An Introduction to Proofs and the Mathematical Vernacular

Martin V. Day
Department of Mathematics
Virginia Tech
Blacksburg, Virginia 24061
http://www.math.vt.edu/people/day/ProofsBook

Dedicated to the memory of my mother:

December 7, 2016

1This work is licensed under the Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/us/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.
Chapter 3

Sets and Functions

This chapter is devoted to the subject of sets, perhaps the most basic of mathematical objects. The theory of sets is a topic of study in its own right, part of the “foundations of mathematics” which involves deep questions of logic and philosophy. (See Section H.2 below for a hint of this.) For our purposes sets and functions are just another part of the language of mathematics. Sections A–F are intended mainly to familiarize you with this part of the vernacular. We will encounter a lot of notation, of which some may be new to you. This notation too is part of the vocabulary. You need to learn to use this language, not be dependent on an example or picture as a way to bypass literacy. Section G introduces some ideas regarding infinite sets. Here we will be going beyond mere terminology and will prove some interesting theorems.

A Notation and Basic Concepts

A set is simply a collection of things. The things in the set are the elements of the set. For instance, the set of multiples of 13 less than 100 is

\[ S = \{13, 26, 39, 52, 65, 78, 91\} \]

This set has exactly 7 elements. Order does not matter; \( S = \{52, 13, 26, 91, 39, 78, 65\} \) describes the same set. The mathematical notation for “is an element of” is ∈. In our example, \( 65 \in S \) but \( 66 \notin S \) (by which we mean 66 is not an element of \( S \)). A set is viewed as a single object, distinct from and of a different type than its elements. Our \( S \) is not a number; it is the new object formed by putting those specific numbers together as a “package.” Even when a set contains only one element, we distinguish between the element and the set containing the element. Thus

\[ 0 \text{ and } \{0\} \]

are two different objects. The first is a number; the second is a set containing a number.

Simple sets are indicated by listing the elements inside braces \{···\} as we have been doing above. When a pattern is clear we might abbreviate the listing by writing “. . .” For instance we might identify the set of prime numbers by writing

\[ P = \{2, 3, 5, 7, 11, 13, . . .\} \]

But to be more precise it is better to use a descriptive specification, such as

\[ P = \{n : n \text{ is a prime number}\} \]

This would be read “the set of \( n \) such that \( n \) is a prime number.” The general form for this kind of set specification is

\[ \{x : \text{criteria}(x)\} \]

where “criteria(\( x \))” is an open sentence specifying the qualifications for membership in the set. For example,

\[ T = \{x : x \text{ is a nonnegative real number and } x^2 - 2 = 0\} \]

\(^1\)Many authors prefer to write “\( \{n\mid . . .\} \), using a vertical bar instead of a colon. That’s just a different notation for the same thing.
is just a cumbersome way of specifying the set \( T = \{ \sqrt{2} \} \). In a descriptive set specification we sometimes limit the scope of the variable(s) by indicating something about it before the colon. For instance we might write our example \( T \) above as

\[
T = \{ x \in \mathbb{R} : 0 \leq x \text{ and } x^2 - 2 = 0 \},
\]

or even

\[
T = \{ x \geq 0 : x^2 - 2 = 0 \},
\]

if we understand \( x \in \mathbb{R} \) to be implicit in \( x \geq 0 \).

There are special symbols\(^2\) for many of the most common sets of numbers:

- the **natural numbers**: \( \mathbb{N} = \{ 1, 2, 3, \ldots \} = \{ n : n \text{ is a positive integer} \} \),
- the **integers**: \( \mathbb{Z} = \{ \ldots -3, -2, -1, 0, 1, 2, 3, \ldots \} = \{ n : n \text{ is an integer} \} \),
- the **rational numbers**: \( \mathbb{Q} = \{ x : x \text{ is a real number expressible as } x = \frac{n}{m}, \text{ for some } n, m \in \mathbb{Z} \} \),
- the **real numbers**: \( \mathbb{R} = \{ x : x \text{ is a real number} \} \), and
- the **complex numbers**: \( \mathbb{C} = \{ z : z = x + iy, \ x, y \in \mathbb{R} \} \).

Intervals are just special types of sets of real numbers, for which we have a special notation.

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

\[
(b, +\infty) = \{ x \in \mathbb{R} : b < x \}.
\]

In particular \((-\infty, +\infty)\) is just another notation for \( \mathbb{R} \).

Another special set is the **empty set**

\[
\emptyset = \{ \} ,
\]

the set with no elements at all. Don’t confuse it with the number 0, or the set containing 0. **All of the following are different mathematical objects.**

\[0, \emptyset, \{0\}, \{\emptyset\}\]

\[\text{B Basic Operations and Properties}\]

**Definition.** Suppose \( A \) and \( B \) are sets. We say \( A \) is a subset of \( B \), and write \( A \subseteq B \), when every element of \( A \) is also an element of \( B \). In other words, \( A \subseteq B \) means that

\[
x \in A \text{ implies } x \in B.
\]

For instance, \( \mathbb{N} \subseteq \mathbb{Z} \), but \( \mathbb{Z} \subseteq \mathbb{N} \) is a false statement. (You will often find “\( A \subseteq B \)” written instead of “\( A \subseteq B \)”. They mean the same thing; it’s just a matter of the author’s preference.) No matter what the set \( A \) is, it is always true that

\[
\emptyset \subseteq A.
\]

(This is in keeping with our understanding that vacuous statements are true.) To say \( A = B \) means that \( A \) and \( B \) are the same set, that is they contain precisely the same elements:

\[
x \in A \text{ if and only if } x \in B,
\]

which means the same as

\[A \subseteq B \text{ and } B \subseteq A.\]

This provides the typical way of proving \( A = B \): prove containment both ways. See the proof of Lemma 3.1 part d) below for an example.

Starting with sets \( A \) and \( B \) there are several operations which form new sets from them.

\(^2\)The choice of \( \mathbb{Q} \) for the rational numbers is suggested by \( \mathbb{Q} \) for “quotient.” The choice of \( \mathbb{Z} \) for the integers comes from the German term “zahlen” for numbers.
**Definition.** Suppose $A$ and $B$ are sets. The *intersection* of $A$ and $B$ is the set

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

The *union* of $A$ and $B$ is the set

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

The set difference, $A$ *remove* $B$, is the set

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

We can illustrate these definitions by imagining that $A$ and $B$ are sets of points inside two circles in the plane, and shade the appropriate regions to indicate the newly constructed sets. Such illustrations are called *Venn diagrams*. Here are Venn diagrams for the definitions above.

![Venn Diagrams](image)

Don’t let yourself become dependent on pictures to work with sets. For one thing, not all sets are geometrical regions in the plane. Instead you should try to work in terms of the definitions, using precise logical language. For instance $x \in A \cap B$ means that $x \in A$ and $x \in B$.

**Example 3.1.** For $A = \{1, 2, 3\}$ and $B = \{2, 4, 6\}$ we have

$$A \cap B = \{2\}, \quad A \cup B = \{1, 2, 3, 4, 6\}, \quad A \setminus B = \{1, 3\}, \quad \text{and } B \setminus A = \{4, 6\}.$$  

When two sets have no elements in common, $A \cap B = \emptyset$ and we say the two sets are *disjoint*. When we have three or more sets we say they are disjoint if every pair of them is disjoint. That is *not* the same as saying their combined intersection is empty.

**Example 3.2.** Consider $A = \{1, 2, 3\}$, $B = \{2, 4, 6\}$, and $C = \{7, 8, 9\}$. These are not disjoint (because $A \cap B \neq \emptyset$), even though $A \cap B \cap C = \emptyset$.

