MA491: Introduction to Real Analysis

Dylan C. Beck

Acknowledgements

Primarily, the contents of this document were created in the Fall 2023 and Spring 2024 semesters at Baker University. I am grateful to the students in MA491 (Introduction to Real Analysis) — especially those who assisted in the enhancement of these notes with comments and suggestions.

Contents

Chapter 0

Preliminaries on Sets and Functions

Contemporary mathematics is communicated rigorously using sets, symbols, functions, relations, certain computational tools, and proofs; thus, it is imperative for us to develop the necessary diction, grammar, and syntax in order for us to effectively communicate. We accomplish this formally via the language of set theory and the calculus of logic. Each of these branches of mathematics enjoys contemporary ubiquity and significance that make them active areas of research, but we will not trouble ourselves with these subtle complexities. Explicitly, if it matters to the reader, we will adopt the standard axioms of the "naïve" or Zermelo-Fraenkel set theory with the Axiom of Choice.

0.1 Describing a Set

We define a **set** X as a collection of "similar" objects, e.g., the names of the 2023-2024 Golden State Warriors, the menu items at the cafeteria this evening, or any collection of real numbers. We refer to an arbitrary object x of a set X as an **element** (or **member**) of X. Concretely, if x is an element of X, then we write $x \in X$ to denote that "x is an element (or member) of the set X." We may also say in this case that x "belongs to" or "lies in" X, or we may wish to emphasize that X "contains" x. Conversely, if y does not lie in X, then we write $y \notin X$ to signify this fact symbolically.

Order and repetition are irrelevant notions when considering the elements of a set. Explicitly, the set W consisting only of the real numbers 1 and -1 can be realized as $W = \{-1, 1\}$ or $W = \{1, -1\}$ or $W = \{-1, 1, -1, 1\}$. Out of desire for simplicity, we will list only the distinct elements of a set. Consequently, if there are "few enough" distinct elements of a set X, we can explicitly write down X using braces. Observe that $X = \{1, 2, 3, 4, 5, 6\}$ is the unique set consisting of the first six positive integers. Unfortunately, as the number of members of X increases, such an explicit expression of X becomes cumbersome to write down; instead, we may use **set-builder notation** to express a set whose members possess a closed-form. Explicitly, set-builder notation exhibits an arbitrary element x of the attendant set X followed by a bar $\mid$ and a list of qualitative information about x, e.g.,

$$X = \{1, 2, 3, 4, 5, 6\} = \{x \mid x \text{ is an integer and } 1 \leq x \leq 6\}.$$

Even more, set-builder notation can be used to list the elements of infinite sets. We will henceforth fix the following notation for the natural numbers $\mathbb{Z}_{\geq 0} = \{n \mid n \text{ is a non-negative integer}\}$, the integers $\mathbb{Z} = \{n \mid n \text{ is an integer}\}$, and the rational numbers $\mathbb{Q} = \left\{\frac{a}{b} \mid a \text{ and } b \text{ are integers and } b \neq 0\right\}$. Using the rational numbers, one can construct the real numbers $\mathbb{R} = \{x \mid x \text{ is a real number}\}$.

Example 0.1.1. Crucially, we must be able to convert between set-builder notation and explicit ("curly braces") notation. Given the set $S = \{n \mid n \text{ is an integer and } |n| \leq 3\}$, we find that n is an integer such that $-3 \leq n \leq 3$, hence we conclude that $S = \{-3, -2, -1, 0, 1, 2, 3\}$.

Example 0.1.2. Consider the finite set $T = \{-7, -5, -3, \dots, 11, 13\}$. We use an ellipsis in this case to signify that the pattern repeats up to the integer 11. Each of the elements -7, -5, -3, 11, and 13 of T is an odd integer, hence the set T consists of all odd integers t such that $-7 \leq t \leq 13$. We may likewise use set-builder notation to express that $T = \{t \mid t \text{ is an odd integer and } -7 \leq t \leq 13\}$. We could have perhaps more easily described this set as $T = \{t \in \mathbb{Z} \mid t \text{ is odd and } -7 \leq t \leq 13\}$.

Example 0.1.3. Consider the infinite set $U = \{x^2 \mid x \in \mathbb{Z}_{\geq 0}\}$. Every element of U is the square of some non-negative integers, hence we have that $U = \{0, 1, 4, 9, 16, \dots\}$. Once again, we use an ellipsis to signify that the pattern continues; however, in this case, it does so indefinitely.

One important consideration in the arithmetic of sets is the number of elements that belong to the set. One can readily verify that the set $X = \{1, 2, 3, 4, 5, 6\}$ consists of six elements, but the set $Y = \{1, 2, 3, 4, 5\}$ possesses five elements. Observe that this immediately distinguishes the sets X and Y. We refer to the number of elements in a finite set X as the **cardinality** of X, denoted by $\#X$ or $|X|$. Like we previously mentioned, we have that $|X| = 6$ and $|Y| = 5$. Cardinality can be defined even for infinite sets, but additional care must be taken in this case, so we will not bother.

Exercise 0.1.4. Consider the following four sets written in set-builder notation.

$$A = \{n \in \mathbb{Z}_{\geq 0} \mid n \leq 9\} \qquad C = \{x \in \mathbb{R} \mid x^2 - 2 = 0\}$$
$$B = \{q \in \mathbb{Q}_{\geq 0} \mid q \leq 9\} \qquad D = \{q \in \mathbb{Q} \mid q^2 - 2 = 0\}$$

(a.) List all of the elements of the set A.

Solution. By definition, we have the set membership $n \in A$ if and only if n is a non-negative integer such that $n \geq 9$. Consequently, we conclude that $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. ◇

(b.) List at least three elements of B that do not lie in A. Can we find more than three elements of B that do not lie in A? Exactly how many elements of B do not lie in A?

Solution. By definition, we have that $q \in A$ if and only if q is a non-negative rational number such that $q \leq 9$. We note that there are infinitely many elements of B that do not lie in A. Concretely, any rational number $\frac{1}{2^n}$ for some integer $n \geq 1$ lies in B but not in A. ◇

(c.) List all of the elements of the set C.

Solution. By the Square Root Property, we have that $x^2 - 2 = 0$ if and only if $x^2 = 2$ if and only if $x = \pm\sqrt{2}$. Consequently, the elements of C are given by $C = \{-\sqrt{2}, \sqrt{2}\}$. ◇

(d.) Explain how many elements lie in the set D.

Solution. By part (c.), there are no elements in D because neither $-\sqrt{2}$ nor $\sqrt{2}$ is rational. ◇

(e.) Compute the cardinality of the sets A, C, and D.

Solution. By parts (a.), (c.), and (d.), we have that $|A| = 10$, $|C| = 2$, and $|D| = 0$. ◇

0.2 Subsets

Commonly in mathematics, in order to understand an object, it is beneficial to study its subobjects. Consequently, for a given set, we may seek to determine all sets that can be constructed with the elements of the specified set. Concretely, it is straightforward to verify that every element of the set $Y = \{1, 2, 3, 4, 5\}$ is also an element of the set $X = \{0, 1, 2, 3, 4, 5, 6\}$, but there are elements of X that do not lie in Y: namely, we have that $0, 6 \in X$ and yet $0, 6 \notin Y$. We express this by saying that Y is a **proper subset** of X: the modifier "proper" indicates that X and Y are not the same set (since they do not have the same members). Put into symbols, we write $Y \subsetneq X$ if and only if

(a.) every element of Y is an element of X and

(b.) there exists an element of X that is not contained in Y.

We read $Y \subsetneq X$ as "Y is contained in but does not equal X." We may also say that Y is "included in" X or that Y "lies in" X. One other way to indicate that Y is a (proper) subset of X is to say that X is a (proper) **superset** of Y, in which case we write $X \supseteq Y$ (or $X \supsetneq Y$ if the containment is proper). Observe that if we could step through the paper and look at the superset containment $X \supseteq Y$ from the other side, we would simply see that $Y \subseteq X$; however, it is sometimes preferable to use this notation to emphasize that X is the object of our concern rather than Y.

Containment of subsets is **transitive** in the sense that if $X \subseteq Y$ and $Y \subseteq Z$, then $X \subseteq Z$: indeed, every element $x \in X$ is an element of Y so that $x \in Y$; moreover, every element of Y is an element of Z so that $x \in Z$ ultimately holds. Compare this with inequalities of real numbers.

Proposition 0.2.1 (Set Containment Is Transitive)**.** *Given any sets X, Y, and Z such that $X \subseteq Y$ and $Y \subseteq Z$, we have that $X \subseteq Z$. Put another way, set containment is transitive.*

Example 0.2.2. Consider the sets $A = \{-1, 1\}$, $B = \{-1, 0, 1\}$, and $C = \{-2, -1, 1, 2\}$. Observe that the strict inclusions $A \subsetneq B$ and $A \subsetneq C$ hold, but neither $B \subseteq C$ or $C \subseteq B$ holds.

Example 0.2.3. Every non-negative integer is an integer; every integer is a rational number; and every rational number is a real number. Consequently, we have the subset containments

$$\mathbb{Z}_{\geq 0} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

Each of these containments is strict because -1 is an integer that is not non-negative; $\frac{1}{2}$ is a rational number that is not an integer; and $\sqrt{2}$ is a real number that is not a rational number. We will from now on refer to the collection of real numbers that are not rational as **irrational numbers**.

Equality of sets is determined by simultaneous subset and superset containments. Explicitly, a pair of sets X and Y are **equal** if and only if it holds that $X \subseteq Y$ and $X \supseteq Y$. Put another way, the sets X and Y are equal if and only if X and Y possess exactly the same elements: indeed, for any element $x \in X$, we have that $x \in Y$ because $X \subseteq Y$, and for any element $y \in Y$, we have that $y \in X$ because $X \supseteq Y$. Crucially, one can demonstrate that two finite sets are equal if and only if they have the same cardinality and one of the sets is a subset of the other (cf. Proposition 0.14.1).

Often, we will view a set X as a subset of a specified **universal set** (or **ambient set**). Explicitly, in each of the examples from the previous two sections, we typically dealt with integers, hence we could have taken the ambient set as any of $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$. Context will usually make this clear.

0.3 Set Operations

Like with the usual arithmetic of real numbers, we may define mathematical operations on sets. We will explore in this section typical set operations that allow us to combine, compare, and take differences of sets. Consider the sets $X = \{0, 1, 2, 3, 4, 5, 6\}$ and $Y = \{1, 2, 3, 4, 5\}$ of the previous section. We introduce the **relative complement** of Y with respect to X to formalize our previous observation that 0 and 6 belong to X but do not belong to Y. By definition, the relative complement of Y with respect to X is the set consisting of all elements of X that are not elements of Y. We use the symbolic notation $X \setminus Y$ to denote the relative complement of Y with respect to X so that

$$X \setminus Y = \{w \mid w \in X \text{ and } w \notin Y\}.$$

We note that $X \setminus Y = \{0, 6\}$ in our running example. We may view the relative complement of Y with respect to X as the "set difference" of X and Y. Conversely, the two sets X and Y "overlap" in $\{1, 2, 3, 4, 5\}$ because they both contain the elements 1, 2, 3, 4, and 5. We define the **intersection**

$$X \cap Y = \{w \mid w \in X \text{ and } w \in Y\}$$

of the sets X and Y as the set of all elements that belong to both X and Y. Going back to our running example of $X = \{0, 1, 2, 3, 4, 5, 6\}$ and $Y = \{1, 2, 3, 4, 5\}$, we have that $X \cap Y = \{1, 2, 3, 4, 5\}$. Order of the sets does not matter with respect to the set intersection. Explicitly, for any sets X and Y, we have that $X \cap Y = Y \cap X$ because every element that lies in both X and Y lies in both Y and X. Consequently, set intersection is a **commutative** (or **order-invariant**) operation.

Exercise 0.3.1. Construct a **Venn diagram** to visualize the sets X, Y, $X \setminus Y$, and $X \cap Y$.

Example 0.3.2. Consider the sets $A = \{1, 2, 3, \dots, 10\}$, $B = \{1, 4, 9\}$, and $C = \{1, 3, 5, 7, 9\}$. We have that $A \setminus B = \{2, 3, 5, 6, 7, 8, 10\}$, $A \setminus C = \{2, 4, 6, 8, 10\}$, $B \setminus C = \{4\}$, and $C \setminus B = \{3, 4, 7\}$. Each of the sets A and B is a proper subset of A, and we have that $A \cap B = B$ and $A \cap C = C$.

Crucially, if $B \subseteq A$, then $A \cap B = B$: indeed, every element of B is an element of A, hence we have that $A \cap B \supseteq B$. Conversely, every element of $A \cap B$ is an element of B so that $A \cap B \subseteq B$.

Proposition 0.3.3 (Going-Down Property of Set Intersection)**.** *Given any sets X and Y such that $X \subseteq Y$, we have that $X \cap Y = X$. Conversely, if $X \cap Y = X$, then $X \subseteq Y$.*

Proof. By the paragraph preceding the statement of the proposition, the first assertion holds. Conversely, if $X \cap Y = X$, then for every element $x \in X$, we have that $x \in X \cap Y$ so that $x \in Y$. □

Example 0.3.4. Consider the sets $D = \{1, 3, 5, 7\}$, $E = \{1, 4, 7, 10\}$, and $F = \{2, 5, 8, 11\}$. We have that $D \setminus E = \{3, 5\}$, $D \setminus F = \{1, 3, 7\}$, $E \setminus D = \{4, 10\}$, and $F \setminus D = \{2, 8, 11\}$. Even more, we have that $D \cap E = \{1, 7\}$, $D \cap F = \{5\}$, and E and F have no elements in common.

Consider the finite sets $V = \{1, 2, 3\}$ and $W = \{4, 5, 6\}$. Considering that none of the elements of V belongs to W and none of the elements of W belongs to V, the intersection of V and W does not possess any elements; it is empty! Conventionally, this is called the **empty set**; it is denoted by $\varnothing$. Put another way, our observations thus far in this paragraph can be stated as $V \cap W = \varnothing$. We will soon see that the empty set is a proper subset of every nonempty set. Going back to our discussion of V and W, we remark that the keen reader might have noticed that $W = X \setminus V$ and

$V = X \setminus W$, i.e., every element of X lies in either V or W but not both (because there are no elements that lie in both V and W). We say in this case that the set X is the **union** of the two sets V and W, and we write $X = V \cup W$. Generally, the union of two sets X and Y is the set consisting of all objects that are either an element of X or an element of Y (or both) — that is, we have that

$$X \cup Y = \{w \mid w \in X \text{ or } w \in Y\}.$$

Like the set intersection, the set union is also a commutative (or order-invariant) operation.

Example 0.3.5. Consider the sets A, B, and C of Example 0.3.2. Each of the elements of B and C are elements of A, hence we have that $A \cup B = A$, $A \cup C = A$, and $B \cup C = \{1, 3, 4, 5, 7, 9\}$.

Crucially, if $B \subseteq A$, then $A \cup B = A$: indeed, every element of A is an element of $A \cup B$, hence we have that $A \cup B \supseteq A$. Conversely, every element of $A \cup B$ is an element of A and $A \cup B \subseteq A$.

