(E)$ denotes the set of points that are mapped into E under f . That is,

$$f^{-1}(E) := \{x \in A : f(x) \in E\}$$

The set $f^{-1}(E)$ is called the **preimage** of E under f .

Example 15.2.5 If $f(x) = x^2$ and $E = [0, 1]$, then $f^{-1}(E) = [-1, 1]$.

Preimages of sets interact nicely with set-theoretic operations. For example, given any $E, F \subset B$, we have

- (i) $f^{-1}(E \cup F) = f^{-1}(E) \cup f^{-1}(F)$,
- (ii) $f^{-1}(E \cap F) = f^{-1}(E) \cap f^{-1}(F)$, and
- (iii) $f^{-1}(E^c) = [f^{-1}(E)]^c$.

The equalities in (i) and (ii) also extend to infinite intersections and unions.

15.3 Cardinality and Measure

For finite sets we have a fairly clear idea of what “the same size” means. For infinite sets things get trickier. It is nonetheless valuable to distinguish infinite sets according to size. To this end, when a bijection exists between sets A and B , they are said to have the **same cardinality**, and we write $|A| = |B|$.

Example 15.3.1 If $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_n\}$ then $|A| = |B|$ because $f(a_n) = b_n$ is a bijection. (Remember that elements of a set are by definition distinct.)

More generally, we have the following result:

Fact 15.3.1 If A and B are finite then $|A| = |B|$ if and only if A and B have the same number of elements.

This fact tells us that classifying sets by cardinality generalizes the notion of classifying sets by number of elements.

Fact 15.3.2 For any sets A, B, C we have $|A| = |B|$ and $|B| = |C| \implies |A| = |C|$.

In other words, cardinality is transitive. This follows from the fact that compositions of bijections are bijections (fact 15.2.1).

A nonempty set A is called **finite** if $|A| = |\{1, 2, \dots, n\}|$ for some $n \in \mathbb{N}$. Otherwise, A is called **infinite**.

Sets that either are finite or have the same cardinality as $\mathbb{N}$ are called **countable**.

Example 15.3.2 $E := \{2, 4, \dots\}$ is countable because it can be put in one-to-one correspondence with $\mathbb{N}$ via $f: E \ni m \mapsto m/2 \in \mathbb{N}$.

In essence, sets formed by sequences are countable.

Uncountable sets are sets that are not countable. For example, $\mathbb{R}$ is uncountable, as is any set that contains a nonempty open interval (a, b) . Put simply, you cannot write down a sequence that will exhaust the elements of these sets. The story is similar for $\mathbb{R}^N$. (Such sets are said to have **the power of the continuum**.) For a beautiful proof that $\mathbb{R}$ is uncountable, look up the diagonalization argument of Georg Cantor.

15.3.1 Lebesgue Measure and Sets of Measure Zero

A set $B \subset \mathbb{R}$ is said to have **Lebesgue measure zero** in $\mathbb{R}$ if, given any $\epsilon > 0$, there exists a sequence of intervals $I_n = (a_n, b_n)$ such that B is contained in the union of these intervals and $\sum_{n=1}^{\infty} (b_n - a_n) \leq \epsilon$. The definition extends to $\mathbb{R}^N$ by replacing intervals with rectangles (see page 408 for a definition of the latter).

For example, any countable set has measure zero. To see this in $\mathbb{R}$, let $C := \bigcup_{n=1}^{\infty} \{x_n\}$ with $x_n \in \mathbb{R}$. Fix $\epsilon > 0$. To each x_n we assign an interval (a_n, b_n) containing x_n and satisfying $b_n - a_n < \epsilon 2^{-n}$. The union of intervals contains C , and, in addition, the sum of their lengths is $\epsilon \sum_{n=1}^{\infty} 2^{-n} = \epsilon$.

Uncountable sets can have measure zero too. For example, every linear subspace S with $\dim S < N$ has measure zero in $\mathbb{R}^N$. A more classical example of an uncountable set with measure zero is the Cantor set. Details can be found in many texts on analysis.

The concept of measure zero sets forms part of a branch of mathematics called **measure theory**. We have used aspects of measure theory in this text—for example, in our discussion of σ -algebras in §4.1.1.2 and probabilities in §4.1.2. The basic problem of measure theory is how to provide a well-defined notion of “measure” or “size” to sets. This is easy if we just want to assign area to a circle (area = πr^2) or rectangle, or

volume to a sphere. But what about an arbitrary subset of $\mathbb{R}^N$? Is there a well-defined and widely accepted notion of size for such a set?