We want to say that the complement of a set $A$, to be denoted\(^3\) $\bar{A}$, is the set of all things which are not elements of $A$. But for this to be meaningful we have to know the allowed scope of all possible elements under consideration. For instance, if $A = \{1, 2, 3\}$, is $\bar{A}$ to contain all natural numbers that are not in $A$, or all integers that are not in $A$, or all real numbers that are not in $A$, or...? It depends on the context in which we are working. If the context is all natural numbers, then $\bar{A} = \{4, 5, 6, \ldots\}$. If the context is all integers, then $\bar{A} = \{\ldots, -3, -2, -1, 0, 4, 5, 6, \ldots\}$. If the context is all real numbers $\mathbb{R}$, then

$$\bar{A} = (\infty, -1) \cup (1, 2) \cup (2, 3) \cup (3, +\infty).$$  

The point is that the complement of a set is always determined relative to an understood context of what the scope of all possible elements is.

**Definition.** Suppose $X$ is the set of all elements which are allowed as elements of sets. The *complement* of a set $A \subseteq X$ is

$$\bar{A} = X \setminus A.$$  

Here is a Venn diagram. The full box illustrates $X$. The shaded region is $\bar{A}$.

---

\(^3\)Other common notations are $A^c$ and $A'$.
In the context of the real numbers for instance,
\[
\widetilde{(a, b)} = (-\infty, a] \cup [b, +\infty).
\]

**Lemma 3.1.** Suppose \( X \) is the set of all elements under consideration, and that \( A, B, \) and \( C \) are subsets of \( X \). Then the following hold.

a) \( A \cup B = B \cup A \) and \( A \cap B = B \cap A \).

b) \( A \cup (B \cup C) = (A \cup B) \cup C \) and \( A \cap (B \cap C) = (A \cap B) \cap C \).

c) \( A \cap (B \cap C) = (A \cap B) \cap (A \cap C) \) and \( A \cup (B \cup C) = (A \cup B) \cup (A \cup C) \).

d) \( (A \cup B)^{\sim} = \tilde{A} \cap \tilde{B} \) and \( (A \cap B)^{\sim} = \tilde{A} \cup \tilde{B} \).

e) \( A \cup \tilde{A} = X \) and \( A \cap \tilde{A} = \emptyset \).

f) \( (\tilde{A}) = A \).

Proofs of these statements consist of just translating the symbols into logical language, working out the implications, and then translating back. We will prove the first part of d) as an example.

**Proof.** To prove the first part of d), we prove containment both ways.

Suppose \( x \in (A \cup B)^{\sim} \). This means \( x \in X \) and \( x \not\in A \cup B \). Since \( (x \in A \lor x \in B) \) is false, we know \( x \notin A \) and \( x \notin B \). So we know that \( x \in \tilde{A} \) and \( x \in \tilde{B} \), which means that \( x \in \tilde{A} \cap \tilde{B} \). This shows that \( (A \cup B)^{\sim} \subseteq \tilde{A} \cap \tilde{B} \).

Now assume that \( x \in \tilde{A} \cap \tilde{B} \). This means that \( x \in \tilde{A} \) and \( x \in \tilde{B} \), and so \( x \in X \), \( x \notin A \), and \( x \notin B \). Since \( x \) is in neither \( A \) nor \( B \), \( x \notin A \cup B \). We find therefore that \( x \in (A \cup B)^{\sim} \). This shows that \( x \in \tilde{A} \cap \tilde{B} \subseteq (A \cup B)^{\sim} \).

Having proven containment both ways we have established that \( (A \cup B)^{\sim} = \tilde{A} \cap \tilde{B} \). \( \square \)

**Problem 3.1** If \( x \) and \( y \) are real numbers, \( \min(x, y) \) refers to the smaller of the two numbers. If \( a, b \in \mathbb{R} \), prove that
\[
\{x \in \mathbb{R} \mid x \leq a\} \cap \{x \in \mathbb{R} \mid \min(x, a) \leq b\} = \{x \in \mathbb{R} \mid x \leq \min(a, b)\}.
\]

**Problem 3.2** Prove part c) of the above lemma.

**Problem 3.3** Prove that for any two sets, \( A \) and \( B \), the three sets
\( A \cap B, A \setminus B, \) and \( B \setminus A \)
are disjoint.

\textbf{Problem 3.4} The \textit{symmetric difference} of two sets $A, B$ is defined to be

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

a) Draw a Venn diagram to illustrate $A \triangle B$.

b) Prove that $\overline{A} \triangle B = (A \cap B) \cup (\overline{A} \cap \overline{B})$.

c) Is it true that $\overline{A} \triangle B = \overline{A} \triangle \overline{B}$? Either prove or give a counterexample.

d) Draw a Venn diagram for $(A \triangle B) \triangle C$.

\textbf{Indexed Families of Sets}

Sometimes we need to work with many sets whose descriptions are similar to each other. Instead of giving them all distinct names, \textit{“A, B, C, . . .”} it may be more convenient to give them the same name but with a subscript which takes different values.

\textit{Example 3.3.} Let

$$A_1 = (-1, 1), \quad A_2 = (-\frac{1}{2}, \frac{1}{2}), \quad A_3 = (-\frac{1}{3}, \frac{1}{3}) \quad \ldots ;$$

in general for $k \in \mathbb{N}$,

$$A_k = (-\frac{1}{k}, \frac{1}{k}).$$

In the above example, the \textit{“k”} in \textit{“$A_k$”} is the \textit{index}. The set of allowed values for the index is called the \textit{index set}. In this example the index set is $\mathbb{N}$. The collection $A_k, k \in \mathbb{N}$ is an example of an \textit{indexed family} of sets. We can form the union of an indexed family, but instead of writing

$$A_1 \cup A_2 \cup A_3 \cup \ldots$$

we write

$$\bigcup_{k \in \mathbb{N}} A_k$$

or, for this particular index set, $\bigcup_{k=1}^{\infty} A_k$.

Similarly,

$$\bigcap_{k \in \mathbb{N}} A_k = A_1 \cap A_2 \cap A_3 \cap \ldots ,$$

We will see some indexed families of sets in the proof of the Schroeder-Bernstein Theorem below. Here are some simpler examples.

\textit{Example 3.4.} Continuing with Example 3.3, we have $\bigcup_{k=1}^{\infty} A_k = (-1, 1)$, and $\bigcap_{k \in \mathbb{N}} A_k = \{0\}$.

\textit{Example 3.5.} $\mathbb{R} \setminus \mathbb{Z} = \bigcup_{k \in \mathbb{Z}} I_k$, where $I_k = (k, k + 1)$.

There is no restriction on what can be used for an index set.