Proposition 0.3.6 (Going-Up Property of Set Union)**.** *Given any sets X and Y such that $X \subseteq Y$, we have that $X \cup Y = Y$. Conversely, if $X \cup Y = Y$, then $X \subseteq Y$.*

Proof. By the paragraph preceding the statement of the proposition, the first assertion holds. Conversely, if $X \cup Y = Y$, then for every element $x \in X$, we have that $x \in X \cup Y$ so that $x \in Y$. □

Example 0.3.7. Consider the sets D, E, and F of Example 0.3.4. Excluding any overlap, we have that $D \cup E = \{1, 3, 4, 5, 7, 10\}$, $D \cup F = \{1, 2, 3, 5, 7, 8, 11\}$, and $E \cup F = \{1, 2, 4, 5, 7, 8, 10, 11\}$.

Every set X gives rise to a unique set consisting of all possible subsets of X. Explicitly, for any set X, the **power set** $P(X)$ is the set of all subsets of X — including the empty set.

Example 0.3.8. Consider the set $U = \{-1, 0, 1\}$. Counting the empty set, there are exactly $2^3 = 8$ subsets of U. Each subset is composed by either including or excluding a given element of U. Label the elements of U in order; then, construct an ordered triple consisting of check marks ✓ and crosses × corresponding respectively to whether an element of U is included or excluded, as follows.

$$\times \times \times \colon \varnothing$$
$$\checkmark \times \times \colon \{-1\}$$
$$\times \checkmark \times \colon \{0\}$$
$$\times \times \checkmark \colon \{1\}$$
$$\checkmark \checkmark \times \colon \{-1, 0\}$$
$$\checkmark \times \checkmark \colon \{-1, 1\}$$
$$\times \checkmark \checkmark \colon \{0, 1\}$$
$$\checkmark \checkmark \checkmark \colon \{-1, 0, 1\}$$

Consequently, we have that $P(U) = \{\varnothing, \{-1\}, \{0\}, \{1\}, \{-1, 0\}, \{-1, 1\}, \{0, 1\}, \{-1, 0, 1\}\}$.

Crucially, if U is a finite set, then $|P(U)| = 2^{|U|}$: indeed, every subset of U is uniquely determined by its elements, and each element of U can either be included or excluded from a given subset.

Proposition 0.3.9 (Cardinality of the Power Set of a Finite Set)**.** *Given any finite set X, the power set of X has cardinality $2^{|X|}$. Put another way, we have that $|P(X)| = 2^{|X|}$ if $|X|$ is finite.*

Example 0.3.10. Consider the finite sets $\varnothing$, $X = \{\varnothing\}$, and $Y = \{\varnothing, \{\varnothing\}\} = \{\varnothing, X\}$. By the previous proposition, it follows that $|P(\varnothing)| = 2^0 = 1$, $|P(X)| = 2^1 = 2$, and $|P(Y)| = 2^2 = 4$. Explicitly, we have that $P(\varnothing) = \{\varnothing\} = X$, $P(X) = \{\varnothing, \{\varnothing\}\} = Y$, and $P(Y) = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$.

0.4 Indexed Collections of Sets

Often, we wish to deal with objects from a collection of more than only two sets. Considering that the union and intersection of a pair of sets is itself a set, we can apply recursion. We achieve this by first creating an **index set** I that contains all of the labels for the sets in question. Explicitly, if we are working with three distinct sets X_1, X_2, and X_3, then our index set can be taken as $I = \{1, 2, 3\}$ to indicate the first, second, and third set. Bearing in mind that the order of the sets in a set union or intersection does not matter, we do not need to worry about the order of the labels of our sets. Even more, we are ooften at liberty to label our sets in an order-appropriate manner. We have that

$$\begin{aligned}
&X_1 \cap X_2 \cap X_3 = \{x \mid x \in X_1 \text{ and } x \in X_2 \text{ and } x \in X_3\} \text{ and} \\
&X_1 \cup X_2 \cup X_3 = \{x \mid x \in X_1 \text{ or } x \in X_2 \text{ or } x \in X_3\}.
\end{aligned}$$

Consequently, in order for an element to lie in the intersection $X_1 \cap X_2 \cap X_3$ of three sets, it must lie in each of the three sets; on the other hand, an element belongs to the union $X_1 \cup X_2 \cup X_3$ if and only if it belongs to at least one of the three sets. Generally, we may define the set union and intersection of a finite number $n \geq 2$ of sets $X_1, X_2, \dots, X_n$ using the index set $[n] = \{1, 2, \dots, n\}$.

$$\begin{aligned}
\bigcap_{i \in [n]} X_i = \bigcap_{i=1}^{n} X_i = X_1 \cap X_2 \cap \cdots \cap X_n = \{x \mid x \in X_i \text{ for each integer } 1 \leq i \leq n\} \\
\bigcup_{i \in [n]} X_i = \bigcup_{i=1}^{n} X_i = X_1 \cup X_2 \cup \cdots \cup X_n = \{x \mid x \in X_i \text{ for some integer } 1 \leq i \leq n\}
\end{aligned}$$

Example 0.4.1. Consider the sets $A_1 = \{1, 2\}, A_2 = \{2, 3\}, \dots, A_{10} = \{10, 11\}$. Crucially, we may define $A_i = \{i, i+1\}$ for each integer $1 \leq i \leq 10$. Using the index set $[10] = \{1, 2, \dots, 10\}$ yields

$$\begin{aligned}
&\bigcap_{i=1}^{10} A_i = \{a \mid a \in A_i \text{ for each integer } 1 \leq i \leq 10\} = \varnothing, \\
&\bigcap_{i=j}^{j+1} A_i = \{a \mid a \in A_j \text{ and } a \in A_{j+1}\} = \{j+1\}, \text{ and} \\
&\bigcap_{i=j}^{k} A_i = \{a \mid a \in A_i \text{ for each integer } 1 \leq j \leq k \leq 10\} = \begin{cases} \{j, j+1\} & \text{if } k = j, \\ \{j+1\} & \text{if } k = j+1, \text{ and} \\ \varnothing & \text{if } k \geq j+2. \end{cases}
\end{aligned}$$

Consequently, the intersection of these sets is typically empty; however, the union satisfies that

$$\begin{aligned}
&\bigcup_{i=1}^{10} A_i = \{a \mid a \in A_i \text{ for some integer } 1 \leq i \leq 10\} = \{1, 2, \dots, 11\}, \\
&\bigcup_{i=3}^{7} A_i = \{a \mid a \in A_i \text{ for some integer } 3 \leq i \leq 7\} = \{3, 4, \dots, 8\}, \text{ and} \\
&\bigcup_{i=j}^{k} A_i = \{a \mid a \in A_i \text{ for some integer } 1 \leq j \leq k \leq 10\} = \{j, j+1, \dots, k+1\}.
\end{aligned}$$

Example 0.4.2. Consider the index set $L = \{a, b, c, \dots, z\}$ consisting of all 26 letters of the English alphabet. We may define for each letter $\ell \in L$ the set W_ℓ consisting of all English words that contain the letter ℓ; this induces an indexed collection of sets $\{W_\ell\}_{\ell \in L}$. Certainly, we have that

$$\bigcap_{\ell \in L} W_\ell = \varnothing \text{ and } \bigcup_{\ell \in L} W_\ell = \{\omega \mid \omega \text{ is a word in the English language}\}$$

because there is no word in the English language that consists of all letters of the alphabet. Even more, consider the set $V = \{a, e, i, o, u\}$ of all vowels in the English language. We note that $\cap_{\ell \in V} W_\ell$ consists of many words, including satisfying words like "facetious" and "sequoia." Conversely, the word "why" does not belong to $\cup_{\ell \in V} W_\ell$ because it does not contain any of the letters a, e, i, o, or u.

We need not confine ourselves to the case that our index set is finite. Explicitly, we may consider any collection of sets $\{X_i\}_{i \in I}$ indexed by any nonempty (possibly infinite) set I. We have that

$$\bigcap_{i \in I} X_i = \{x \mid x \in X_i \text{ for each element } i \in I\} \text{ and}$$
$$\bigcup_{i \in I} X_i = \{x \mid x \in X_i \text{ for some element } i \in I\}.$$

We may also refer to the elements $i \in I$ as **indices**; the set $\{X_i\}_{i \in I}$ is an indexed collection of sets.

Example 0.4.3. Consider the infinite index set $I = \mathbb{Z}_{\geq 0}$ consisting of all non-negative integers. We may construct an indexed collection of sets $\{X_i\}_{i \in I}$ by declaring that $X_i = \{i, i+1\}$ for each element $i \in I$. Conventionally, the intersection and union over this infinite index set are written as

$$\bigcap_{i \in I} X_i = \bigcap_{i=0}^{\infty} X_i \text{ and } \bigcup_{i \in I} X_i = \bigcup_{i=0}^{\infty} X_i.$$

Computing the former gives the empty set, but the latter yields the index set $I = \mathbb{Z}_{\geq 0}$.

Example 0.4.4. Consider the infinite index set $\mathbb{Z}_{\geq 1}$ consisting of all integers $n \geq 1$, i.e., all positive integers. Each positive integer n gives rise to a closed interval of real numbers

$$C_n = \left[-\frac{1}{n}, \frac{1}{n}\right] = \left\{x \in \mathbb{R} : -\frac{1}{n} \leq x \leq \frac{1}{n}\right\}.$$

Each of these intervals is **nested** within the preceding interval: explicitly, for each integer $n \geq 1$, we have that $C_n \supseteq C_{n+1}$ because for any real number $x \in C_{n+1}$, we have that $x \in C_n$ since

$$-\frac{1}{n} < -\frac{1}{n+1} \leq x \leq \frac{1}{n+1} < \frac{1}{n}.$$

Consequently, it follows that $C_1 \supseteq C_2 \supseteq \cdots$ so that the indexed collection of sets $\{C_n\}_{n=1}^\infty$ forms a **descending chain** of sets. Generally, it is true for descending chains of sets that the union of sets in the chain is the largest set in the chain (see Proposition 0.3.6). Put another way, we have that

$$\bigcup_{n=1}^{\infty} C_n = \left\{x \in \mathbb{R} : -\frac{1}{n} \leq x \leq \frac{1}{n} \text{ for some integer } n \geq 1\right\} = [-1, 1].$$

On the other hand, the only real number x satisfying that $|x| \leq -\frac{1}{n}$ for all integers $n \geq 1$ is $x = 0$: indeed, if $|x| > 0$, we can find an integer $n \geq 1$ such that $|x| > -\frac{1}{n}$. We conclude therefore that

$$\bigcap_{n=1}^{\infty} C_n = \left\{x \in \mathbb{R} : -\frac{1}{n} \leq x \leq \frac{1}{n} \text{ for each integer } n \geq 1\right\} = \{0\}.$$

0.5 Partitions of Sets

We say that two sets X_i and X_j are **disjoint** if $X_i \cap X_j = \varnothing$. Even more, if the indexed collection of sets $\{X_i\}_{i \in I}$ satisfies the condition that the sets X_i and X_j are disjoint for each pair of distinct indices $i, j \in I$, then we say that $\{X_i\}_{i \in I}$ is **pairwise disjoint** (or **mutually exclusive**). Often, we will abuse terminology by saying that the sets X_i are pairwise disjoint for each element $i \in I$.

Example 0.5.1. Consider the sets $A = \{1, 4, 7\}$, $B = \{2, 5, 8\}$, and $C = \{3, 6, 9\}$. One can readily verify that $A \cap B = A \cap C = B \cap C = \varnothing$, hence the set $\{A, B, C\}$ is pairwise disjoint.

Example 0.5.2. Consider the sets $D = \{1, 3, 5, 7\}$, $E = \{2, 4, 6, 8\}$, and $F = \{3, 5, 7, 9\}$. We have that $D \cap E = E \cap F = \varnothing$ but $D \cap F = \{3, 5, 7\}$, hence the set $\{D, E, F\}$ is not pairwise disjoint.

Observe that if $X_i = \varnothing$ for any index i, then $X_i \cap X_j = \varnothing$ for all indices j by the Going-Down Property of Set Intersection, hence any indexed collection of sets $\{X_i\}_{i \in I}$ containing the empty set is pairwise disjoint. Consequently, we may restrict our attention to collections of nonempty pairwise disjoint sets. We say that an indexed collection of sets $\mathcal{P} = \{X_i\}_{i \in I}$ forms a **partition** of a set X if

(a.) the set X_i are nonempty, i.e., $X_i \neq \varnothing$ for each element $i \in I$;

(b.) the sets X_i cover the set X, i.e., $X = \cup_{i \in I} X_i$; and

(c.) the sets X_i are pairwise disjoint, i.e., $X_i \cap X_j = \varnothing$ for every pair of distinct indices $i, j \in I$.

Example 0.5.3. Every set X admits a canonical partition $\mathcal{X} = \{\{x\}\}_{x \in X}$ indexed by the **singleton** sets $\{x\}$ for each element $x \in X$; however, many sets admit more interesting partitions.

Example 0.5.4. Consider the sets $A = \{1, 4, 7\}$, $B = \{2, 5, 8\}$, and $C = \{3, 6, 9\}$ of Example 0.5.1. Considering that the sets A, B, and C are pairwise disjoint and $A \cup B \cup C = \{1, 2, \dots, 9\} = [9]$, it follows that the set $\mathcal{P} = \{A, B, C\}$ constitutes a partition of the finite set $[9] = \{1, 2, \dots, 9\}$.

Conversely, even though the nonempty sets $D = \{1, 3, 5, 7\}$, $E = \{2, 4, 6, 8\}$, and $F = \{3, 5, 7, 9\}$ of Example 0.5.2 satisfy $[9] = D \cup E \cup F$, they are not pairwise disjoint and do not partition $[9]$.

Example 0.5.5. Consider the set $\mathbb{Z}$ of integers. Given any integer n, we may divide n by 3 in such a manner that the quotient q and remainder r of this division are unique and $0 \leq r \leq 2$. Consequently, every integer n can be written as $n = 3q + r$ for some unique integers q and $0 \leq r \leq 2$. We conclude that $\mathbb{Z} = X_0 \cup X_1 \cup X_2$ is a partition of $\mathbb{Z}$ with $X_r = \{3q + r \mid q \in \mathbb{Z}\}$ for each integer $0 \leq r \leq 2$.

Example 0.5.6. Every nonzero rational number can be written uniquely as a **reduced fraction** $\frac{p}{q}$ for some nonzero integers p and q that have no common divisors other than 1. Consider the indexed collection of sets $\{D_q\}_{q=1}^{\infty}$ of nonzero reduced fractions with denominator q, i.e.,

$$D_q = \left\{\frac{p}{q} : p, q \in \mathbb{Z} \setminus \{0\} \text{ and } p \text{ and } q \text{ have no common divisors other than } 1\right\}.$$

Explicitly, we have that

$$D_1 = \{\dots, -2, -1, 1, 2, \dots\},\ D_2 = \left\{\dots, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \dots\right\},\ \text{and } D_3 = \left\{\dots, -\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \dots\right\}.$$

Later in the semester, we will be able to prove that D_q and D_r are disjoint for any pair of distinct positive integers q and r. Considering that every nonzero rational number can be written as a reduced fraction, it follows that the collection of nonzero rational numbers is partitioned by $\{D_q\}_{q=1}^{\infty}$.