The answer to this question is affirmative. The so-called **Lebesgue outer measure** is a map, typically denoted λ^* , that gives a natural notion of size to any set. It works by approximating arbitrary sets by simpler sets, like rectangles, and agrees with standard notions of area or volume for common geometric objects. The only problem is that it isn't always additive, in the sense that

$$\lambda^*(A \cup B) = \lambda^*(A) + \lambda^*(B) \quad (15.1)$$

can fail. To get around this problem, we exclude any sets that cause (15.1) to fail. That is, we define the **Lebesgue measurable sets** to be those set on which additivity holds. This causes no difficulty since the discarded sets never come up in day-to-day analysis. In fact we usually focus on a smaller class of well-behaved sets, called the **Borel sets**. These were introduced in §4.1.1.2.

15.4 Real-Valued Functions

For any set A , if $f: A \rightarrow \mathbb{R}$ then f is called a **real-valued function**. If f and g are real-valued functions, then $f + g$ is defined by $(f + g)(x) = f(x) + g(x)$, while αf is defined by $(\alpha f)(x) = \alpha f(x)$. A **maximizer** of f on A is a point $a^* \in A$ such that $f(a^*) \geq f(a)$ for all $a \in A$. The value $f(a^*)$ is called the **maximum** of f on A . A **minimizer** of f on A is a point $b^* \in A$ such that $f(b^*) \leq f(a)$ for all $a \in A$. The value $f(b^*)$ is called the **minimum** of f on A .

Monotone increasing transformations of functions do not affect maximizers. To see this, let $f: \mathbb{R} \rightarrow \mathbb{R}$ and let h be **monotone increasing**, in the sense that if $x \leq x'$, then $h(x) \leq h(x')$, and let g be the function defined by $g(a) = h(f(a))$. Any maximizer of f on A is also a maximizer of g on A . To see why, let $a \in A$. Since a^* is a maximizer of f , it must be the case that $f(a) \leq f(a^*)$. Since h is monotone increasing, this implies that $h(f(a)) \leq h(f(a^*))$. Given that a was chosen arbitrarily, we have now shown that $g(a^*) \geq g(a)$ for all $a \in A$. In other words, a^* is a maximizer of g on A .

15.4.1 Sup and Inf

Maximizers and minimizers don't always exist. To illustrate, consider $f: (0, 1) \rightarrow (0, 1)$ defined by $f(x) = x$. The function f has neither maximizer nor minimizer on $(0, 1)$. For example, no $a \in (0, 1)$ is a minimizer because $b := a/2$ is in $(0, 1)$ and $f(b) < f(a)$.

To negotiate this kind of problem, we can use the notion of supremum instead. If A is a set, then the **supremum** $s := \sup A$ is the unique $s \in \mathbb{R}$ such that (i) $a \leq s$ for

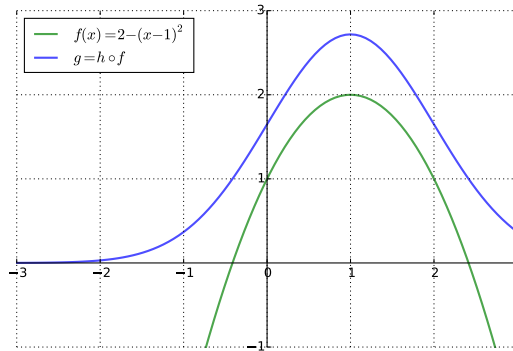


Figure 15.8 Increasing transform $h(x) = \exp(x/2)$ preserves the maximizer

every $a \in A$ and (ii) there exists a sequence $\{x_n\} \subset A$ such that $x_n \rightarrow s$. For example, 1 is the supremum of both $(0, 1)$ and $[0, 1]$. The **infimum** $i := \inf A$ is the unique $i \in \mathbb{R}$ such that (i) $a \geq i$ for every $a \in A$ and (ii) there exists a sequence $\{x_n\} \subset A$ such that $x_n \rightarrow i$. For example, 0 is the infimum of both $(0, 1)$ and $[0, 1]$.

One can show that the supremum and infimum of any bounded set A exist. The boundedness requirement can be dropped when the values $-\infty$ and ∞ are admitted as a possible infima and suprema.

Returning to our original example with $f(x) = x$, while $\max_{x \in (0,1)} f(x)$ is not well-defined, $\sup_{x \in (0,1)} f(x) := \sup\{f(x) : x \in (0,1)\} = \sup(0,1) = 1$.