\textit{Example 3.6.} Let $U = (0, \infty)$, the set of positive real numbers. For each $r \in U$ let

$$S_r = \{(x, y) : x, y \text{ are real numbers with } |x|^r + |y|^r = 1\}.$$
The figure at right illustrates some of the sets $S_r$. The set $\cup_{r \in U} S_r$ is somewhat tricky to describe:

$$\cup_{r \in U} S_r = (B \setminus A) \cup C,$$

where

$$B = \{(x, y) : |x| < 1 \text{ and } |y| < 1\}$$

$$A = \{(x, y) : x = 0 \text{ or } y = 0\}$$

$$C = \{(0, 1), (0, -1), (1, 0), (-1, 0)\}.$$  

Problem 3.5

a) For each $x \in \mathbb{R}$ let $C_x = \{y : x^2 + y^2 \leq 1\}$. What is $\cup_{x \in \mathbb{R}} C_x$? What is $\cap_{x \in \mathbb{R}} C_x$?

b) Let $M_n = \{k \in \mathbb{N} : k = nm \text{ for some integer } m > 1\}$. What is $\cup_{n \in \mathbb{N}} M_n$?

c) Suppose $A \subseteq \mathbb{R}$ and for each $\varepsilon > 0$ let $I_\varepsilon = \{a \in \mathbb{R} : (a - \varepsilon, a + \varepsilon) \subseteq A\}$. Is it true that $A = \cup_{\varepsilon > 0} I_\varepsilon$? Explain.

d) Let $S_\varepsilon = \{n \in \mathbb{N} : \sin(n) > 1 - \varepsilon\}$. For $\varepsilon > 0$ is $S_\varepsilon$ finite or infinite? What is $\cap_{\varepsilon > 0} S_\varepsilon$?

Problem 3.6 For each $x \in \mathbb{R}$, let $P_x$ be the set

$$P_x = \{y : y = x^n \text{ for some } n \in \mathbb{N}\}.$$  

a) There are exactly three values of $x$ for which $P_x$ is a finite set. What are they?

b) Find $\cap_{0 < x < 1} P_x$ and $\cup_{0 < x < 1} P_x$.

c) For a positive integer $N$, find $\cap_{k=1}^N P_{2^k}$. Find $\cap_{k=1}^\infty P_{2^k}$.

C Product Sets

Sets don’t care about the order of their elements. For instance,

$$\{1, 2, 3\} \text{ and } \{3, 1, 2\}$$

are the same set. You might think of a set as an unordered list of elements. An ordered list of numbers is a different kind of thing. For instance if we are thinking of $(1, 2)$ and $(2, 1)$ as the coordinates of points in the plane, then the order matters. When we write $(x, y)$ we mean the ordered pair of numbers. The use of parentheses indicates that we mean ordered pair, not set.

If $A$ and $B$ are sets, we can consider ordered pairs $(a, b)$ where $a \in A$ and $b \in B$.

**Definition.** Suppose $A$ and $B$ are sets. The set of all such ordered pairs $(a, b)$ with $a \in A$ and $b \in B$ is called the **Cartesian product** of $A$ and $B$:

$$A \times B = \{(a, b) : a \in A, \ b \in B\}.$$
Don’t let the use of the word “product” and the notation “×” mislead you here — there is no multiplication involved. The elements of \( A \times B \) are a different kind of object than the elements of \( A \) and \( B \). For instance the elements of \( \mathbb{R} \) and \( \mathbb{Z} \) are individual numbers (the second more limited that the first), but an element of \( \mathbb{R} \times \mathbb{Z} \) is an ordered pair of two numbers (not the result of multiplication). For instance \((\pi, -3) \in \mathbb{R} \times \mathbb{Z}\).

**Example 3.7.** Let \( X = \{-1, 0, 1\} \) and \( Y = \{\pi, e\} \). Then
\[
X \times Y = \{(-1, \pi), (0, \pi), (1, \pi), (-1, e), (0, e), (1, e)\}.
\]

We can do the same thing with more than two sets: for a set of ordered triples we would write
\[
A \times B \times C = \{(a, b, c) : a \in A, b \in B, c \in C\}.
\]

**Example 3.8.** If \( \Gamma = \{a, b, c, d, e, \ldots, z\} \) is the English alphabet (thought of as a set) then \( \Gamma \times \Gamma \times \Gamma \times \Gamma \) is the (notorious) set of four-letter words (including ones of no known meaning).

When we form the Cartesian product of the same set with itself, we often write
\[
A^2
\]
as an alternate notation for \( A \times A \).

So the coordinates of points in the plane make up the set \( \mathbb{R}^2 \). The set of all possible coordinates for points in three dimensional space is \( \mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \). The set of four-letter words in the example above is \( \Gamma^4 \). The next lemma lists some basic properties of Cartesian products.

**Lemma 3.2.** Suppose \( A, B, C, D \) are sets. The following hold.

a) \( A \times (B \cup C) = (A \times B) \cup (A \times C) \).

b) \( A \times (B \cap C) = (A \times B) \cap (A \times C) \).

c) \( (A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D) \).

d) \( (A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D) \).

As an example of how this sort of thing is proven, here is a proof of part b).

**Proof:** Suppose \((x, y) \in A \times (B \cap C)\). This means \(x \in A\) and \(y \in B \cap C\). Since \( y \in B \) it follows that \((x, y) \in A \times B\). Similarly, since \( y \in C \) it follows that \((x, y) \in A \times C\) as well. Since \((x, y)\) is in both \(A \times B\) and \(A \times C\), we conclude that \((x, y) \in (A \times B) \cap (A \times C)\). This proves that \(A \times (B \cap C) \subseteq (A \times B) \cap (A \times C)\).

Suppose \((x, y) \in (A \times B) \cap (A \times C)\). Since \((x, y) \in A \times B\) we know that \(x \in A\) and \(y \in B\). Since \((x, y) \in A \times C\) we know that \(x \in A\) and \(y \in C\). Therefore \(x \in A\) and \(y \in B \cap C\), and so \((x, y) \in A \times (B \cap C)\).

This proves that \((A \times B) \cap (A \times C) \subseteq A \times (B \cap C)\).

Having proven containment both ways, this proves b) of the lemma. \(\square\)

The next example shows why part d) is only a subset relation, not equality, in general.

**Example 3.9.** Consider \( A = \{a\}, B = \{b\}, C = \{c\}, D = \{d\} \) where \(a, b, c, d\) are distinct. Observe that \((a, d)\) is an element of \((A \cup C) \times (B \cup D)\), but not of \((A \times B) \cup (C \times D)\).

**Problem 3.7** Prove part c) of the above lemma.
D The Power Set of a Set

We can also form sets of sets. For instance if \( A = \{1\} \), \( B = \{1, 2\} \) and \( C = \{1, 2, 3\} \), we can put these together into a new set

\[ \mathcal{F} = \{A, B, C\} . \]

This \( \mathcal{F} \) is also a set, but of a different type than \( A \), \( B \), or \( C \). The elements of \( C \) are integers; the elements of \( \mathcal{F} \) are sets of integers, which are quite different. We have used a script letter “\( \mathcal{F} \)” to emphasize the fact that it is a different kind of set than \( A \), \( B \), and \( C \). While it is true that \( A = \{1\} \in \mathcal{F} \), it is not true that \( 1 \in \mathcal{F} \).

Starting with any set, \( A \) we can form its power set, namely the set of all subsets of \( A \).

Definition. Suppose \( A \) is a set. The power set of \( A \) is

\[ \mathcal{P}(A) = \{B : B \subseteq A\} . \]

As an example

\[ \mathcal{P}(\{1, 2, 3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} . \]

Observe that \( B \in \mathcal{P}(A) \) means that \( B \subseteq A \). We will have something interesting to say about the power set of an infinite set in Section G.2.

We could keep going, and form the power set of the power set, and so forth. But if you carelessly wander too far down that path you will find yourself in a logical quagmire! See Section H.2.

E Relations

There are many ways we can compare two mathematical objects of the same type. Here are some examples.

- inequality for real numbers: \( x \leq y \),
- containment for sets: \( A \subseteq B \),
- division for integers: \( n \mid m \) (defined on page 15 above),
- equality of integers mod \( k \): \( n \equiv_k m \) (defined in Section D of the next chapter).