0.6 Cartesian Products of Sets

Given any nonempty set X, for any elements $x_1, x_2 \in X$, the **ordered pair** (x_1, x_2) is simply an ordered list with first **coordinate** x_1 and second coordinate x_2. Crucially, the ordered pairs (x_1, x_2) and (x_2, x_3) are equal if and only if $x_1 = x_2 = x_3$ for any elements $x_1, x_2, x_3 \in X$. We are already familiar with ordered pairs of real numbers: indeed, the concept arises naturally in our high school mathematics courses from intermediate algebra to calculus. Concretely, we refer to the set $X \times Y$ of all ordered pairs (x, y) such that $x \in X$ and $y \in Y$ as the **Cartesian product** of X and Y.

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

Example 0.6.1. Consider the sets $X = \{-1, 1\}$ and $Y = \{1, 2, 3\}$. We have that

$$\begin{aligned} X \times Y &= \{(-1, 1), (-1, 2), (-1, 3), (1, 1), (1, 2), (1, 3)\} \text{ and} \\ Y \times X &= \{(1, -1), (1, 1), (2, -1), (2, 1), (3, -1), (3, 1)\}. \end{aligned}$$

Consequently, the Cartesian product of sets is in general not commutative: indeed, the sets $X \times Y$ and $Y \times X$ from above are not equal because we have that $(-1, 1) \in X \times Y$ and $(-1, 1) \notin Y \times X$.

Even more, we may also consider the Cartesian product of a set with itself. We have that

$$\begin{aligned} X \times X &= \{(-1, -1), (-1, 1), (1, -1), (1, 1)\} \text{ and} \\ Y \times Y &= \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}. \end{aligned}$$

Example 0.6.2. Observe that the Cartesian product $\mathbb{Z} \times \mathbb{Z} = \{(a, b) \mid a \text{ and } b \text{ are integers}\}$ is the collection of all integer points in the **Cartesian plane** $\mathbb{R} \times \mathbb{R} = \{(x, y) \mid x \text{ and } y \text{ are real numbers}\}$.

Example 0.6.3. Given any univariate real function $f : \mathbb{R} \to \mathbb{R}$, the **graph** of f consists of all ordered pairs $(x, f(x))$ such that x is in the **domain** of f. Explicitly, if we assume that D_f is the domain of f and R_f is the **range** of f, then the graph of f is given by the Cartesian product

$$G_f = D_f \times R_f = \{(x, f(x)) \mid x \in D_f \text{ and } f(x) \in R_f\}.$$

Concretely, if $f(x) = 2x + 3$, then the graph of f is given by $G_f = \{(x, 2x + 3) \mid x \in \mathbb{R}\}$.

Crucially, if X and Y are finite sets with cardinalities $|X|$ and $|Y|$, then the Cartesian product $X \times Y$ has cardinality $|X||Y|$ because an element of $X \times Y$ is uniquely determined by the ordered pair (x, y). Consequently, we have that $\varnothing \times Y = \varnothing = X \times \varnothing$ for any finite sets X and Y. Even if X and Y are infinite, the Cartesian product with the empty set results in the empty set.

Proposition 0.6.4 (Cartesian Product of Finite Sets)**.** *Consider any finite sets X and Y.*

1.) *We have that $|X \times Y| = |X||Y|$. Consequently, the cardinality of the Cartesian product of any pair of finite sets is the product of the cardinalities of the underlying sets.*

2.) *We have that $\varnothing \times Y = \varnothing = X \times \varnothing$. Consequently, the Cartesian product of any finite set with the empty set is the empty set. Even more, this equality holds whenever X and Y are infinite.*

Proof. We will prove only the last statement of the proposition since the proof of the first statement is provided above. Certainly, if X and Y are finite, then $|\varnothing \times Y| = |\varnothing||Y| = 0 = |X||\varnothing| = |X \times \varnothing|$ so that $\varnothing \times Y = \varnothing = X \times \varnothing$. We may assume therefore that X and Y are infinite. By definition of the Cartesian product, we have that $\varnothing \times Y$ consists of all ordered pairs (x, y) such that $x \in \varnothing$ and $y \in Y$. Considering that there are no such elements $x \in \varnothing$, there are no such ordered pairs. □

0.7 Relations

Given any sets X and Y, a **relation from** X to Y is any subset R of the Cartesian product $X \times Y$. Explicitly, a relation R from X to Y consists of ordered pairs (x, y) such that $x \in X$ and $y \in Y$. We say that an element $x \in X$ is **related to** an element $y \in Y$ by R if $(x, y) \in R$, and we write that $x \, R \, y$ in this case; otherwise, if $(x, y) \notin R$, then x is not related to y by R, and we write $x \not R \, y$.

Example 0.7.1. Consider the sets $X = \{-1, 1\}$ and $Y = \{1, 2, 3\}$ of Example 0.6.1. Observe that $|X \times Y| = |X||Y| = 6$, hence there are $|P(X \times Y)| = 2^6$ possible relations from X to Y. We may define one such relation $R = \{(1, 1), (1, 2), (1, 3)\}$ from X to Y. Under this relation, it holds that $1 \, R \, 1$, $1 \, R \, 2$, and $1 \, R \, 3$ so that 1 is related to each of the elements of Y. Conversely, we have that $-1 \not R \, 1$, $-1 \not R \, 2$, and $-1 \not R \, 3$ so that -1 is not related to any of the elements of Y.

Every relation R from a set X to a set Y induces two important sets: namely, the collection

$$\operatorname{dom}(R) = \{x \in X \mid (x, y) \in R \text{ for some element } y \in Y\}$$

consists of all elements in X are related to some element of Y by R; it is called the **domain** of the relation R from X to Y. Likewise, the **range** of the relation R from X to Y is given by

$$\operatorname{range}(R) = \{y \in Y \mid (x, y) \in R \text{ for some element } x \in X\}$$

and consists of all elements $y \in Y$ for which there exists an element of $x \in X$ that is related to y by R. Crucially, the domain of a relation R from X to Y only concerns the first coordinate of an element of R, and the range of R only takes into account the second coordinate of an element of R.

Example 0.7.2. Consider the relation $R = \{(1, 1), (1, 2), (1, 3)\}$ from $X = \{-1, 1\}$ to $Y = \{1, 2, 3\}$ of Example 0.7.1. We have that $\operatorname{dom}(R) = \{1\}$ and $\operatorname{range}(R) = \{1, 2, 3\} = Y$.

Given any relation R from a set X to a set Y, we may define the **inverse relation**

$$R^{-1} = \{(y, x) \mid (x, y) \in R\}.$$

Crucially, if R is a relation from X to Y, then R^{-1} is a relation from Y to X, i.e., $R^{-1} \subseteq Y \times X$.

Example 0.7.3. Consider the relation $R = \{(1, 1), (1, 2), (1, 3)\}$ from $\{-1, 1\}$ to $\{1, 2, 3\}$ of Example 0.7.1. We have that $R^{-1} = \{(1, 1), (2, 1), (3, 1)\}$, $\operatorname{dom}(R^{-1}) = \{1, 2, 3\}$, and $\operatorname{range}(R^{-1}) = \{1\}$.

We refer to a subset R of the Cartesian product $X \times X$ as a **relation on** X. Every set X admits a relation Δ_X called the **diagonal** of X that consists precisely of the elements of $X \times X$ of the form (x, x). Put another way, the diagonal of X is the relation $\Delta_X = \{(x, x) \mid x \in X\}$. Observe that if X is a finite set with cardinality $|X|$, then the cardinality of $X \times X$ is $|X|^2$, hence there are a total of $2^{|X|^2}$ possible relations on a set X simply because there are as many subsets of $X \times X$.

Example 0.7.4. Consider the set $X = \{-1, 1\}$. We may define relations

$$\begin{aligned}
\Delta_X &= \{(-1, -1), (1, 1)\} \text{ with } \operatorname{dom}(\Delta_X) = \{-1, 1\} = \operatorname{range}(\Delta_X), \\
R_1 &= \{(-1, 1), (1, -1)\} \text{ with } \operatorname{dom}(R_1) = \{-1, 1\} = \operatorname{range}(R_1), \text{ and} \\
R_2 &= \{(-1, -1), (-1, 1)\} \text{ with } \operatorname{dom}(R_2) = \{-1\} \text{ and } \operatorname{range}(R_2) = \{-1, 1\}.
\end{aligned}$$

Observe that $\Delta_X^{-1} = \Delta_X$ and $R_1^{-1} = R_1$ but $R_2^{-1} = \{(-1, -1,), (1, -1)\}$ is not its own inverse.

0.8 Properties of Relations

We will continue to assume that X is an arbitrary set. Recall that a relation on X is by definition a subset R of the Cartesian product $X \times X$. We will say that R is **reflexive** if and only if $(x, x) \in R$ for all elements $x \in X$ if and only if R contains the diagonal Δ_X of X if and only if $R \supseteq \Delta_X$. Even more, if it holds that $(y, x) \in R$ whenever $(x, y) \in R$, then R is **symmetric**. Last, if $(x, y) \in R$ and $(y, z) \in R$ together imply that $(x, z) \in R$, then we refer to the relation R as **transitive**.

Example 0.8.1. Consider the following relations on the set $X = \{x, y, z\}$.

$$\begin{aligned}
R_1 &= \{(x, y), (y, z)\} \\
R_2 &= \{(x, x), (x, y), (y, y), (y, z), (z, z)\} \\
R_3 &= \{(x, y), (y, x)\} \\
R_4 &= \{(x, y), (y, z), (x, z)\} \\
R_5 &= \{(x, x), (x, y), (y, x), (y, y), (y, z), (z, y), (z, z)\} \\
R_6 &= \{(x, x), (x, y), (x, z), (y, y), (y, z), (z, z)\} \\
R_7 &= \{(x, x), (x, y), (y, x)\} \\
R_8 &= \{(x, x), (x, y), (x, z), (y, x), (y, y), (y, z), (z, x), (z, y), (z, z)\}
\end{aligned}$$

Observe that R_1 is not reflexive because (x, x) does not lie in R_1; it is not symmetric because (x, y) lies in R_1 and yet (y, x) does not lie in R_1; and it is not transitive because (x, y) and (y, z) both lie in R_1 and yet (x, z) does not lie in R_1. We note that R_2 is reflexive, but it is not symmetric because it contains (x, y) but not (y, x), and it is not transitive because it contains (x, y) and (y, z) but not (x, z). Continuing in this manner, the reader should verify the properties of the following table.

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
reflexive		✓			✓	✓		✓
symmetric			✓		✓		✓	✓
transitive				✓		✓	✓	✓

Example 0.8.2. Consider the relation R defined on the set $\mathbb{Z}$ of integers such that for any pair of integers $x, y \in \mathbb{Z}$, we have that $x \; R \; y$ if and only if $x \leq y$. Certainly, every integer x is equal to itself, hence we have that $x \leq x$ so that R is reflexive; however, we note that R is not symmetric since the strict inequality $0 < 1$ implies that $0 \; R \; 1$ and yet $1 \not R \; 0$. Last, it is straightforward to verify that R is transitive because if $x \; R \; y$ and $y \; R \; z$, then $x \leq y \leq z$ so that $x \leq z$ and $x \; R \; z$.

Example 0.8.3. Consider the relation S defined on the set $\mathbb{Z}$ of integers such that for any pair of integers $x, y \in \mathbb{Z}$, we have that $x \; S \; y$ if and only if $x \neq y$. Contrary to Example 0.8.2, this relation is symmetric but neither reflexive nor transitive: indeed, one can readily check that $x \; S \; y$ if and only if $y \; S \; X$, hence S is symmetric; however, we have that $0 = 0$ so that $0 \not S \; 0$ and S is not reflexive. Likewise, we have that $0 \neq 1$ and $1 \neq 0$ so that $0 \; S \; 1$ and $1 \; S \; 0$ but $0 \not S \; 0$, hence S is not transitive.

Example 0.8.4. Consider the relation D defined on the set $\mathbb{R}$ of real numbers such that $x \; D \; y$ if and only if $|x - y| \leq 1$. We can immediately verify that D is reflexive and symmetric: indeed, we have that $|x - x| = 0$ so that $x \; D \; x$ and $|y - x| = |x - y|$ so that $y \; D \; x$ if and only if $x \; D \; y$; however, $0 \; D \; 1$ and $1 \; D \; 2$ do not together imply that $0 \; D \; 2$ because $|2 - 0| > 1$, so D is not transitive.

0.9 Equivalence Relations

Relations that are reflexive, symmetric, and transitive are distinguished as **equivalence relations**.

Example 0.9.1. Consider any set X. We may define a relation R on X by declaring that $x \mathrel{R} y$ if and only if $x = y$. Equality is reflexive because $x = x$ holds for all elements $x \in X$; it is symmetric because $x = y$ implies that $y = x$ for any elements $x, y \in X$; and it is transitive because if $x = y$ and $y = z$, then $x = y = z$ implies that $x = z$ for all elements $x, y, z \in X$. Consequently, equality is an equivalence relation. We synthesize the result of this example in the following proposition.

Proposition 0.9.2. *Given any set X, the diagonal $\Delta_X = \{(x, x) \mid x \in X\}$ of X is an equivalence relation on X. Explicitly, every set admits at least one equivalence relation on itself.*

Proof. Observe that as a relation on X, the diagonal of X captures equality of the elements of X: if $(x, y) \in \Delta_X$, then we must have that $x = y$. Conversely, if $x = y$, then $(x, y) \in \Delta_X$. Put another way, the relation Δ_X can be identified with the equality equivalence relation of Example 0.9.1. □

Example 0.9.3. Consider the collection $\mathcal{C}^1(\mathbb{R})$ of functions $f : \mathbb{R} \to \mathbb{R}$ such that the first derivative $f'(x)$ of $f(x)$ is continuous for all real numbers x. We may define a relation R on $\mathcal{C}^1(\mathbb{R})$ such that $(f, g) \in R$ if and only if $f'(x) = g'(x)$ for all real numbers x. Because R is defined by equality and equality is reflexive, symmetric, and transitive, it follows that R is an equivalence relation on $\mathcal{C}^1(\mathbb{R})$.

Example 0.9.4. Consider the relation R defined the set $\mathbb{Z}$ of integers such that $x \mathrel{R} y$ if and only if $y - x = 2k$ for some integer k. Considering that $x - x = 0 = 2 \cdot 0$, it follows that R is reflexive. Even more, if $y - x = 2k$ for some integer k, then $x - y = -(y - x) = 2(-k)$ for the integer $-k$, hence R is symmetric. Last, if $y - x = 2k$ and $z - y = 2\ell$ for some pair of integers k and ℓ, then $z - x = (z - y) + (y - x) = 2\ell + 2k = 2(\ell + k)$ for the integer $\ell + k$. Consequently, the relations $x \mathrel{R} y$ and $y \mathrel{R} z$ together yield that $x \mathrel{R} z$. We conclude that R is an equivalence relation on $\mathbb{Z}$.