Each of these is an example of what we call a relation between two objects of the same type. If \( X \) is a set, a relation on \( X \) is really an open sentence \( R(x, y) \) taking two variables \( x, y \in X \). For the frequently used relations we have special symbols (like the examples above) that we write between the two arguments instead of in front of them: \( \leq \) instead of \( \leq(x, y) \). We will follow this pattern and write \( x R y \) instead of \( R(x, y) \) for our discussion below.

A statement of relation \( x R y \) does not refer to some calculation we are to carry out using \( x \) and \( y \). It is simply a statement which has a well-defined truth value for each pair \( (x, y) \in X \times X \). Whether it is true or false depends on the specific choices for \( x, y \in X \). For instance consider again the inequality relation, \( \leq \). The statement \( x \leq y \) is true for \( x = 2 \) and \( y = 3 \), but false for \( x = 3 \) and \( y = 2 \). (This also shows that order matters; in general \( x R y \) is not the same statement as \( y R x \).)

We can make up all kinds of strange relations; you will find several in the problems below. The important ones are those which describe a property which is significant for some purpose, like the examples above. Most useful relations have one or more of the following properties.

Definition. Suppose \( R \) is a relation on a set \( X \) and \( x, y, z \in X \).

- \( R \) is called reflexive when \( x R x \) is true for all \( x \in X \).
- \( R \) is called symmetric when \( x R y \) is equivalent to \( y R x \).
- \( R \) is called transitive when \( x R y \) and \( y R z \) together imply \( x R z \).

Example 3.10.

- Inequality \( (\leq) \) on \( \mathbb{R} \) is transitive and reflexive, but not symmetric.
• Strict inequality (<) on \( \mathbb{R} \) is transitive, but not reflexive or symmetric.

• Define the coprime relation \( C \) on \( \mathbb{N} \) so that \( n \ C \ m \) means that \( n \) and \( m \) share no common positive factors other than 1. Then \( C \) is symmetric, but not reflexive or transitive.

Example 3.11. Define the lexicographic order relation on \( \mathbb{R}^2 \) by \( (x, y) \ L (u, v) \) when \( x < u \) or \( (x = u \text{ and } y \leq v) \). Then \( L \) is transitive and reflexive, but not symmetric. Let’s write out the proof that \( L \) is transitive.

Suppose \( (x, y) \ L (u, v) \) and \( (u, v) \ L (w, z) \), where \( x, y, u, v, w, z \in \mathbb{R} \). Our goal is to show that \( (x, y) \ L (w, z) \).

We know that \( x \leq u \) and \( u \leq w \). If either of these is a strict inequality, then \( x < w \) and therefore \( (x, y) \ L (w, z) \). Otherwise \( x = u = w, y \leq v, \) and \( v \leq z \). But then \( x = w \) and \( y \leq z \), which imply \( (x, y) \ L (w, z) \). So in either case we come to the desired conclusion.

Problem 3.8 For each of the following relations on \( \mathbb{R} \), determine whether or not it is reflexive, symmetric, and transitive and justify your answers.

a) \( x \triangle y \) means \( xy = 0 \).

b) \( x \triangledown y \) means \( xy \neq 0 \).

c) \( x \rtriangleright y \) means \( |x - y| < 5 \).

d) \( x \circ y \) means \( x^2 + y^2 = 1 \).

Definition. A relation \( R \) is called an equivalence relation when it is symmetric, reflexive, and transitive.

Equivalence relations are especially important because they describe some notion of “sameness.” For instance, suppose we are thinking about angles \( \theta \) in the plane. We might start by saying an angle is any real number \( x \in \mathbb{R} \). But that is not quite what we mean: \( \pi/3 \) and \( 7\pi/3 \) are different as numbers, but they are the same as angles, because they differ by a multiple of \( 2\pi \). The next example defines an equivalence relation on \( \mathbb{R} \) that expresses this idea of sameness as angles.

Example 3.12. Let \( \odot \) be the relation on \( \mathbb{R} \) defined by

\[
    x \odot y \text{ means that there exists } k \in \mathbb{Z} \text{ so that } x = 2k\pi + y.
\]

For instance \( \pi/3 \odot 7\pi/3 \), because \( \pi/3 = 2(-1)\pi + 7\pi/3 \) so that the definition holds with \( k = -1 \).

This is an equivalence relation, as we will now check. Since \( x = 2 \cdot 0 \cdot \pi + x \), the relation is reflexive. If \( x = 2k\pi + y \), then \( y = 2(-k)\pi + x \) and \( -k \in \mathbb{Z} \) if \( k \) is. This proves symmetry. If \( x = 2k\pi + y \) and \( y = 2m\pi + z \), then \( x = 2(k + m)\pi + z \), showing that the relation is transitive.

Definition. Suppose \( R \) is an equivalence relation on \( X \) and \( x \in X \). The equivalence class of \( x \) is the set

\[
    [x]_R = \{ y \in X : xRy \}.
\]

The \( y \in [x]_R \) are called representatives of the equivalence class.

When it is clear what equivalence relation is intended, often we leave it out of the notation and just write \( [x] \) for the equivalence class. The equivalence classes “partition” \( X \) into to the sets of mutually equivalent elements.

Example 3.13. Continuing with Example 3.12, the equivalence classes of \( \odot \) are sets of real numbers which differ from each other by multiples of \( 2\pi \).

\[
[x] = \{ \ldots, x - 4\pi, x - 2\pi, \pi, x, x + 2\pi, x + 4\pi, \ldots \}.
\]

For instance \( \pi/3 \) and \( 7\pi/3 \) belong to the same equivalence class, which we can refer to as either \( [\pi/3] \) or \( [7\pi/3] \) — both refer to the same set of numbers. But \( [\pi/3] \) and \( [\pi] \) are different.
One of the uses of equivalence relations is to define new mathematical objects by disregarding irrelevant properties or information. The equivalence relation \( x \circ y \) of the above examples defines exactly what we mean by saying \( x \) and \( y \) represent the “same angle.” If we have a particular angle in mind, then the set of all the \( x \) values corresponding to that angle form one of the equivalence classes of \( \circ \). If we want a precise definition of what an “angle” actually is (as distinct from a real number), the standard way to do it is say that an angle \( \theta \) is an equivalence class of \( \circ \): \( \theta = [x] \). The \( x \in \theta \) are the representatives of the angle \( \theta \). The angle \( \theta \) is the set of all \( x \) which represent the same angle. In this way we have defined a new kind of object (angle) by basically gluing together all the equivalent representatives of the same object and considering that glued-together lump (the equivalence class) as the new object itself. This will be the basis of our discussion of the integers mod \( m \) in Section D of the next chapter.

Problem 3.9  Show that each of the following is an equivalence relation.

a) On \( \mathbb{N} \times \mathbb{N} \) define \((j, k) \parallel (m, n)\) to mean \(jn = km\). (From [10].)

b) On \( \mathbb{R} \times \mathbb{R} \) define \((x, y) \Uparrow (u, v)\) to mean \(x^2 - y = u^2 - v\).

c) On \((0, \infty)\) define \(x \lessdot y\) to mean \(x/y \in \mathbb{Q}\).

For a) and b), give a geometrical description of the the equivalence classes.