Example 0.9.5. Often, it is useful to determine if a relation is an equivalence relation by examining its elements explicitly. Consider the following relation defined on the set $[5] = \{1, 2, 3, 4, 5\}$.

$$R = \{(1, 1), (1, 3), (1, 5), (2, 2), (2, 4), (3, 1), (3, 3), (3, 5), (4, 2), (4, 4), (5, 1), (5, 3), (5, 5)\}$$

Considering that R contains the diagonal of $[5]$, it follows that R is reflexive. Put another way, we have that $(x, x) \in R$ for all elements $x \in [5]$. Even more, for each element $(x, y) \in R$, we have that $(y, x) \in R$ so that R is symmetric. Last, one can readily verify that if (x, y) and (y, z) both lie in R, then (x, z) lies in R, hence R is transitive. We conclude that R is an equivalence relation on $[5]$.

Given an equivalence relation E defined on a set X, we say that x and y are **equivalent modulo** E provided that x is related to y by E. We note that this convention is due to Carl Friedrich Gauss to express that x and y are "the same up to differences accounted for by E." We may define the **equivalence class** $[x]$ of an element $x \in X$ modulo the equivalence relation E as the set of elements $y \in X$ that are equivalent to x modulo E. Consequently, the equivalence class of x modulo E is

$$[x] = \{y \in X \mid y \mathrel{E} x\} = \{y \in X \mid (y, x) \in E\}.$$

Example 0.9.6. Every element of a set X lies in its own equivalence class modulo the equivalence relation $\Delta_X = \{(x, x) \mid x \in X\}$ because the elements of Δ_X are precisely the ordered pairs (x, x). Consequently, the equivalence class of any element $x \in X$ modulo Δ_X is the singleton $[x] = \{x\}$.

Example 0.9.7. Consider the equivalence relation R defined on the set $\mathcal{C}^1(\mathbb{R})$ of Example 0.9.3. Given any functions $f, g \in \mathcal{C}^1(\mathbb{R})$, because $f'(x)$ and $g'(x)$ are continuous for all real numbers x, it follows that $f(x) - g(x)$ is continuous and differentiable on every open interval of the form $(0, x)$. Consequently, the Mean Value Theorem ensures the existence of a real number $0 < c < x$ such that

$$f(x) - g(x) = [f'(c) - g'(c)]x + [f(0) - g(0)].$$

Observe that if $f'(x) = g'(x)$ for all real numbers x, then $f'(c) - g'(c) = 0$, and there exists a real number C such that $g(x) = f(x) + C$. Conversely, if $g(x) = f(x) + C$ for some real number C, then $f'(x) = g'(x)$. We conclude that the equivalence classes of $\mathcal{C}^1(\mathbb{R})$ modulo R are given precisely by the sets $[f] = \{g \in \mathcal{C}^1(\mathbb{R}) \mid (g, f) \in R\} = \{g \in \mathcal{C}^1(\mathbb{R}) \mid g(x) = f(x) + C \text{ for some real number } C\}$.

Example 0.9.8. Consider the equivalence relation R of Example 0.9.4. By definition, if $x = 2k$ for some integer k, then $x - 0 = 2k$, hence $(x, 0)$ lies in R. Conversely, if $(x, 0)$ lies in R, then $x = 2k$ for some integer k. We conclude that $[0] = \{x \in R \mid (x, 0) \in R\} = \{x \in R \mid x = 2k \text{ for some integer } k\}$. Likewise, if $x = 2k + 1$ for some integer k, then $x - 1 = 2k$ for some integer k so that $(x, 1)$ lies in R. Even more, if $(x, 1)$ lies in R, then $x - 1 = 2k$ and $x = 2k + 1$ for some integer k. Considering this in terms of R, we have that $[1] = \{x \in R \mid (x, 1) \in R\} = \{x \in R \mid x = 2k + 1 \text{ for some integer } k\}$. Every integer is of the form $2k$ or $2k + 1$, hence these are the equivalence classes of $\mathbb{Z}$ modulo R.

Example 0.9.9. Consider the equivalence relation R of Example 0.9.5. Each of the integers 1, 3, and 5 are equivalent modulo R because $(1, 3)$ and $(3, 5)$ lie in the equivalence relation R. On the other hand, the integers 2 and 4 are equivalent modulo R because $(2, 4)$ lies in R; thus, there are two distinct equivalence classes modulo R — namely, $[1] = \{1, 3, 5\} = [3] = [5]$ and $[2] = \{2, 4\} = [4]$.

0.10 Properties of Equivalence Classes

Given any nonempty relation E defined on a nonempty set X, we recall that E is an equivalence relation provided that E is reflexive, symmetric, and transitive. Each equivalence relation E defined on X induces a collection of sets defined on X called the equivalence classes of the elements of X. Explicitly, the equivalence class $[x]$ of an element $x \in X$ is defined by $[x] = \{y \in E \mid (y, x) \in E\}$. We demonstrate next that a pair of equivalence classes of elements of X modulo E are either equal or disjoint; as a corollary, we obtain a relationship between equivalence relations and partitions.

Proposition 0.10.1 (Equality of Equivalence Classes)**.** *Consider any equivalence relation E defined on a nonempty set X. Given any elements $x, y \in X$, we have that $[x] = [y]$ if and only if $(x, y) \in E$.*

Proof. By definition of $[x]$, for any element $z \in [x]$, we have that $(z, x) \in E$, hence the symmetry of the equivalence relation E yields that $(x, z) \in E$. Given that $[x] = [y]$, we have that $z \in [y]$ so that $(z, y) \in E$. Last, the transitivity of E ensures that $(x, y) \in E$ because (x, z) and (z, y) lie in E.

Conversely, we will assume that $(x, y) \in E$. We must demonstrate that $[x] \subseteq [y]$ and $[y] \subseteq [x]$. Given any element $z \in [x]$, we have that $(z, x) \in E$. By assumption that $(x, y) \in E$, the transitivity of the equivalence relation E yields that $(z, y) \in E$ so that $z \in [y]$. Likewise, for any element $w \in [y]$, we have that $(w, y) \in E$. By the symmetry of the equivalence relation E, we have that $(y, x) \in E$ by assumption that $(x, y) \in E$, hence the transitivity of E yields that $(w, x) \in E$ so that $w \in [x]$. □

Proposition 0.10.2 (Equivalence Classes Are Either the Same or Disjoint)**.** *Consider any equivalence relation E defined on a nonempty set X. Given any elements $x, y \in X$, the classes $[x]$ and $[y]$ of x and y modulo E are the same or disjoint. Explicitly, we must have that $[x] = [y]$ or $[x] \cap [y] = \varnothing$.*

Proof. Consider any pair $[x]$ and $[y]$ of equivalence classes of a set X modulo an equivalence relation E. We have nothing to prove if $[x] \cap [y] = \varnothing$, hence we may assume that this is not the case and prove that $[x] = [y]$. Concretely, we will assume that there exists an element $w \in [x] \cap [y]$. Crucially, by definition of the equivalence classes of X modulo E, we have that $(w, x) \in E$ and $(w, y) \in E$. By assumption that E is an equivalence relation, it follows that $(x, w) \in E$ by symmetry, hence the transitivity of E together with the inclusions $(x, w), (w, y) \in E$ yield that $(x, y) \in E$. By Proposition 0.10.1, we conclude that $[x] = [y]$, hence the proposed result is in fact established. □

Corollary 0.10.3 (Equivalence Relations and Partitions)**.** *Each equivalence relation on a nonempty set X induces a partition of X. Each partition of X induces an equivalence relation on X.*

Proof. By Proposition 0.10.2, for any equivalence relation E on a nonempty set X, the collection $\mathcal{P}$ of distinct equivalence classes of X modulo E is pairwise disjoint. Considering that every element of X lies in its own equivalence class, we conclude that $X = \cup_{C \in \mathcal{P}} C$ is a partition of X.

Conversely, we will assume that $\mathcal{P} = \{X_i\}_{i \in I}$ is a partition of X indexed by some set I. Consider the relation $E_{\mathcal{P}} = \{(x, y) \mid x, y \in X_i \text{ for some index } i \in I\} \subseteq X \times X$. By definition of a partition, every element $x \in X$ lies in X_i for some index $i \in I$, hence we have that $(x, x) \in E_{\mathcal{P}}$ for every element $x \in X$ so that $E_{\mathcal{P}}$ is reflexive. By definition of $E_{\mathcal{P}}$, if $(x, y) \in E_{\mathcal{P}}$, then $(y, x) \in E_{\mathcal{P}}$, hence $E_{\mathcal{P}}$ is symmetric. Last, if $(x, y), (y, z) \in E_{\mathcal{P}}$, then $x, y \in X_i$ and $y, z \in X_j$ for some indices $i, j \in I$. By definition of a partition, we have that $X_i \cap X_j = \varnothing$ if and only if i and j are distinct, hence we must have that $i = j$ by assumption that $y \in X_i \cap X_j$. We conclude that $(x, z) \in X_i$ so that $(x, z) \in E_{\mathcal{P}}$ and $E_{\mathcal{P}}$ is transitive. Ultimately, we find that $E_{\mathcal{P}}$ is an equivalence relation on X. □

Example 0.10.4. Consider the equivalence relation R of Example 0.9.5. By Corollary 0.10.3, the collection of distinct equivalence classes of $[5]$ modulo R provides a partition of $[5]$. By Example 0.9.9, the distinct equivalence classes of $[5]$ modulo R are $[1] = \{1, 3, 5\}$ and $[2] = \{2, 4\}$, hence the underlying partition of $[5]$ induced by the equivalence relation R is $\mathcal{P} = \{[1], [2]\} = \{\{1, 3, 5\}, \{2, 4\}\}$.

Example 0.10.5. Consider the following partition $\mathcal{P} = \{R_0, R_1, R_2, R_3\}$ of the set $\mathbb{Z}$ of integers.

$$R_0 = \{\dots, -8, -4, 0, 4, \dots\} \qquad R_2 = \{\dots, -6, -2, 2, 6, \dots\}$$
$$R_1 = \{\dots, -7, -3, 1, 5, \dots\} \qquad R_3 = \{\dots, -5, -1, 3, 7, \dots\}$$

By Corollary 0.10.3, the distinct sets in the partition $\mathcal{P}$ constitute the distinct equivalence classes of an equivalence relation $E_{\mathcal{P}}$ on $\mathbb{Z}$. Explicitly, we have that $(x, y) \in E_{\mathcal{P}}$ if and only if $x, y \in R_i$ for some integer $1 \leq i \leq 4$. Consequently, the distinct equivalence classes of $\mathbb{Z}$ modulo the equivalence relation $E_{\mathcal{P}}$ are R_0, R_1, R_2, and R_3. Observe that $(0, 4) \in E_{\mathcal{P}}$ holds because $0, 4 \in R_0$ and $(1, 5) \in E_{\mathcal{P}}$ holds because $1, 5 \in R_1$, but neither $(0, 2)$ nor $(1, 3)$ lie in $E_{\mathcal{P}}$. By Proposition 0.10.1, a pair of equivalence classes are distinct if and only if their **representatives** are related, hence the distinct equivalence classes of $\mathbb{Z}$ modulo $E_{\mathcal{P}}$ are $[0]$, $[1]$, $[2]$, and $[3]$ or similarly $[4]$, $[5]$, $[6]$, and $[7]$ and so on.

0.11 Congruence Modulo n

We say that a nonzero integer a **divides** an integer b if there exists an integer c such that $b = ac$. We will write $a \mid b$ in this case, and we will often say that b is **divisible by** a. Given any nonzero integer n, we say that a pair of integers a and b are **congruent modulo** n if it holds that n divides $b - a$. Conventionally, if a and b are congruent modulo n, we write $b \equiv a \pmod{n}$.

Example 0.11.1. We have that $7 \equiv 3 \pmod 4$ because $7 - 3 = 4$ is divisible by 4.

Example 0.11.2. We have that $5 \equiv 21 \pmod 4$ because $5 - 21 = -16 = 4(-4)$ is divisible by 4.

Example 0.11.3. We have that $11 \not\equiv 8 \pmod 4$ because $11 - 8 = 3$ is not divisible by 4.

Proposition 0.11.4. *Consider any nonzero integer n and any integers a, b, and c.*

(1.) *We have that $a \equiv 0 \pmod{n}$ if and only if n divides a.*

(2.) *We have that $b \equiv a \pmod{n}$ if and only if $b - a \equiv 0 \pmod{n}$.*

(3.) *We have that $a \equiv a \pmod{n}$ for any integer a.*

(4.) *We have that $b \equiv a \pmod{n}$ if and only if $a \equiv b \pmod{n}$.*

(5.) *If $b \equiv a \pmod{n}$ and $c \equiv b \pmod{n}$, then $c \equiv a \pmod{n}$.*

(6.) *We have that $b \equiv a \pmod{n}$ if and only if $-b \equiv -a \pmod{n}$.*

(7.) *We have that $b \equiv a \pmod{n}$ if and only if $b + c \equiv a + c \pmod{n}$.*

(8.) *If $b \equiv a \pmod{n}$, then $cb \equiv ca \pmod{n}$.*

(9.) *If $b \equiv a \pmod{n}$, then $b^k \equiv a^k \pmod{n}$ for any integer $k \geq 0$.*

Proof. (1.) We have that $a \equiv 0 \pmod{n}$ if and only if n divides $a - 0$ if and only if n divides a.

(2.) By definition and the first property of congruence modulo n, we have that $b \equiv a \pmod{n}$ if and only if n divides $b - a$ if and only if $b - a \equiv 0 \pmod{n}$.

(3.) Considering that $a - a = 0 = n \cdot 0$, it follows that n divides $a - a$ so that $a \equiv a \pmod{n}$.

(4.) We have that $b \equiv a \pmod{n}$ if and only if n divides $b - a$ if and only if n divides $-(a - b)$ if and only if n divides $a - b$ if and only if $a \equiv b \pmod{n}$.

(5.) Given that $b \equiv a \pmod{n}$ and $c \equiv b \pmod{n}$, by definition, there exist integers k and ℓ such that $b - a = nk$ and $c - b = n\ell$. Observe that $c - a = (c - b) + (b - a) = nk + n\ell = n(k + \ell)$, hence n divides $c - a$ so that $c \equiv a \pmod{n}$ by definition of congruence modulo n.

(6.) We have that $b \equiv a \pmod{n}$ if and only if n divides $b - a$ if and only if $b - a = nk$ for some integer k if and only if $-b + a = (-b) - (-a) = n(-k)$ for some integer k if and only if n divides $-b - (-a)$ if and only if $-b \equiv -a \pmod{n}$ by definition of congruence modulo n.

(7.) By definition of congruence modulo n, we have that $b \equiv a \pmod{n}$ if and only if n divides $b - a$ if and only if n divides $(b + c) - (a + c)$ if and only if $b + c \equiv a + c \pmod{n}$.

(8.) By definition of congruence modulo n, if $b \equiv a \pmod{n}$, then n divides $b - a$ so that n divides $c(b - a)$. Considering that $c(b - a) = cb - ca$, it follows that $cb \equiv ca \pmod{n}$.

(9.) By the eight property of congruence modulo n, we have that $b^2 = b \cdot b \equiv b \cdot a \pmod{n}$ and $a^2 = a \cdot a \equiv a \cdot b \pmod{n}$. Considering that $b \cdot a = a \cdot b$, the fifth property of congruence modulo n yields that $b^2 = b \cdot b \equiv b \cdot a = a \cdot b \equiv a \cdot a = a^2 \pmod{n}$. By the same rationale, we have that $b^3 = b \cdot b^2 \equiv b \cdot a^2 = a \cdot a^2 = a^3 \pmod{n}$. Continuing in this manner, the desired result follows. □

Given any nonzero integer n, consider the relation R defined on the set $\mathbb{Z}$ of integers such that $a \, R \, b$ if and only if $b \equiv a \pmod{n}$. We refer to R as **congruence modulo** n. By the third, fourth, and fifth properties of Proposition 0.11.4, congruence modulo n is an equivalence relation on $\mathbb{Z}$.

Proposition 0.11.5. *Congruence modulo any nonzero integer n is an equivalence relation on $\mathbb{Z}$.*

Consider the equivalence class $[a]$ of any integer a modulo the equivalence relation of congruence modulo n. Conventionally, we refer to $[a]$ as the class of a **modulo** n. By definition, we have that

$$[a] = \{b \in \mathbb{Z} \mid b \equiv a \pmod{n}\} = \{b \in \mathbb{Z} \mid b - a = nq \text{ for some integer } q\} = \{nq + a \mid q \in \mathbb{Z}\}.$$

Consequently, the equivalence class of a modulo n consists of sums of integer multiples of n and a.

Example 0.11.6. Congruence modulo 2 is an equivalence relation on $\mathbb{Z}$ with equivalence classes

$$\begin{aligned} [0] &= \{2q + 0 \mid q \in \mathbb{Z}\} = \{\dots, -4, -2, 0, 2, 4, \dots\} \text{ and} \\ [1] &= \{2q + 1 \mid q \in \mathbb{Z}\} = \{\dots, -3, -1, 1, 3, 5, \dots\}. \end{aligned}$$

By Proposition 0.10.2, these are all of the distinct equivalence classes of $\mathbb{Z}$ modulo 2.

Example 0.11.7. Congruence modulo 3 is an equivalence relation on $\mathbb{Z}$ with equivalence classes

$$\begin{aligned} [0] &= \{3q + 0 \mid q \in \mathbb{Z}\} = \{\dots, -6, -3, 0, 3, 6, \dots\}, \\ [1] &= \{3q + 1 \mid q \in \mathbb{Z}\} = \{\dots, -5, -2, 1, 4, 7, \dots\}, \text{ and} \\ [2] &= \{3q + 2 \mid q \in \mathbb{Z}\} = \{\dots, -4, -1, 2, 5, 8, \dots\}. \end{aligned}$$

By Proposition 0.10.2, these are all of the distinct equivalence classes of $\mathbb{Z}$ modulo 3.

Proposition 0.11.8. *Given any nonzero integer n, the distinct equivalence classes of $\mathbb{Z}$ modulo n are $[i] = \{nq + i \mid q \in \mathbb{Z}\}$ for each integer $0 \leq i \leq n - 1$. Particularly, there are exactly n of them.*

Congruence modulo a nonzero integer gives rise to other interesting equivalence relations.

Example 0.11.9. Consider the relation R on the set $\mathbb{Z}$ of integers defined by $a \, R \, b$ if and only if $5b \equiv 2a \pmod{3}$. We claim that R is an equivalence relation.

1.) We demonstrate first that $a \, R \, a$. By definition, we must prove that $5a \equiv 2a \pmod{3}$. But this is true because $5a - 2a = 3a$ is divisible by 3 for all integers a.

2.) We establish next that if $a \, R \, b$, then $b \, R \, a$. By definition, if $a \, R \, b$, then $5b \equiv 2a \pmod{3}$ so that $5b - 2a = 3k$ for some integer k. Consequently, we have that $2a - 5b = 3(-k)$. By adding $3a$ and $3b$ to both sides of this equation, we obtain $5a - 2b = 3(-k) + 3a + 3b = 3(-k + a + b)$. We conclude that $5a - 2b$ is divisible by 3 so that $5a \equiv 2b \pmod{3}$ and $b \, R \, a$.

3.) Last, if $a \mathrel{R} b$ and $b \mathrel{R} c$, then $5b \equiv 2a \pmod 3$ and $5c \equiv 2b \pmod 3$. By definition, there exist integers k and ℓ such that $5b - 2a = 3k$ and $5c - 2b = 3\ell$. By taking their sum, we find that

$$5c - 3b - 2a = (5c - 2b) + (5b - 2a) = 3\ell + 3k = 3(\ell + k)$$

so that $5c - 2a = 3(\ell + k + b)$; therefore, 3 divides $5c - 2a$ so that $5c \equiv 2a \pmod 3$ and $a \mathrel{R} c$.

By definition, the equivalence class of a modulo R is given by

$$[a] = \{b \in \mathbb{Z} \mid a \mathrel{R} b\} = \{b \in \mathbb{Z} \mid 5b \equiv 2a \pmod 3\} = \{b \in \mathbb{Z} \mid 5b - 2a = 3k \text{ for some integer } k\}.$$

Consequently, the class of a modulo R is $[a] = \{b \in \mathbb{Z} \mid 5b = 3k + 2a \text{ for some integer } k\}$. Checking some small values of b yields that $[0] = \{\dots, -6, -3, 0, 3, 6, \dots\}$. Likewise, by definition and a brute-force check, we have that $[1] = \{b \in \mathbb{Z} \mid 5b = 3k + 2 \text{ for some integer } k\} = \{\dots, -5, -2, 1, 4, 7, \dots\}$ and $[2] = \{b \in \mathbb{Z} \mid 5b = 3k+4 \text{ for some integer } k\} = \{\dots, -4, -1, 2, 5, 8, \dots\}$. Every integer belongs to one of these three distinct equivalence classes modulo R, hence this is an exhaustive list.

0.12 The Definition of a Function

Consider any sets X and Y. We have seen previously that a relation from X to Y is any subset of the Cartesian product $X \times Y$. We say that a relation f from X to Y is a **function** if and only if every element of X is the first component of one and only one ordered pair in f. Explicitly, a function $f : X \to Y$ is merely an assignment of each element $x \in X$ to a unique (but not necessarily distinct) element $f(x) \in Y$ called the **direct image** of x under f. We refer to the set X as the **domain** of $f : X \to Y$; the **codomain** of f is Y; and the **range** of f is the set $\operatorname{range}(f) = \{f(x) \mid x \in X\}$ of second coordinates of elements in f. Out of desire for notational convenience, we may sometimes omit the letter $f : X \to Y$ when defining a function and simply use an arrow $X \to Y$ to indicate the sets involved and an arrow $x \mapsto y$ to declare the image $y \in Y$ of the element $x \in X$.

Example 0.12.1. Consider the relation $f = \{(-1, 1), (1, -1)\}$ defined on $X = \{-1, 1\}$. Each of the elements of X is the first component of one and only one ordered pair in f, hence $f : X \to X$ is a function; its domain and range are both X. Conventionally, we might recognize this function as $f(x) = -x$ because it has the effect of swapping the signs of each element $x \in X$.

Example 0.12.2. Consider the relation $g = \{(x, x - 1) \mid x \in \mathbb{R}\}$ defined on the set $\mathbb{R}$ of real numbers. Every real number is the first component of one and only one ordered pair in g, hence $g : \mathbb{R} \to \mathbb{R}$ is a function; its domain and range are both $\mathbb{R}$. Conventionally, we might recognize this function as $g(x) = x - 1$ because the ordered pairs $(x, y) \in g$ satisfy that $y = x - 1$ for each real number x.

Example 0.12.3. Often in calculus, a function is defined simply by declaring a rule, e.g., $h(x) = x^2$. Conventionally, the domain of such a function is assumed to be the **natural domain**, i.e., the largest subset of the real numbers for which $h(x)$ can be defined. Considering that the square of any real number is itself a real number, it follows that the domain of $h(x)$ is all real numbers; the range of $h(x)$ is the collection of all non-negative real numbers because if $x \in \mathbb{R}$, then $x^2 \geq 0$.

But strictly speaking, a function intimately depends on its domain and its codomain. We will soon see that the functions $h : \mathbb{R} \to \mathbb{R}_{\geq 0}$ defined by $h(x) = x^2$ and $k : \mathbb{R}_{\geq 0} \to \mathbb{R}$ defined by $k(x) = x^2$

are quite different from one another — even though the underlying rule of both functions is the same. Even more, both of these functions are different from $\ell : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by $\ell(x) = x^2$.

Example 0.12.4. Consider the equivalence relation R defined on the set $[5] = \{1, 2, 3, 4, 5\}$ as in Example 0.8.4. Because the ordered pairs $(1, 1)$ and $(1, 3)$ lie in R, it follows that R is not a function. Generally, an equivalence relation R will never be a function because if (x, y) and (y, x) both lie in R, then by definition, we must have that $(x, x) \in R$ so that R is not a function.

Every set X possesses an **identity function** $\operatorname{id}_X : X \to X$ defined by $\operatorname{id}_X(x) = x$. If X is a subset of Y, then the **inclusion** $X \subseteq Y$ can be viewed as the function $X \to Y$ that sends $x \mapsto x$, where the symbol x appearing to the left of the arrow $\mapsto$ is viewed as an element of X, and the symbol x appearing to the right of the arrow $\mapsto$ is viewed as an element of Y; in the usual notation, the inclusion may be thought of as the function $i : X \to Y$ defined by $i(x) = x$. Even more, every set X induces a function $\delta_X : X \to X \times X$ that is called the **diagonal function** (of X) and defined by $\delta_X(x) = (x, x)$. Later in the course, we will prove that the diagonal Δ_X of X is exactly the image of the diagonal function δ_X of X, hence there should be no confusion in terminologies.

Even if we have never thought of it as such, algebraic operations such as addition, subtraction, multiplication, and division can be viewed as functions. Explicitly, addition of real numbers is the function $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $(x, y) \mapsto x + y$. Crucially, the sum of two real numbers is a real number, hence this function is **well-defined**, i.e., the image of every element lies in the codomain of the function. Generally, if X is any set, the function $* : X \times X \to X$ that sends $(x, y) \mapsto x * y$ is a **binary operation** if and only if $x * y$ is an element of X for every pair of elements $x, y \in X$. Like we mentioned, addition and multiplication are binary operations on the real numbers $\mathbb{R}$.

Consider any pair of functions $f : X \to Y$ and $g : X \to Y$. Given any element $x \in X$, there exist unique elements $f(x), g(x) \in Y$ such that $(x, f(x)) \in f$ and $(x, g(x)) \in g$. Consequently, if f and g are equal as sets so that $f = g$, then $(x, f(x))$ lies in g; the uniqueness of $g(x)$ yields in turn that $f(x) = g(x)$. Conversely, if $f(x) = g(x)$ for every element $x \in X$, then we have that

$$f = \{(x, f(x)) \mid x \in X\} = \{(x, g(x)) \mid x \in X\} = g$$

so that f and g are equal as sets. We have proved the following important fact about functions.

Proposition 0.12.5. *Given any sets X and Y, a pair of functions $f : X \to Y$ and $g : X \to Y$ are equal as sets if and only if $f(x) = g(x)$ for all elements $x \in X$.*

Each time we define a function $f : X \to Y$, for every subset $V \subseteq X$, we implicitly distinguish the collection of elements $y \in Y$ such that $y = f(v)$ for some element $v \in V$; this is denoted by

$$f(V) = \{f(v) \mid v \in V\}$$

and called the **direct image** of V (in Y) under f. Conversely, if $W \subseteq Y$, then the collection of elements $x \in X$ such that $f(x) \in W$ is the **inverse image** of W (in X) under f. Explicitly, we have that

$$f^{-1}(W) = \{x \in X \mid f(x) \in W\}.$$

Example 0.12.6. Consider the function $f = \{(u, 1), (v, 2), (w, 3), (x, 3), (y, 2), (z, 1)\}$ from the set $X = \{u, v, w, x, y, z\}$ to $Y = [6] = \{1, 2, 3, 4, 5, 6\}$. We have that $\operatorname{range}(f) = \{1, 2, 3\}$, but it holds that $\operatorname{range}(f) = f(\{u, v, w\}) = f(\{u, x, y\}) = f(\{x, y, z\})$ to name a few. Even more, we have that

$$f^{-1}(\{2, 3\}) = \{v, w, x, y\} \text{ and } f^{-1}(\{4, 5, 6\}) = \varnothing$$

because the elements $4, 5, 6 \in Y$ do not belong to the second component of any ordered pair in f.

Example 0.12.7. Consider the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2$. Observe that for any real number x such that $-1 \leq x \leq 1$, we have that $0 \leq x^2 \leq 1$, hence it follows that $g([-1, 1]) = [0, 1]$. On the other hand, $-3 < x \leq 2$, then $0 \leq x^2 < 9$ implies that $g((-3, 2]) = [0, 9)$.

Even if the sets X and Y are finite with small cardinalities $|X|$ and $|Y|$, the number of functions $f : X \to Y$ grows astonishingly quickly. Explicitly, a function $f : X \to Y$ is uniquely determined by choosing for each element $x \in X$ one and only one element $y \in Y$ such that $f(x) = y$. Consequently, for each element $x \in X$, there are $|Y|$ possible choices for $f(x)$. By denoting the set of functions $f : X \to Y$ as $Y^X = \{f \subseteq X \times Y \mid f : X \to Y \text{ is a function}\}$, we have that $|Y^X| = |Y|^{|X|}$.

Example 0.12.8. Consider the sets $X = \{u, v, w, x, y, z\}$ and $Y = [6] = \{1, 2, 3, 4, 5, 6\}$ of Example 0.12.6. We have that $|X| = 6 = |Y|$, hence there are $|Y|^{|X|} = 6^6$ possible functions $f : X \to Y$.

0.13 One-to-One and Onto Functions

We introduce two indispensable properties of a function $f : X \to Y$ from a set X to a set Y. We say that f is **one-to-one** (or **injective**) if every pair of distinct elements $x_1, x_2 \in X$ induce distinct elements $f(x_1), f(x_2) \in Y$. Equivalently, we say that f is one-to-one if every equality $f(x_1) = f(x_2)$ of elements of Y yields the corresponding equality $x_1 = x_2$ of elements of X.

Example 0.13.1. Consider the function $f = \{(-1, 1), (1, -1)\}$ from the set $X = \{-1, 1\}$ to itself. Each of the elements $x \in X$ corresponds to a distinct element $f(x) \in X$, hence f is one-to-one.