Problem 3.10  Suppose \( \vdash \) is an equivalence relation on \( X \). For \( x, y \in X \) let \([x]\) and \([y]\) be their equivalence classes with respect to \( \vdash \).

a) Show that \( x \in [x] \).

b) Show that either \([x] = [y]\), or \([x]\) and \([y]\) are disjoint.

c) Show that \( x \vdash y \) iff \([x] = [y]\).

d) Consider the relation \( \Diamond \) on \( \mathbb{R} \) from Problem 3.8. (Recall that this is not an equivalence relation.) Define \([x] = \{y \in \mathbb{R} : x \Diamond y\}\).

Which of a), b) and c) above are true if \([x]\) and \([y]\) are replaced by \(\langle x\rangle\) and \(\langle y\rangle\) and \(\vdash\) is replaced by \(\Diamond\)?
Example 3.14. Here are some examples to make the point that the elements of the domain and codomain can be more complicated than just numbers.

a) Let \( \mathbb{R}^{2 \times 2} \) denote the set of \( 2 \times 2 \) matrices with real entries, \( M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \) and \( \det(M) = \text{the usual determinant:} \)

\[
\det(M) = m_{11}m_{22} - m_{12}m_{21}.
\]

Then \( \det : \mathbb{R}^{2 \times 2} \to \mathbb{R} \).

b) Let \( C([0,1]) \) be the set of all continuous functions \( f : [0,1] \to \mathbb{R} \). We can view the integral \( I(f) = \int_0^1 f(x) \, dx \) as a function \( I : C([0,1]) \to \mathbb{R} \).

c) For any \( c = (c_1,c_2,c_3) \in \mathbb{R}^3 \) we can form a quadratic polynomial \( P(c) = c_3x^2 + c_2x + c_1 \). We can view \( P \) as a function \( P : \mathbb{R}^3 \to C([0,1]) \) from ordered triples to continuous functions.

Definition. Suppose \( f : A \to B \) is a function. The range of \( f \) is the set

\[
\text{Ran}(f) = \{ b : \text{there exists } a \in A \text{ such that } b = f(a) \}.
\]

Don’t confuse the codomain of a function with its range. The range of \( f \) is the set of values in \( b \in B \) for which there actually does exist an \( a \in A \) with \( f(a) = b \). The codomain \( B \) can be any set containing the range as a subset. In general there is no presumption that \( \text{Ran}(f) \) is all of \( B \). What we understand the codomain to be does affect how we answer some questions, as we will see shortly. When \( \text{Ran}(f) = B \) we say \( f \) is onto, or surjective; see the definition below. In particular whether \( f \) is onto or not depends on what we understand the codomain to be.

Definition. Suppose \( f : A \to B \) is a function.

a) \( f \) is said to be surjective (or a surjection, or onto) when for every \( b \in B \) there exists an \( a \in A \) for which \( b = f(a) \).

b) \( f \) is said to be injective (or an injection, or one-to-one) when \( a = a' \) is necessary for \( f(a) = f(a') \).

c) When \( f \) is both surjective and injective, we say \( f \) is bijective (or a bijection).

Example 3.15. Suppose we consider the function \( f(x) = x^2 \) for the following choices of domain \( A \) and codomain \( B \): \( f : A \to B \).

a) \( A = [0, \infty) \) and \( B = \mathbb{R} \). This makes \( f \) injective but not surjective. To prove that it is injective, suppose \( a, a' \in A \) and \( f(a) = f(a') \). This means \( a, a' \geq 0 \) and \( a^2 = (a')^2 \). It follows that \( a = a' \) just as in the proof of Proposition 1.3. This proves that \( f \) is injective. To see that \( f \) is not surjective, simply observe that \(-1 \in B \) but there is no \( a \in A \) with \( f(a) = -1 \).

b) Now take \( A = \mathbb{R} \) and \( B = \mathbb{R} \). For these choices \( f \) is still not surjective, but is not injective either, because \( f(1) = f(-1) \).

c) Consider \( A = \mathbb{R} \) and \( B = [0, \infty) \). Compared to b) all we have done is change what we understand the codomain to be. As in b) \( f \) is not injective, but now is surjective, because every \( b \in B \) does have a square root: \( a = \sqrt{b} \) is in \( A \) and \( f(a) = b \).

When we have two functions, \( f \) and \( g \), we can sometimes follow one by the other to obtain the composition of \( f \) with \( g \).
**Definition.** Suppose \( f : A \to B \) and \( g : C \to D \) are functions, and that \( B \subseteq C \). Their **composition** is the function \( g \circ f : A \to D \) defined by
\[
 g \circ f(a) = g(f(a)) \quad \text{for all } a \in A.
\]

Observe that we must have \( B \subseteq C \) in order for \( g(f(a)) \) to be defined.

**Example 3.16.** Let \( f : [0, \infty) \to \mathbb{R} \) be defined by \( f(x) = \sqrt{x} \) and \( g : \mathbb{R} \to \mathbb{R} \) defined by \( g(x) = \frac{x^2}{x^2 + 1} \). Then \( g \circ f(x) = \frac{x}{x^2 + 1} \), for \( x \geq 0 \).

**Proposition 3.3.** Suppose \( f : A \to B \) and \( g : C \to D \) are functions with \( B \subseteq C \).

a) If \( g \circ f \) is injective then \( f \) is injective.

b) If \( g \circ f \) is surjective then \( g \) is surjective.

c) If \( B = C \) and both \( f \) and \( g \) are bijective, then \( g \circ f \) is bijective.

**Problem 3.11** Prove the Proposition.

**Problem 3.12**

a) For part a) of the Proposition, give an example to show that \( g \) can fail to be injective.

b) For part b) of the Proposition, give an example to show that \( f \) can fail to be surjective.

c) In part c) of the Proposition, show that it is possible for \( f \) and \( g \) to fail to be bijections even if \( g \circ f \) is.

Suppose \( f : A \to B \) is a function. We think of \( f \) as “sending” each \( a \in A \) to a \( b = f(a) \in B \): \( a \to b \).

What happens if we try to reverse all these arrows: \( a \leftarrow b \); does that correspond to a function \( g : B \to A \)? The answer is “yes” precisely when \( f \) is a bijection, and the resulting function \( g \) is called its inverse.

**Definition.** Suppose \( f : A \to B \) is a function. A function \( g : B \to A \) is called an **inverse** function to \( f \) when \( a = g(f(a)) \) for all \( a \in A \) and \( b = f(g(b)) \) for all \( b \in B \). When such a \( g \) exists we say that \( f \) is **invertible** write \( g = f^{-1} \).

**Example 3.17.** The usual exponential function \( \exp : \mathbb{R} \to (0, \infty) \) is invertible. Its inverse is the natural logarithm: \( \exp^{-1}(x) = \ln(x) \).

**Example 3.18.** Consider \( X = (-2, \infty), \ Y = (-\infty, 1) \) and \( f : X \to Y \) defined by \( f(x) = \frac{x}{x^2 + 2} \). We claim that \( f \) is invertible and its inverse \( g : Y \to X \) is given by \( g(y) = \frac{2y}{1 - y} \). To verify this we need to examine both \( g \circ f \) and \( f \circ g \). First consider \( g \circ f \). For any \( x \in X \) we can write
\[
 f(x) = \frac{x}{x + 2} = 1 - \frac{2}{x + 2}.
\]

Since \( x + 2 > 0 \) it follows that \( 1 - \frac{2}{x + 2} < 1 \). Therefore \( f(x) \in Y \) for every \( x \in X \), and so \( g(f(x)) \) is defined. Now we can calculate that for every \( x \in X \),
\[
 g(f(x)) = \frac{2x}{1 - \frac{2}{x + 2}} = \frac{2x}{x + 2 - x} = \frac{2x}{2} = x.
\]

Similar considerations apply to \( f(g(y)) \). For any \( y \in Y, 1 - y > 0 \). Since \( 2y > 2y - 2 = -2(1 - y) \), we see that \( g(y) = \frac{2y}{1 - y} > -2 \), so that \( g(y) \in X \). For each \( y \in Y \) we can now check that
\[
 f(g(y)) = \frac{2y}{\frac{2y}{1 - y} + 2} = \frac{2y}{2y + 2(1 - y)} = \frac{y}{1} = y.
\]
Problem 3.13 Suppose \(a, b, c, d\) are positive real numbers. What is the largest subset \(X \subseteq \mathbb{R}\) on which \(f(x) = \frac{ax^3 + bx}{cx^3 + dx}\) is defined? Show that \(f : X \to \mathbb{R}\) is injective provided \(ad \neq bc\). What is its range \(Y\)? Find a formula for \(f^{-1} : Y \to X\).