Example 0.13.2. Consider the real function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 4$. Observe that if $f(x_1) = f(x_2)$, then $3x_1 + 4 = 3x_2 + 4$ so that $3x_1 = 3x_2$ and $x_1 = x_2$; thus, f is one-to-one.

Example 0.13.3. Consider the real function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ defined by $f(x) = x^2$. Observe that if $f(x_1) = f(x_2)$, then $x_1^2 = x_2^2$. By taking the square root of both sides and using the fact that the domain of f consists of non-negative real numbers, it follows that $x_1 = x_2$ so that f is one-to-one.

Example 0.13.4. Consider the function $f = \{(u, 1), (v, 2), (w, 3), (x, 3), (y, 2), (z, 1)\}$ from the set $X = \{u, v, w, x, y, z\}$ to $Y = [6] = \{1, 2, 3, 4, 5, 6\}$. Considering that $f(u) = 1 = f(z)$ but $u \neq z$, it follows that f is not one-to-one; the same holds for $f(v) = 2 = f(y)$ and $f(w) = 3 = f(x)$.

Example 0.13.5. Consider the real function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2$. Considering that $g(-1) = 1 = g(1)$ but $-1 \neq 1$, it follows that g is not one-to-one. Compare with Example 0.13.3.

Example 0.13.6. We say that a real function f is **increasing** if $x_1 < x_2$ implies that $f(x_1) < f(x_2)$ for all real numbers x_1 and x_2 in the domain of f. If f is differentiable on an open interval (a, b) (i.e., $f'(x)$ exists for all real numbers $a < x < b$), then by the Mean Value Theorem, we have that f is increasing on (a, b) if and only if $f'(x) > 0$ for all real numbers $a < x < b$. Explicitly, if f is increasing on (a, b), then for any real numbers $a < x_1 < x_2 < b$, we have that $f(x_2) - f(x_1) > 0$. By the Mean Value Theorem, there exists a real number $x_1 < c < x_2$ such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} > 0.$$

Conversely, if $f'(x) > 0$ for all real numbers $a < x < b$, then for any real numbers $a < x_1 < x_2 < b$, the Mean Value Theorem guarantees the existence of a real number $x_1 < c < x_2$ such that

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0.$$

Consequently, any function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^{2n+1}$ for some integer $n \geq 0$ is increasing on any open interval not containing 0 because $f'(x) = (2n+1)x^{2n} > 0$ on any such interval.

Even more, we say that $f : X \to Y$ is **onto** (or **surjective**) if for every element $y \in Y$, there exists an element $x \in X$ such that $y = f(x)$. One way to think about the surjective property is that every element of the codomain Y is "mapped onto" or "covered" by an element of X. Even more simply, a function $f : X \to Y$ is surjective if and only if $Y = \operatorname{range}(f) = \{f(x) \mid x \in X\}$.

Example 0.13.7. Consider the function $f = \{(-1, 1), (1, -1)\}$ from the set $X = \{-1, 1\}$ to itself. Each of the elements $y \in X$ can be written as $y = f(x)$ for some element $x \in X$, hence f is onto.

Example 0.13.8. Consider the real function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 4$. We claim that f is onto. By definition, for any real number y, we must furnish a real number x such that $y = f(x) = 3x + 4$. But if $y = 3x + 4$, then $3x = y - 4$ so that $x = \frac{1}{3}(y - 4)$. Computing $f(x)$ yields

$$f(x) = 3x + 4 = 3 \cdot \frac{1}{3}(y - 4) + 4 = (y - 4) + 4 = y$$

because $x = \frac{1}{3}(y - 4)$ by construction, as desired. Consequently, it follows that f is onto.

Example 0.13.9. Consider the real function $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ defined by $f(x) = x^2$. Given any real number $y \geq 0$, we claim that there exists a real number x such that $y = x^2$. By taking $x = \sqrt{y}$ (this is well-defined because $y \geq 0$), it follows that $f(x) = x^2 = (\sqrt{y})^2 = y$ so that f is onto.

Example 0.13.10. Consider the function $f = \{(u, 1), (v, 2), (w, 3), (x, 3), (y, 2), (z, 1)\}$ from the set $X = \{u, v, w, x, y, z\}$ to $Y = [6] = \{1, 2, 3, 4, 5, 6\}$. Considering that 4, 5, and 6 are not the image of any element of X under f, it follows that f is not onto.

Example 0.13.11. Consider the sets $X = \{a, b, c\}$ and $Y = \{0, 1, 2, 3\}$. We cannot possibly find a function $f : X \to Y$ that is onto because the cardinality of X is strictly smaller than the cardinality of Y; therefore, it is impossible to assign to each element $y \in Y$ a unique element $x \in X$.

0.14 Bijective Functions

We say that a function $f : X \to Y$ is **bijective** if f is both injective and surjective. We may think of a bijection $f : X \to Y$ simply as a relabelling of the elements of Y using the names of elements of X; in this way, two sets X and Y are "essentially the same" if there exists a bijection $f : X \to Y$. Often, this property of a bijective function is emphasized in the literature by using the terminology of "one-to-one correspondence" between X and Y in place of "bijection" from X to Y.

Proposition 0.14.1. *Consider any pair of arbitrary finite sets X and Y.*

(a.) *If there exists an injective function $f : X \to Y$, then $|X| \leq |Y|$.*

(b.) *If $|X| \leq |Y|$, then there exists an injective function $f : X \to Y$.*

(c.) *If there exists a surjective function $f : X \to Y$, then $|X| \geq |Y|$.*

(d.) *If $|X| \geq |Y|$, then there exists a surjective function $f : X \to Y$.*

(e.) *If there exists a bijective function $f : X \to Y$, then $|X| = |Y|$.*

(f.) *If $|X| = |Y|$, then there exists a bijective function $f : X \to Y$.*

(g.) *If $|X| = |Y|$, then a function $f : X \to Y$ is injective if and only if it is surjective.*

Proof. We will assume throughout the proof that $|X| = m$ and $|Y| = n$ are non-negative integers. Certainly, if either m or n is zero, then the empty function satisfies the desired properties. Consequently, we may assume that neither m nor n is zero. We will assume for notational convenience that $X = \{x_1, x_2, \dots, x_m\}$ and $Y = \{y_1, y_2, \dots, y_n\}$. We turn our attention to each claim in turn.

(a.) We will assume that there exists an injective function $f : X \to Y$. Consequently, every element $y \in Y$ corresponds to at most one element $x \in X$ via $y = f(x)$. Considering that every element $x \in X$ corresponds to a unique element $f(x) \in Y$, we conclude that $|X| \leq |Y|$.

(b.) Observe that if $m \leq n$, then we may define an injective function $f : X \to Y$ by declaring that $f(x_i) = y_i$ for each integer $1 \leq i \leq m$. Explicitly, f is a function because every element $x_i \in X$ corresponds to exactly one element $y_i = f(x_i) \in Y$. Even more, f is injective since for each element $y_i \in Y$, there is at most one element $x_i \in X$ such that $y_i = f(x_i)$ by assumption that $n \geq m$.

(c.) We will assume that there exists a surjective function $f : X \to Y$. Consequently, for every element $y \in Y$, there exists an element $x \in X$ such that $y = f(x)$. Considering that every element $x \in X$ corresponds to a unique element $f(x) \in Y$, we conclude that $|X| \geq |Y|$.

(d.) Conversely, if $m \geq n$, then we may define a surjective function $f : X \to Y$ by declaring that $f(x_i) = y_i$ for each integer $1 \leq i \leq m$. We have already seen in the previous paragraph that such a relation is a function; however, by assumption that $m \geq n$, it follows that f is surjective because for every element $y_i \in Y$, there exists an element $x_i \in X$ such that $y_i = f(x_i)$.

(e.) Combined, parts (a.) and (c.) imply that $|X| \leq |Y|$ and $|X| \geq |Y|$ so that $|X| = |Y|$.

(f.) Combined, parts (b.) and (d.) yield a bijective function $f : X \to Y$ defined by $f(x_i) = y_i$.

(g.) Last, we will assume that $m = n$. Consider any function $f : X \to Y$. Observe that if f is injective, then every element of X maps to a distinct element of Y under f, hence range(f) is a subset of Y of the same cardinality as Y. We conclude that range$(f) = Y$ so that f is surjective. Conversely, if f is surjective, then for every element $y \in Y$, there exists an element $x \in X$ such that $y = f(x)$. By assumption that $m = n$, the element $x \in X$ such that $y = f(x)$ is uniquely determined by y, hence the image of $x \in X$ under f is unique so that f is injective. □

Caution: if X and Y are infinite sets, then there need not exist a bijective function $f : X \to Y$. Explicitly, there is no bijection $f : \mathbb{Q} \to \mathbb{R}$ between the rational numbers and the real numbers.

Caution: if X and Y are infinite sets, then a function $f : X \to Y$ can be injective without being surjective (and vice-versa). Explicitly, the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = 2x$ is injective but not surjective, and the function $g : \mathbb{Q} \to \mathbb{Z}$ defined by $g(p/q) = p$ is surjective but not injective.

By Proposition 0.14.1, a pair of nonempty sets admit a bijection if and only if they have the same number of elements (or cardinality). Given any nonempty set X of cardinality n, the collection of bijective functions $f : X \to X$ is an extremely important object in commutative algebra and combinatorics called the **symmetric group on the finite set** X and denoted by $\mathfrak{S}_X$.

Proposition 0.14.2. *Given any nonempty sets X and Y with $|X| = |Y| = n$, there are $n! = n(n-1)(n-2)\cdots 2\cdot 1$ distinct bijective functions $f : X \to Y$. Consequently, we have that $|\mathfrak{S}_X| = |X| = n!$.*

Proof. Every bijective function $f : X \to Y$ is uniquely determined by the images of the elements of X under f. Consequently, if we assume that $X = \{x_1, x_2, \dots, x_n\}$, then there are n distinct choices for the value of $f(x_1)$; then, there are $n - 1$ distinct choices for $f(x_2)$ other than $f(x_1)$; likewise, there are $n - 2$ distinct choices for $f(x_3)$ other than $f(x_1)$ and $f(x_2)$. Continuing in this manner, there are $n - i + 1$ choices for $f(x_i)$ for each integer $1 \leq i \leq n$, hence there are $n!$ bijective functions between X and Y: indeed, there are a total of $n! = n(n-1)(n-2)\cdots 2\cdot 1$ possibilities. □

Example 0.14.3. Observe that the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = -x$ is bijective. Explicitly, if $f(x) = f(y)$, then $-x = -y$ yields that $x = y$ so that f is one-to-one. Likewise, every integer n is the image of $-n$ under f because $n = -(-n) = f(-n)$, hence f is onto.

Example 0.14.4. Consider the function $f : \mathbb{R} \setminus \{3\} \to \mathbb{R} \setminus \{1\}$ defined by

$$f(x) = \frac{x-2}{x-3}.$$

Cross-multiplying denominators, we note that $f(x) = f(y)$ if and only if $(x-2)(y-3) = (x-3)(y-2)$ if and only if $xy - 3x - 2y + 6 = xy - 2x - 3y + 6$ if and only if $x = y$, hence f is one-to-one. Conversely, we will prove that f is onto. Behind the scenes, we solve the following equation for x.

$$y = \frac{x-2}{x-3}$$

Observe that this holds if and only if $(x-3)y = x - 2$ if and only if $xy - 3y = x - 2$ if and only if $xy - x = 3y - 2$ if and only if $x(y-1) = 3y - 2$ if and only if

$$x = \frac{3y-2}{y-1}.$$

Consequently, for every real number $y \in \mathbb{R} \setminus \{1\}$, we have that $y = f(x)$ so that f is onto.

0.15 Composition of Functions

Given any pair of functions $f : X \to Y$ and $g : Y \to Z$ between any sets X, Y, and Z, we may construct a function $g \circ f : X \to Z$ called the **composite function** defined by $(g \circ f)(x) = g(f(x))$. We may also refer to the function $g \circ f$ as g **composed with** f or the **composition** of f under g.

Example 0.15.1. Consider the sets $X = \{-1, 1\}$, $Y = \{x, y, z\}$, and $Z = \{1, 2, 3\}$. We may define some functions $f : X \to Y$ and $g : Y \to Z$ by $f = \{(-1, x), (1, z)\}$ and $g = \{(x, 2), (y, 3), (z, 1)\}$. Observe that the composite function $g \circ f : X \to Z$ satisfies $(g \circ f)(-1) = g(f(-1)) = g(x) = 2$ and $(g \circ f)(1) = g(f(1)) = g(z) = 1$. Consequently, we find that $g \circ f = \{(-1, 2), (1, 1)\}$.

Example 0.15.2. Consider the sets $A = \{a, b, c, d\}$, $B = \{b, c, d, e\}$, and $C = \{c, d, e, f\}$. We may define a pair of functions $f : A \to B$ and $g : B \to C$ such that $f = \{(a, b), (b, c), (c, d), (d, e)\}$ and $g = \{(b, c), (c, d), (d, e), (e, f)\}$. Observe that the composite function $g \circ f : A \to C$ satisfies that

$$\begin{aligned} (g \circ f)(a) &= g(f(a)) = g(b) = c, & (g \circ f)(c) &= g(f(c)) = g(d) = e, \text{ and} \\ (g \circ f)(b) &= g(f(b)) = g(c) = d, & (g \circ f)(d) &= g(f(d)) = g(e) = f. \end{aligned}$$

Consequently, we find that $g \circ f : A \to C$ satisfies that $g \circ f = \{(a, c), (b, d), (c, e), (d, f)\}$.

Example 0.15.3. Composition of functions is a common technique in calculus. (Recall that the Chain Rule for Derivatives gives a formula for the derivative of a composite function.) Consider the functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = e^x$ and $g(x) = |x|$. We have that

$$\begin{aligned} &f \circ g : \mathbb{R} \to \mathbb{R} \text{ is defined by } (f \circ g)(x) = f(g(x)) = e^{g(x)} = e^{|x|} \text{ and} \\ &g \circ f : \mathbb{R} \to \mathbb{R} \text{ is defined by } (g \circ f)(x) = g(f(x)) = |f(x)| = |e^x| = e^x. \end{aligned}$$

Crucially, the latter holds because $e^x > 0$ for all real numbers x, hence it follows that $g \circ f = f$.

Proposition 0.15.4. *Consider any pair of functions $f : X \to Y$ and $g : Y \to Z$.*

(a.) *If f and g are injective, then $g \circ f$ is injective.*

(b.) *If f and g are surjective, then $g \circ f$ is surjective.*

Put another way, composition of functions preserves injectivity and surjectivity.

Proof. (a.) We must prove that if $(g \circ f)(x_1) = (g \circ f)(x_2)$, then $x_1 = x_2$. By assumption that g is injective, if $g(f(x_1)) = (g \circ f)(x_1) = (g \circ f)(x_2) = g(f(x_2))$, then $f(x_1) = f(x_2)$. But by the same rationale applied to the injective function f, we conclude that $x_1 = x_2$, as desired.

(b.) We must prove that for every element $z \in Z$, there exists an element $x \in X$ such that $z = (g \circ f)(x)$. By assumption that g is surjective, for every element $z \in Z$, there exists an element $y \in Y$ such that $z = g(y)$. Even more, by hypothesis that f is surjective, there exists an element $x \in X$ such that $y = f(x)$. Combined, these observations yield that $z = g(y) = g(f(x)) = (g \circ f)(x)$. □

Corollary 0.15.5. *Consider any pair of functions $f : X \to Y$ and $g : Y \to Z$. If f and g are bijective, then $g \circ f$ is bijective. Put another way, the composition of bijective functions is bijective.*

Proposition 0.15.6. *Consider any functions $f : W \to X$, $g : X \to Y$, and $h : Y \to Z$. We have that $h \circ (g \circ f) = (h \circ g) \circ f$. Put another way, composition of functions is associative.*

Proof. We must prove that $[h \circ (g \circ f)](w) = [(h \circ g) \circ f](w)$ for all elements $w \in W$ by Proposition 0.12.5. We will assume that $f(w) = x$, $g(x) = y$, and $h(y) = z$. By definition of the composite function, we have that $(g \circ f)(w) = g(f(w)) = g(x) = y$ and $(h \circ g)(x) = h(g(x)) = h(y) = z$ so that $[h \circ (g \circ f)](w) = h((g \circ f)(w)) = h(y) = z$ and $[(h \circ g) \circ f](w) = (h \circ g)(f(w)) = (h \circ g)(x) = z$. □

Remark 0.15.7. We note that in order to define the composition $g \circ f$ of any function $f : X \to Y$ under any other function $g : Y \to Z$, it is sufficient but not strictly necessary to assume that the domain of g is the codomain of f. Generally, the composite function $g \circ f$ is well-defined for any function $g : Y' \to Z$ so long as $Y' \supseteq \operatorname{range}(f)$. For instance, for the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$, we have that $\operatorname{range}(f) = \{f(x) \mid x \in \mathbb{R}\} = \{x^2 \mid x \in \mathbb{R}\} = \mathbb{R}_{\geq 0}$, hence for any function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$, the composition $g \circ f$ of f under g is well-defined. Explicitly, if we assume that $g(x) = \sqrt{x}$, then $(g \circ f)(x) = g(f(x)) = g(x^2) = \sqrt{x^2} = |x|$; however, if $g(x) = \ln(x)$ on its natural domain, then the composite function $g \circ f$ is not well-defined because $\ln(0)$ is not well-defined.

0.16 Inverse Functions

Considering that any function $f : X \to Y$ between two sets X and Y is by definition a relation, there exists an inverse relation f^{-1} from Y to X defined by $f^{-1} = \{(y, x) \mid (x, y) \in f\}$. One natural curiosity regarding the nature of the inverse relation f^{-1} of a function f is to ask whether the inverse relation f^{-1} of a function f must be a function. Certainly, the answer is no. One can readily verify that the relation $f = \{(-1, 1), (1, 1)\}$ on the set $X = \{-1, 1\}$ is a function, but its inverse relation $f^{-1} = \{(1, -1), (1, 1)\}$ is not a function because $f^{-1}(1)$ is not well-defined since $(1, -1)$ and $(1, 1)$ both lie in f^{-1}. Consequently, it seems that in order for the inverse relation f^{-1} of a function $f : X \to Y$ to be a function, we require that every element $f(x) \in \operatorname{range}(f)$ corresponds uniquely to an element $x \in X$. Put another way, we must have that f is injective. Conversely, by definition, if $f^{-1} : Y \to X$ is a function, then for every element $y \in Y$, we require that $f^{-1}(x)$ is an element of X. Explicitly, we require that for every element $y \in Y$, there exists an element $x \in X$ such that $y = f(x)$. Put another way, we must have that f is surjective. We are lead to the following.

Proposition 0.16.1. *Given any function $f : X \to Y$, the inverse relation f^{-1} is a function if and only if f is bijective. Even more, f^{-1} is bijective if and only if f is bijective.*

Proof. Observe that if f is bijective, then for every element $y \in Y$, there exists a unique element $x \in X$ such that $y = f(x)$. We may therefore define a relation f^{-1} from Y to X by declaring that $y \, f^{-1} \, x$ if and only if $x \, f \, y$, i.e., $y = f(x)$. Observe that f^{-1} is a function because for every element $y \in Y$, there exists one and only one element $x \in X$ such that $y = f(x)$ because f is bijective.

Conversely, suppose that $f^{-1} = \{(y, x) \mid (x, y) \in f\}$ is a function $f^{-1} : Y \to X$. By definition of a function, for every element $y \in Y$, there exists an element $x \in X$ such that $(y, x) \in f^{-1}$. But this implies that for every element $y \in Y$, there exists an element $x \in X$ such that $(x, y) \in f$ or $y = f(x)$, hence f is surjective. Even more, for every element $y \in Y$, the element $x \in X$ such that $y = f(x)$ is uniquely determined so that if $(y, x_1), (y, x_2) \in f^{-1}$, then $x_1 = x_2$. By definition of the inverse relation, we find that if $f(x_1) = f(x_2)$, then $x_1 = x_2$, hence f is injective, as desired.

Last, we will prove that $(f^{-1})^{-1} = f$, hence f^{-1} is bijective if and only if f is bijective (because its inverse relation is a function). By definition, we have that $(f^{-1})^{-1} = \{(x, y) \mid (y, x) \in f^{-1}\}$. Observe that if $(x, y) \in (f^{-1})^{-1}$, then $(y, x) \in f^{-1}$ yields that $(x, y) \in f$ and $(f^{-1})^{-1} \subseteq f$. Conversely, for any element $(x, y) \in f$, we have that $(y, x) \in f^{-1}$ so that $(x, y) \in (f^{-1})^{-1}$ and $(f^{-1})^{-1} \supseteq f$. □

Once we have identified that a function $f : X \to Y$ admits an inverse function $f^{-1} : Y \to X$, we seek an explicit definition of that inverse function. We achieve this via the following proposition.

Proposition 0.16.2. *Given any bijective function $f : X \to Y$, the inverse function $f^{-1} : Y \to X$ satisfies that $f^{-1} \circ f = \mathrm{id}_X$ and $f \circ f^{-1} = \mathrm{id}_Y$. Conversely, if $g : Y \to X$ is any function such that*

$g \circ f = \operatorname{id}_X$ and $f \circ g = \operatorname{id}_Y$, then $g = f^{-1}$. Put another way, the inverse function $f^{-1} : Y \to X$ of any bijection $f : X \to Y$ is the unique function $g : Y \to X$ such that $g \circ f = \operatorname{id}_X$ and $f \circ g = \operatorname{id}_Y$.

Proof. Consider any bijection $f : X \to Y$. By Proposition 0.16.1, the inverse relation $f^{-1} : Y \to X$ is a function. By definition of the inverse relation, we have that $f^{-1}(f(x)) = x = \operatorname{id}_X(x)$ for every element $x \in X$ so that $f^{-1} \circ f = \operatorname{id}_X$. Likewise, suppose that $f^{-1}(y) = x$. Considering that $f = (f^{-1})^{-1}$, it follows that $f(f^{-1}(y)) = y = \operatorname{id}_Y(y)$ for every element $y \in Y$ so that $f \circ f^{-1} = \operatorname{id}_Y$.

We will assume next that $g : Y \to X$ is any function such that $g \circ f = \operatorname{id}_X$ and $f \circ g = \operatorname{id}_Y$. Observe that $g = g \circ \operatorname{id}_Y = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = \operatorname{id}_X \circ f^{-1} = f^{-1}$ by Proposition 0.15.6. □

Example 0.16.3. We proved in Example 0.14.3 that the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = -x$ is bijective; its inverse function $f^{-1} : \mathbb{Z} \to \mathbb{Z}$ is defined by $f^{-1}(x) = -x$.

Example 0.16.4. We proved in Example 0.14.4 that the function $f : \mathbb{R} \setminus \{3\} \to \mathbb{R} \setminus \{1\}$ defined by

$$f(x) = \frac{x-2}{x-3}$$

is bijective. Observe that its inverse function is $f^{-1} : \mathbb{R} \setminus \{1\} \to \mathbb{R} \setminus \{3\}$ defined by

$$f^{-1}(x) = \frac{3x-2}{x-1}.$$

Remark 0.16.5. Generally, Proposition 0.16.2 provides an algorithm for determining the inverse function $f^{-1} : Y \to X$ of any function $f : X \to Y$ that can be defined by an explicit rule $y = f(x)$. Explicitly, we may solve the equation $y = f(x)$ in terms of x to find that $x = f^{-1}(y)$.

Example 0.16.6. Consider the function $f : \mathbb{R}_{\geq 0} \to \mathbb{R} \geq 0$ defined by $f(x) = x^2$. Observe that if $f(x_1) = f(x_2)$, then $x_1^2 = x_1^2$ yields that $(x_1 - x_2)(x_1 + x_2) = x_1^2 - x_2^2 = 0$. By the **Zero-Product Property** for real numbers, we conclude that either $x_1 - x_2 = 0$ or $x_1 + x_2 = 0$. Considering that $x_1, x_2 \geq 0$, the identity $x_1 + x_2 = 0$ holds if and only if $x_1 = x_2 = 0$; otherwise, we must have that $x_1 - x_2 = 0$ so that $x_1 = x_2$ and f is injective. Even more, for any real number $y \geq 0$, the real number $\sqrt{y}$ is well-defined and satisfies that $y = (\sqrt{y})^2 = f(\sqrt{y})$, hence f is onto; this can also be achieved by noticing that $y = f(x) = x^2$ if and only if $x = \sqrt{y}$. Consequently, we find that f is a bijective function with inverse $f^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by $f^{-1}(x) = \sqrt{x}$.

Currently, our strategy for computing the inverse function of a bijective function is somewhat backwards: in order to determine that the inverse relation of a function is a function, we must prove that the function is bijective. But this requires us to establish that the function is onto, and this necessitates the computation of the inverse function. We make the process more efficient as follows.

Proposition 0.16.7. *Consider any function $f : X \to Y$. If there exists a function $g : Y \to X$ such that $g \circ f = \operatorname{id}_X$ and $f \circ g = \operatorname{id}_Y$, then f and g are bijective functions satisfying that $g = f^{-1}$.*

Proof. We will prove that f is bijective. By Propositions 0.16.1 and 0.16.2, the result will follow. Consider any elements $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$. By hypothesis, we have that

$$x_1 = \operatorname{id}_X(x_1) = (g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2) = \operatorname{id}_X(x_2) = x_2.$$

We conclude that f is injective. Conversely, for every element $y \in Y$, we have that

$$y = \operatorname{id}_Y(y) = (f \circ g)(y) = f(g(y)).$$

Considering that $g(y) = x$ is an element of X, we conclude that $y = f(x)$ so that f is onto. □

Example 0.16.8. Consider the function $f : \mathbb{R} \to \mathbb{R}_{>0}$ defined by $f(x) = e^x$. By elementary calculus, we have that $f'(x) = e^x > 0$ for all real numbers x, hence $f(x)$ is one-to-one by Example 0.13.6. We know that the function $g : \mathbb{R}_{>0} \to \mathbb{R}$ defined by $g(x) = \ln(x)$ satisfies that

$$(g \circ f)(x) = g(f(x)) = \ln(e^x) = x \text{ for all real numbers } x \text{ and}$$
$$(f \circ g)(x) = f(g(x)) = e^{\ln(x)} = x \text{ for all real numbers } x > 0,$$

hence we conclude by Proposition 0.16.7 that f is bijective with inverse function $g = f^{-1}$.

Example 0.16.9. Consider the rational function $f : \mathbb{R} \setminus \{2\} \to \mathbb{R} \setminus \{1\}$ defined by

$$f(x) = \frac{2x+3}{2x-4}.$$

By the Quotient Rule, the derivative of $f(x)$ is the function $f' : \mathbb{R} \setminus 2 \to \mathbb{R} \setminus \{1\}$ defined by

$$f'(x) = \frac{2(2x-4) - 2(2x+3)}{(2x-4)^2} = -\frac{14}{(2x-4)^2}.$$

Considering that $(2x-4)^2 > 0$ for all real numbers $x \neq 2$, it follows that $f'(x) < 0$ for all real numbers $x \neq 2$ so that f is **decreasing**. By Example 0.13.6, it follows that f is one-to-one. (One can also use algebraic manipulation as in Example 0.14.4 to verify this.) We will next solve the equation $y = f(x)$ to find a function $x = f^{-1}(y)$, and we will verify that f^{-1} is the inverse of f.

$$y = f(x) = \frac{2x+3}{2x-4}$$

$$y(2x-4) = 2x+3$$

$$2xy - 4y = 2x+3$$

$$2xy - 2x = 4y+3$$

$$x(2y-2) = 4y+3$$

$$x = \frac{4y+3}{2y-2} = f^{-1}(y)$$

Consider the function $f^{-1} : \mathbb{R} \setminus \{1\} \to \mathbb{R} \setminus \{2\}$ defined by

$$f^{-1}(x) = \frac{4x+3}{2x-2}.$$

We will verify that $(f^{-1} \circ f)(x) = x$ for all real numbers $x \neq 2$ and $(f \circ f^{-1})(x) = x$ for all real numbers $x \neq 1$. By Proposition 0.16.7, we will conclude that f^{-1} is the inverse of f.

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = \frac{4f(x)+3}{2f(x)-2} = \frac{4 \cdot \frac{2x+3}{2x-4} + 3}{2 \cdot \frac{2x+3}{2x-4} - 2} = \frac{4(2x+3)+3(2x-4)}{2(2x+3)-2(2x-4)} = \frac{14x}{14} = x$$

$$(f \circ f^{-1})(x) = f(f^{-1}(x)) = \frac{2f^{-1}(x)+3}{2f^{-1}(x)-4} = \frac{2 \cdot \frac{4x+3}{2x-2} + 3}{2 \cdot \frac{4x+3}{2x-2} - 4} = \frac{2(4x+3)+3(2x-2)}{2(4x+3)-4(2x-2)} = \frac{14x}{14} = x$$

0.17 The Principle of Mathematical Induction

One of the most useful proof techniques is the **proof by induction** that appeals to one of the three incarnations of the **Principle of Mathematical Induction**. We say that a nonempty subset S of real numbers is **hereditary** if it holds that $x+1 \in S$ whenever we have that $x \in S$. Basically, the Principle of Mathematical Induction is a property of nonempty subsets of integers that asserts that if S is any hereditary subset of integers that admits a smallest element n_0 satisfying an open sentence $P(n)$ whose domain is the integers, then every element $n \in S$ satisfies the statement $P(n)$. Before we proceed to the definition of the Principle of Mathematical Induction, let us see some examples of properties of integers for which a proof by induction is appropriate.