Problem 3.14 Prove that if an inverse function exists, then it is unique.

Problem 3.15 Suppose that \(f : A \to B\) and \(g : B \to A\) such that \(g(f(x)) = x\) for all \(x \in A\).

a) Show by example that \(f\) need not be surjective and \(g\) need not be injective. Show that \(f(g(y)) = y\) for all \(y \in B\) fails for your example.

b) Show that the following are equivalent.

1. \(f\) is surjective.
2. \(g\) is injective.
3. \(f(g(y)) = y\) for all \(y \in B\).

Theorem 3.4. A function \(f : X \to Y\) is a bijection if and only if it has an inverse function.

Proof. Suppose \(f\) is a bijection. For any \(y \in Y\), since \(f\) is a surjection there exists \(x \in X\) with \(f(x) = y\). Since \(f\) is injective, this \(x\) is unique. Thus each \(y \in Y\) determines a unique \(x \in X\) for which \(f(x) = y\). We can therefore define a function \(g : Y \to X\) by

\[
g(y) = x\text{ for that }x \in X \text{ with } f(x) = y.
\]

We claim that \(g\) is an inverse function to \(f\). To see that, consider any \(x \in X\) and let \(y = f(x)\). Then \(y \in Y\) and by definition of \(g\) we have \(x = g(y)\). In other words \(x = g(f(x))\) for all \(x \in X\). Next consider any \(y \in Y\) and let \(x = g(y)\). Then \(x \in X\) and by definition of \(g\) we know \(f(x) = y\). Thus \(y = f(g(y))\) for all \(y \in Y\). We see that \(g\) is an inverse function to \(f\).

Now assume that there does exist a function \(g : Y \to X\) which is an inverse to \(f\). We need to show that \(f\) is a bijection. To see that it is surjective, consider any \(y \in Y\) and take \(x = g(y)\). Then \(f(x) = f(g(y)) = y\). Thus all of \(Y\) is in the range of \(f\). To see that \(f\) is injective, consider \(x, x' \in X\) with \(f(x) = f(x')\). Then \(x = g(f(x)) = g(f(x')) = x'\). Hence \(f\) is indeed injective.

Problem 3.16 Let \(f : (-1,1) \to \mathbb{R}\) be the function given by \(f(x) = \frac{x}{1-x^2}\). Prove that \(f\) is a bijection, and that its inverse \(g : \mathbb{R} \to (-1,1)\) is given by by

\[
g(y) = \begin{cases} 1 - \sqrt{1 + 4y^2} & \text{for } y \neq 0, \\ 0 & \text{for } y = 0. \end{cases}
\]

One way to do this is to show \(f\) is injective, surjective, and solve \(y = f(x)\) for \(x\) and see that you get the formula \(x = g(y)\). An alternate way is to verify that \(g(f(x)) = x\) for all \(x \in (-1,1)\), that \(f(g(y)) = y\) for all \(y \in \mathbb{R}\), and then appeal to Theorem 3.4.
Images and Preimages of Sets

We usually think of a function \( f : A \to B \) as sending elements \( a \) of \( A \) to elements \( b = f(a) \) of \( B \). But sometimes we want to talk about what happens to all the elements of a subset at once. We can think of \( f \) as sending a set \( E \subseteq A \) to the set \( f(E) \subseteq B \) defined by

\[
f(E) = \{ b \in B : \ b = f(e) \text{ for some } e \in E \}.
\]

We call \( f(E) \) the image of \( E \) under \( f \). We can do the same thing in the backward direction. If \( G \subseteq B \), the preimage of \( G \) under \( f \) is the set

\[
f^{-1}(G) = \{ a \in A : \ f(a) \in G \}.
\]

We do not need \( f \) to have an inverse function to be able to do this! In other words \( f^{-1}(G) \) is defined for sets \( G \subseteq B \) even if \( f^{-1}(b) \) is not defined for elements \( b \in B \). The meaning of “\( f^{-1}(\cdot) \)” depends on whether what is inside the parentheses is a subset or an element of \( B \).

**Example 3.19.** Consider \( f : \mathbb{R} \to \mathbb{R} \) defined by \( f(x) = x^2 \).

Then

- \( f((-2,3)) = [0,9) \). In other words the values \( 0 \leq y < 9 \) are the only real numbers which arise as \( y = f(x) \) for \(-2 < x < 3\).
- \( f^{-1}([1,2]) = [-\sqrt{2},-1] \cup [1, \sqrt{2}] \). In other words the values of \( x \) for which \( 1 \leq f(x) \leq 2 \) are those for which \( 1 \leq |x| \leq \sqrt{2} \).
- \( f^{-1}([-2,-1)) = \emptyset \). In other words there are no values of \( x \) for which \(-2 \leq f(x) \leq -1 \).

**Warning:** This function is neither injective nor surjective — \( f^{-1} \) does not exist as a function. However \( f^{-1} \) of sets is still defined. \( f^{-1}(2) \) is not defined, but \( f^{-1}(\{2\}) \) is \((-\sqrt{2}, \sqrt{2})\).

**Problem 3.17** Consider \( f : \mathbb{Z} \to \mathbb{Z} \) defined by \( f(n) = n^2 - 7 \). What is \( f^{-1}(\{ k \in \mathbb{Z} : \ k \leq 0 \}) \)?

**Problem 3.18** Suppose \( f : X \to Y \) and \( A, B \subseteq Y \). Prove that

- a) \( A \subseteq B \) implies \( f^{-1}(A) \subseteq f^{-1}(B) \).
- b) \( f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B) \).
- c) \( f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B) \).

**Problem 3.19** Suppose \( f : X \to Y \), \( A, B \subseteq Y \) and \( C, D \subseteq X \).

- a) Show that it is not necessarily true that \( f^{-1}(A) \subseteq f^{-1}(B) \) implies \( A \subseteq B \).
- b) Show that it is not necessarily true that \( f(C) \cap f(D) = f(C \cap D) \).
- c) If \( f \) is injective, show that it is true that \( f(C) \cap f(D) = f(C \cap D) \).
G  Cardinality of Sets

Now that we have summarized all the standard ideas about sets and functions, we are going to use them to discuss the basic idea of the size of a set. Most of us have no trouble with what it would mean to say that a set “has 12 elements” for instance, or even that a set “is finite.”

Example 3.20. Let $A = \{a, b, c, \ldots, z\}$ be our usual alphabet, and $B = \{1, 2, 3, \ldots, 26\}$. There is a simple bijection $f(a) = 1, f(b) = 2, f(c) = 3, \ldots$ in which $f$ assigns to each letter its position in the usual ordering of the alphabet. This $f$ provides a one-to-one pairing of letters with numbers $1 \leq n \leq 26$:

- a goes with 1
- b goes with 2
- c goes with 3
- \vdots
- z goes with 26.

Of course this works because both sets have exactly 26 elements.

In general when there is a bijection between two sets $A$ and $B$, that means that $A$ and $B$ have exactly the same number of elements. The bijection provides a way to pair the elements of $A$ with the elements of $B$, a “one-to-one correspondence” between the two sets. Here is the formal definition for this idea.

**Definition.** Two sets, $A$ and $B$, are called *equipotent* if there exists a bijection $f : A \rightarrow B$. We will write $A \simeq B$ to indicate that $A$ and $B$ are equipotent.

Two equipotent sets are often said to have *the same cardinality*. The idea of equipotence gives a precise meaning to “have the same number of elements” even when the number of elements is infinite. This agrees with our intuitive idea of “same size” for finite sets, as the example above illustrated. Things becomes more interesting for infinite sets.

Example 3.21. Let $A = (0, 1)$ and $B = (0, 2)$. The function $f : A \rightarrow B$ defined by $f(x) = 2x$ is a bijection, as is simple to check. So $f$ provides a one-to-one correspondence between $A$ and $B$, showing that these two intervals are equipotent.

The same reasoning applied to Problem 3.16 says that $(-1, 1)$ is equipotent to $\mathbb{R}$. In both these examples we have two sets $A$ and $B$ which are “of the same size” in terms of equipotence, even though in terms of geometrical size (length) one is larger than the other.

G.1  Finite Sets

What exactly do we mean by saying as set $A$ is a finite set? One way to express that idea is to say we can count its elements, “this is element #1, this is element #2, …” and get done at some point, “…this is element #n, and that accounts for all of them.” That counting process consists of picking a function $f : \{1, 2, 3, \ldots, n\} \rightarrow A$; $f(1)$ is the element of $A$ we called #1, $f(2)$ is the element we called #2, and so on. The fact that we accounted for all of them means that $f$ is surjective. The fact that we didn’t count the same element more than once means that $f$ is injective. We can turn our counting idea of finiteness into the following definition.

**Definition.** A set $A$ is said to be *finite* if it is either empty or equipotent to $\{1, 2, 3, \ldots, n\}$ for some $n \in \mathbb{N}$. A set which is not finite is called *infinite*.

Could it happen that a set is equipotent to $\{1, 2, 3, \ldots, n\}$ as well as to $\{1, 2, 3, \ldots, m\}$ for two different integers $n$ and $m$? We all know that this is *not* possible.

**Lemma 3.5** (Pigeon Hole Principle). Suppose $n, m \in \mathbb{N}$ and $f : \{1, 2, 3, \ldots, n\} \rightarrow \{1, 2, 3, \ldots, m\}$ is a function.

---

5There are many different words people use for this: equinumerous, equipollent, equinumerous, congruent, equivalent. Many use the symbol “$,\approx$,” but since that suggests approximately in some sense, we have chosen “$,\simeq$” instead.

6Although this name sounds like something out of Winnie the Pooh, it’s what everyone calls it.
a) If $f$ is injective then $n \leq m$.

b) If $f$ is surjective then $n \geq m$.

We could write a proof of this (using the Well-Ordering Principle of the integers from the next chapter), but it is tricky because what we are proving seems so obvious. Most of the effort would go toward sorting out exactly what we can and cannot assume about subsets of $\mathbb{Z}$. Instead we will take it for granted.

**Corollary 3.6.** If $f : \{1, 2, 3, \ldots, n\} \to \{1, 2, 3, \ldots, m\}$ is a bijection ($n, m \in \mathbb{N}$), then $n = m$.

**Problem 3.20** Prove that if there is an injection $f : A \to A$ which is not surjective, then $A$ is not a finite set.

G.2 Countable and Uncountable Sets

To say that “infinite” means “not finite” is simple enough. But now the big question: are all infinite sets equipotent to each other? This is where things get interesting: the answer is “no!” In fact there are infinitely many nonfinite cardinalities, as we will see. But a point to make first is that we are beyond our intuition here. We depend on proofs and counterexamples to know what is true and what is not. Few people can guess their way through this material.

**Example 3.22.** We will exhibit bijections for the following in class.

- $\{2k : k \in \mathbb{N}\} \sim \{2, 3, 4, \ldots\} \sim \mathbb{N} \sim \mathbb{Z}$.
- $\mathbb{R} \sim (0, 1) \sim (0, 1]$.

**Problem 3.21** Find an injective function $f : \mathbb{N} \to [0, 1)$ (the half-open unit interval).

The next example shows that it is possible for two infinite sets to fail to be equipotent.

**Example 3.23.** $\mathbb{N}$ and $[0, 1)$ are not equipotent! To see this, consider any function $f : \mathbb{N} \to [0, 1)$. We will show that $f$ is not surjective, by identifying a $y \in [0, 1)$ which is not in $\text{Ran}(f)$. We will do this by specifying the decimal representation $y = d_1d_2d_3\ldots$ where each digit $d_i \in \{0, 1, 2, \ldots, 9\}$. Consider the decimal representation of $f(1)$. We want to choose $d_1$ to be different than the first digit in the decimal representation of $f(1)$. For instance, if $f(1) = .352011\ldots$ we could choose $d_1 = 7$, since $7 \neq 3$. To be systematic about it, if $f(1) = .3\ldots$ take $d_1 = 7$, and otherwise take $d_1 = 3$. If $f(2) = .3\ldots$ take $d_2 = 7$, and otherwise take $d_2 = 3$. If $f(k) = .3\ldots$ (a 3 in the $k$th position) take $d_k = 7$, and $d_3 = 3$ otherwise. This identifies a specific sequence of digits $d_i$ which we use to form the decimal representation of $y$. Then $y \in [0, 1)$, and for every $k$ we know $y \neq f(k)$ because their decimal expansions differ in the $k$th position. Thus $y$ is not in the range of $f$. Hence $f$ is not surjective.

**Definition.** A set $A$ is called **countably infinite** when $A \sim \mathbb{N}$. We say $A$ is **countable** when $A$ is either finite or countably infinite. A set which is not countable is called **uncountable**.

**Theorem 3.7** (Cantor’s Theorem). If $A$ is a set, there is no surjection $F : A \to \mathcal{P}(A)$.

(Note that we are using an upper case “$F$” for this function to help us remember that $F(a)$ is a subset, not an element, of $A$.)

It’s easy to write down an injection: $F(a) = \{a\}$. Essentially Cantor’s Theorem says that $\mathcal{P}(A)$ always has greater cardinality than $A$ itself. So the sets $\mathbb{N}$, $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, \ldots will be an unending list of sets, each with greater cardinality than the one before it! This is why we said above there are infinitely many nonfinite cardinalities. Here is the proof of Cantor’s Theorem.

69
Proof. Suppose \( F : A \to \mathcal{P}(A) \) is a function. In other words, for each \( a \in A \), \( F(a) \) is a subset of \( A \). Consider

\[
C = \{ a \in A : a \notin F(a) \}.
\]

Clearly \( C \subseteq A \), so \( C \in \mathcal{P}(A) \). We claim that there is no \( b \in A \) for which \( F(b) = C \). If such a \( b \) existed then either \( b \in C \) or \( b \notin C \). But \( b \in C \) means that \( b \notin F(b) = C \), which is contradictory. And \( b \notin C \) would mean that \( b \in F(b) = C \), again a contradiction. Thus \( C = F(b) \) leads to a contradiction either way. Hence no such \( b \) can exist.

\[ \square \]

**Problem 3.22** Let \( F : \mathbb{R} \to \mathcal{P}(\mathbb{R}) \) be defined by

\[
F(x) = \begin{cases} 
\emptyset & \text{if } -1 \leq x \leq 0, \\
(-x, x^2) & \text{otherwise.}
\end{cases}
\]

Find the set \( C \) described in the proof of Theorem 3.7.

---

### G.3 The Schroeder-Bernstein Theorem

The definition of equipotence gives a precise meaning for saying two sets have the same size. It is pretty easy to express what we mean by a set \( A \) being at least as large as \( B \), the existence of an injection \( f : B \to A \). This implies that \( B \) is equipotent to a subset of \( A \), which seems to agree with our idea of \( A \) being at least as large as \( B \). Now, if \( A \) is at least as large as \( B \) and \( B \) is at least as large as \( A \), we might naturally expect that \( A \) and \( B \) must be equipotent. That seems natural, but is it true? Yes, it is – this is the Schroeder-Bernstein Theorem. But to prove it is not a simple task.

**Theorem 3.8** (Schroeder-Bernstein Theorem). Suppose \( X \) and \( Y \) are sets and there exist functions \( f : X \to Y \) and \( g : Y \to X \) which are both injective. Then \( X \) and \( Y \) are equipotent.

**Example 3.24.** To illustrate how useful this theorem can be, let's use it to show that \( \mathbb{Z} \simeq \mathbb{Q} \). It's easy to exhibit an injection \( f : \mathbb{Z} \to \mathbb{Q} \); just use \( f(n) = n \). It's also not too hard to find an injection \( g : \mathbb{Q} \to \mathbb{Z} \). Given \( q \in \mathbb{Q} \), start by writing it as \( q = \pm \frac{n}{m} \), where \( n, m \) are nonnegative integers with no common factors. Using this representation, define \( g(q) = \pm 2^m 3^m \). It is clear that this is also an injection. Theorem 3.8 now implies that \( \mathbb{Z} \simeq \mathbb{Q} \). By Example 3.22, it follows that \( \mathbb{Q} \) is countable!

**Problem 3.23** Use Theorem 3.8 to show \( [0, 1] \simeq (0, 1) \).

---

**Proof.** Suppose \( f : X \to Y \) and \( g : Y \to X \) are both injective. Let \( Y_0 = f(X) \), the range of \( f \). Then \( f \) is a bijection from \( X \) to \( Y_0 \), so \( X \simeq Y_0 \). Although \( Y_0 \) is only a subset of \( Y \) we will show that there is a bijection \( Y_0 \simeq Y \).

Let \( B_0 = Y \setminus Y_0 \), the part of \( Y \) that \( f \) misses. Then recursively define

\[
B_1 = f(g(B_0)), \ B_2 = f(g(B_1)), \ldots, B_{n+1} = f(g(B_n)), \ldots
\]

For \( n \geq 1 \) note that \( B_n \) is a subset of \( Y_0 \) and therefore disjoint from \( B_0 \). Define \( h : Y \to Y_0 \) by

\[
h(y) = \begin{cases} 
f(g(y)) & \text{if } y \in \bigcup_{n=0}^{\infty} B_n \\
y & \text{otherwise.}
\end{cases}
\]

To see that \( h \) is surjective consider any \( y_0 \in Y_0 \). If \( y_0 \) does not belong to \( \bigcup_{n=0}^{\infty} B_n \) then \( y_0 = h(y_0) \). Suppose \( y_0 \) does belong to \( \bigcup_{n=0}^{\infty} B_n \). Since \( y_0 \in Y_0 \) which is disjoint from \( B_0 \) it must be that \( y_0 \in B_n \) for some \( n \geq 1 \). By definition of \( B_n \) this means \( y_0 = f(g(y)) \) for some \( y \in B_{n-1} \). Thus \( y_0 = h(y) \) for some \( y \in Y \).
To see that $h$ is injective suppose that $h(y) = h(y')$ for some $y, y' \in Y$. Observe that $h(\cup_0^\infty B_n) = \cup_1^\infty B_n$. Suppose $y \notin \cup_0^\infty B_n$ then $h(y') = h(y) = y$, and therefore $y' \notin \cup_0^\infty B_n$. That implies that $y' = h(y') = h(y) = y$. Next suppose $y \in \cup_0^\infty B_n$. Then it must be that $y' \in \cup_0^\infty B_n$ as well, else $y' = h(y') = h(y) \in \cup_0^\infty B_n$ would be a contradiction. But then $f(g(y)) = h(y) = h(y') = f(g(y'))$. Since $f \circ g$ is injective it follows that $y = y'$.

Having shown that $h$ is a bijection, we know it has an inverse $h^{-1} : Y_0 \to Y$. The composition $h^{-1} \circ f$ is a bijection from $X$ to $Y$, proving the theorem.

H Perspective: The Strange World at the Foundations of Mathematics

We close this chapter with a brief look at a couple of issues from the foundations of mathematics.

H.1 The Continuum Hypothesis

The collection of possible cardinalities of sets form what are called the \textit{cardinal numbers}. H.1

We have mentioned that a logical swamp awaits those who wander too far along the path of forming sets of sets of sets of... This is a real mind-bender — take a deep breath, and read on.

Let $S$ be the set of \textit{all} sets. It contains sets, and sets of sets, and sets of sets of sets and so on. If $A$ and $B$ are sets then $A, B \in S$, and it is possible that $A \in B$, e.g. $B = \{A, \ldots\}$. In particular we can ask whether $A \in A$. Consider then the set

$$\mathcal{R} = \{A \in S : A \notin A\}.$$

Now we ask the “killer” question: is $\mathcal{R} \in \mathcal{R}$? If the answer is “yes,” then the definition of $\mathcal{R}$ says that $\mathcal{R} \notin \mathcal{R}$, meaning the answer is “no.” And if the answer is “no,” the definition of $\mathcal{R}$ says that $\mathcal{R} \in \mathcal{R}$, meaning the answer is “yes.” Either answer to our question is self-contradictory.

What is going on here? We seem to have tied ourselves in a logical knot. This conundrum is called \textit{Russell’s Paradox}. A paradox is not quite the same as a logical contradiction or impossibility. Rather it is an \textit{apparent} contradiction, which typically indicates something wrong or inappropriate about our reasoning. Sometimes a paradox is based on a subtle/clever misuse of words. Here we are allowing ourselves to mingle the ideas of set and element too freely, opening the door to the above paradoxical discussion that Bertrand Russell brewed up. Historically, Russell’s Paradox showed the mathematical world that they had not yet fully worked out what set theory actually consisted of, and sent them back to the task of trying to decide which statements about sets are legitimate and which are not. This led to the development of axioms which govern formal set theory (the Zermelo-Fraenkel axioms) which prevent the misuse of language that Russell’s paradox illustrates. (In the next chapter we will see what we mean by “axioms” in the context of the integers.) This is a difficult
topic in mathematical logic which we cannot pursue further here. Our point here is only that careless or sloppy use of words can tie us in logical knots, and a full description of what constitutes appropriate usage is a major task.

These things do not mean that there is something wrong with set theory, and that we are just choosing to go ahead in denial. Rather they mean that we have to be careful not to be too cavalier with the language of sets. For virtually all purposes, if we limit ourselves to two or three levels of elements, sets, sets of sets, and sets of sets of sets (but stop at some point) we will be able to sleep safely, without fear of these dragons from the logical abyss. See Chapter 4, Section E for a bit more on these topics.