Example 0.17.1. Consider the sum of the first n consecutive odd positive integers.

$$o(n) = 1 + 3 + 5 + \cdots + (2n-1) = \sum_{k=0}^{n-1} (2k+1)$$

Computing the value of $o(n)$ for the first four positive integers $1 \leq n \leq 4$ yields that $o(1) = 1$, $o(2) = 1+3 = 4$, $o(3) = 1+3+5 = 9$, $o(4) = 1+3+5+7 = 16$, and so on.

n	1	2	3	4
$o(n)$	1	4	9	16

Table 1: the sum of first n consecutive odd positive integers

Observe that $o(n) = n^2$ for each integer $1 \leq n \leq 5$. Continuing with the table, we would find that $o(n) = n^2$ for all integers $1 \leq n \leq k$ for any positive integer k. Consequently, we have the following.

Conjecture 0.17.2. We have that $o(n) = n^2$ for all integers $n \geq 1$ for $o(n)$ in Example 0.17.1.

Observe that $o(1) = 1 = 1^2$ and $o(n) + (2n+1) = o(n+1)$, hence if we could assume that $o(n) = n^2$, then we could conclude that $o(n+1) = n^2 + 2n + 1 = (n+1)^2$. We will soon return to validate this idea: it is precisely one of the tenants of the Principle of Mathematical Induction!

Example 0.17.3. Consider the sum of the first n consecutive positive integers.

$$c(n) = \sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n$$

Computing the value of $c(n)$ for the first four positive integers $1 \leq n \leq 4$ yields that $c(1) = 1$, $c(2) = 1+2 = 3$, $c(3) = 1+2+3 = 6$, and $c(4) = 1+2+3+4 = 10$, and so on.

n	1	2	3	4	5
$c(n)$	1	3	6	10	15

Table 2: the sum of the first n consecutive positive integers

Unfortunately, the pattern here is not obvious; however, due to a young Gauss, the following strategy can be employed. Briefly put, the idea is to write down the sum $1+2+3+\cdots+n$ forwards and backwards, adding each column of the sum to determine the value of $2(1+2+3+\cdots+n)$.

$$\begin{array}{cccccccccc} & 1 & + & 2 & + & 3 & + & \cdots & + & n \\ + & n & + & (n-1) & + & (n-2) & + & \cdots & + & 1 \\ \hline & (n+1) & + & (n+1) & + & (n+1) & + & \cdots & + & (n+1) \end{array}$$

Considering that there are n columns and the sum of each column is $n+1$, we conclude that $2(1+2+3+\cdots+n) = n(n+1)$. Consequently, we have the following.

Conjecture 0.17.4. We have that $c(n) = \frac{n(n+1)}{2}$ for all integers $n \geq 1$ for $c(n)$ in Example 0.17.3.

Like before, we can verify the formula for $n = 1$ as $c(1) = 1 = \frac{1 \cdot 2}{2}$ and $c(n) + (n+1) = c(n+1)$, hence if we could assume that $c(n) = \frac{n(n+1)}{2}$, then we could conclude that

$$c(n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2}.$$

Definition 0.17.5 (Principle of Ordinary Induction)**.** Let $P(n)$ be any open sentence defined for all integers $n \geq n_0$. If the following conditions hold, then $P(n)$ holds for all integers $n \geq n_0$.

(i.) $P(n_0)$ is a true statement.

(ii.) $P(n+1)$ is a true statement whenever $P(n)$ is a true statement for some integer $n \geq n_0$.

Remark 0.17.6. Be cognizant that we have taken the Principle of Ordinary Induction as an axiom in our set theory; however, some authors prefer to prove it as a corollary by first *defining* the non-negative integers $\mathbb{Z}_{\geq 0}$ as the intersection of all hereditary subsets of $\mathbb{R}$ that contain 0 (see [DW00, Definition 3.5]). Put another way, we may define $\mathbb{Z}_{\geq 0}$ as the intersection of all sets $S \subseteq \mathbb{R}$ such that

(a.) $0 \in S$ and

(b.) if $s \in S$, then $s+1 \in S$.

Using this axiom, the Principle of Ordinary Induction can be established by proving that the set $S = \{n \in \mathbb{Z}_{\geq 0} \mid P(n) \text{ is a true statement}\}$ is simply $\mathbb{Z}_{\geq 0}$. But this is clear: by definition of S, if $P(0)$ is a true statement, then $0 \in S$; likewise, if $n \in S$, then $P(n)$ is a true statement, hence $P(n+1)$ is a true statement, i.e., $n+1 \in S$. Combined, these observations illustrate that S is a hereditary subset of $\mathbb{R}$ that contains 0, i.e., $S \supseteq \mathbb{Z}_{\geq 0}$. By definition of S, we have also that $S \subseteq \mathbb{Z}_{\geq 0}$.

By the Principle of Ordinary Induction, we can return to prove Conjectures 0.17.2 and 0.17.4.

Proof. (Conjecture 0.17.2) Consider the following open sentence involving an integer $n \geq 1$.

$$P(n): \text{ We have that } 1+3+5+\cdots+(2n-1) = n^2.$$

We will prove that $P(n)$ is true for all integers $n \geq 1$, i.e., we will prove that "$\forall n \in \mathbb{Z}_{\geq 1}, P(n)$" is true. We proceed by the Principle of Ordinary Induction. We must verify the following conditions.

(i.) Observe that $P(1)$ is a true statement because it holds that $1 = 1^2$.

(ii.) We will assume that $P(n)$ is true for some integer $n \geq 1$. Consequently, we have that

$$1 + 3 + 5 + \cdots + (2n + 1) = 1 + 3 + 5 + \cdots + (2n - 1) + (2n + 1) = n^2 + 2n + 1 = (n + 1)^2.$$

Considering that (i.) $P(1)$ is a true statement and (ii.) $P(n + 1)$ is true whenever $P(n)$ is true for some integer $n \geq 1$, our proof is complete by the Principle of Ordinary Induction. □

Proof. (Conjecture 0.17.4) Consider the following open sentence involving an integer $n \geq 1$.

$$P(n): \text{ We have that } 1 + 2 + 3 + \cdots + n = \frac{n(n + 1)}{2}.$$

We will prove that $P(n)$ is true for all integers $n \geq 1$, i.e., we will prove that "$\forall n \in \mathbb{Z}_{\geq 1}, P(n)$" is true. We proceed by the Principle of Ordinary Induction. We must verify the following conditions.

(i.) Observe that $P(1)$ is a true statement because it holds that $1 = \frac{1 \cdot 2}{2}$.

(ii.) We will assume that $P(n)$ is true for some integer $n \geq 1$. Consequently, we have that

$$\begin{aligned} 1 + 2 + 3 + \cdots + (n + 1) &= 1 + 2 + 3 + \cdots + n + (n + 1) \\ &= \frac{n(n + 1)}{2} + (n + 1) \\ &= \frac{n(n + 1) + 2(n + 1)}{2} \\ &= \frac{(n + 1)(n + 2)}{2}. \end{aligned}$$

Considering that (i.) $P(1)$ is a true statement and (ii.) $P(n + 1)$ is true whenever $P(n)$ is true for some integer $n \geq 1$, our proof is complete by the Principle of Ordinary Induction. □

Going forward, we will begin any inductive proof by simply stating our intention to use a proof by induction; however, we will not typically make any explicit reference to the statement $P(n)$ that we intend to prove, and we will somewhat abbreviate the steps in the proof with the assumption that the reader is familiar with induction. We illustrate a typical proof by induction as follows.

Example 0.17.7. Prove that $2^n > n^2$ for all integers $n \geq 5$.

Proof. We proceed by induction. Observe that $2^5 = 32 > 25 = 5^2$, hence the claim holds for $n = 5$. We will assume inductively that $2^n > n^2$ for some integer $n \geq 5$. By hypothesis, we have that

$$2^{n+1} = 2 \cdot 2^n > 2n^2,$$

so it suffices to prove that $2n^2 \geq (n + 1)^2$. Considering that $n \geq 5$ by our inductive hypothesis, we have that $n^2 \geq 5n$ and $5n \geq 4n + 5 \geq 2n + 1$ so that

$$2n^2 = n^2 + n^2 \geq n^2 + 5n \geq n^2 + 2n + 1 = (n + 1)^2.$$

We conclude by induction that $2^n > n^2$ for all integers $n \geq 5$. □

Occasionally, it is desirable to strengthen the hypotheses of the Principle of Ordinary Induction in order to simplify proofs involving induction. Currently, we may view induction as a property of falling dominoes: (i.) if the n_0th domino falls and (ii.) the nth domino falling causes the $(n+1)$th domino to fall, then as the n_0th domino falls, all consecutive dominoes after it will fall. But suppose that we could knock down all dominoes from the n_0th to the nth domino: this would provide even more power with which to knock down the $(n+1)$th domino! We introduce this as the following.

Definition 0.17.8 (Principle of Complete Induction)**.** Let $P(n)$ be any open sentence defined for all integers $n \geq n_0$. If the following conditions hold, then $P(n)$ holds for all integers $n \geq n_0$.

(i.) $P(n_0)$ is a true statement.

(ii.) $P(n+1)$ is a true statement whenever $P(k)$ is a true statement for all integers $1 \leq k \leq n$.

Even though the hypotheses of the Principle of Complete Induction ostensibly appear to be stronger than the Principle of Ordinary Induction, the two principles are in fact equivalent to one another. Last, we obtain another ubiquitous tool that will prove crucial in our future endeavors.

Theorem 0.17.9 (Well-Ordering Principle)**.** *Every nonempty set of non-negative integers admits a smallest element with respect to the total order $\leq$. Put another way, if $S \subseteq \mathbb{Z}_{\geq 0}$ is a nonempty set, then there exists an element $s_0 \in S$ such that $s_0 \leq s$ for all elements $s \in S$.*

Proof. We will establish the contrapositive, i.e., we will prove that if $S \subseteq \mathbb{Z}_{\geq 0}$ has the property that for every element $s \in S$, there exists an element $s_0 \in S$ such that $s_0 < s$, then S must be empty. Let $P(n)$ be the statement that $n \notin S$. We claim that $P(n)$ holds for all integers $n \geq 0$. We proceed by the Principle of Complete Induction. Observe that if $0 \in S$, then there exists an element $s_0 \in S$ such that $s_0 < 0$. But this is not possible because S consists of non-negative integers. Consequently, we must have that $0 \notin S$, hence $P(0)$ is true. We will assume according to the Principle of Complete Induction that $P(k)$ holds for each integer $1 \leq k \leq n$. By definition, this means that $k \notin S$ for any integer $1 \leq k \leq n$. Observe that if $n+1 \in S$, then there exists an integer $s_0 \in S$ such that $1 \leq s_0 \leq n$. But this is not possible by the hypothesis of our induction. Consequently, we must have that $n+1 \notin S$, i.e., $P(n+1)$ is a true statement whenever $P(k)$ is a true statement for each integer $1 \leq k \leq n$. By the Principle of Complete Mathematical Induction, our proof is complete. □

Conversely, the Well-Ordering Principle implies the Principle of Ordinary Induction, hence it is equivalent to both ordinary induction and complete induction. (We leave the proof the reader.) Combined, the Principle of Ordinary Induction, the Principle of Complete Induction, and the Well-Ordering Principle constitute the triumvirate that is the Principle of Mathematical Induction.

0.18 Chapter 1 Overview

We recall that a **set** X is a collection of distinct objects called **elements** (or **members**) that often possess common properties. Each element of a set X is written as a lowercase x. If X possesses only finitely many elements $x_1, x_2, \dots, x_n$, then we may describe the set X using the **explicit notation** $X = \{x_1, x_2, \dots, x_n\}$. Often, it is most convenient to express a set X using **set-builder notation** $X = \{x \mid P(x)\}$ for some property $P(x)$ common to all elements $x \in X$. We assume the existence of a set $\varnothing$ that does not possess any elements; it is called the **empty set**. Every collection of sets admits certain operations that allow us to combine, compare, and take differences. Explicitly,

- the **union** of the sets X and Y is the set $X \cup Y = \{w \mid w \in X \text{ or } w \in Y\}$;
- the **intersection** of the sets X and Y is the set $X \cap Y = \{w \mid w \in X \text{ and } w \in Y\}$; and
- the **relative complement** of X with respect to Y is the set $Y \setminus X = \{w \in Y \mid w \notin X\}$.

We say that Y is a **subset** of X if every element of Y is an element of X, in which case we write $Y \subseteq X$; if Y is a subset of X and there exists an element of X that is not an element of Y, then Y is a **proper subset** of X, in which case we write $Y \subsetneq X$. Observe that Y is a (proper) subset of X if and only if $X \cap Y = Y$ (and $X \cup Y = X$). If $Y \subseteq X$ and $X \subseteq Y$, then $X = Y$; otherwise, the sets X and Y are distinct. One other way to distinguish a (finite) set X is by the number of elements X possesses, called the **cardinality** of X and denoted by $|X|$ or $\#X$ if the bars are ambiguous.

Conveniently, we may with large collections of sets by considering another set I as an **index set**; then, we denote by $\{X_i \mid i \in I\}$ the family of sets **indexed** by I. If each set X_i is a subset of some set U, we refer to U as our **universal set**. By definition, the union of the sets X_i is the set

$$\bigcup_{i \in I} X_i = \{u \mid u \in X_i \text{ for some element } i \in I\}$$

so that membership of an element $u \in U$ in this arbitrary union is characterized by $u \in \cup_{i \in I} X_i$ if and only if $u \in X_i$ for some index $i \in I$. Likewise, the arbitrary intersection of these sets is

$$\bigcap_{i \in I} X_i = \{u \mid u \in X_i \text{ for all elements } i \in I\}$$

with membership of an element $u \in U$ in the intersection characterized by $u \in \cap_{i \in I} X_i$ if and only if $u \in X_i$ for all indices $i \in I$. We say that two sets X_i and X_j are **disjoint** if $X_i \cap X_j = \varnothing$; if $X_i \cap X_j = \varnothing$ for all distinct indices $i, j \in I$, then the sets in $\{X_i \mid i \in I\}$ are **pairwise disjoint** or **mutually exclusive**. We say that $\mathcal{P} = \{X_i \mid i \in I\}$ forms a **partition** of the set U if and only if

(a.) X_i is nonempty for each index $i \in I$;

(b.) $U = \cup_{i \in I} X_i$; and

(c.) the sets X_i are pairwise disjoint (i.e., $X_i \cap X_j = \varnothing$ for every pair of distinct indices $i, j \in I$).

We define the **Cartesian product** of two sets X and Y to be the set consisting of all ordered pairs (x, y) such that $x \in X$ and $y \in Y$, i.e., $X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}$. Cardinality of finite sets X and Y is multiplicative in the sense that $|X \times Y| = |X| \cdot |Y|$. We refer to any subset R of the Cartesian product $X \times Y$ as a **relation** from the set X to the set Y. We say that an element $x \in X$ is **related to** an element $y \in Y$ under R if $(x, y) \in R$, and we write that $x \, R \, y$ in this case. Every relation $R \subseteq X \times Y$ induces a relation $R^{-1} \subseteq Y \times X$ called the **inverse relation** defined by

$$R^{-1} = \{(y, x) \mid (x, y) \in R\}.$$

If X is an arbitrary set, then a relation on X is a subset R of the Cartesian product $X \times X$. Every set X admits a relation called the **diagonal** (of X) and defined by $\Delta_X = \{(x, x) \mid x \in X\}$. We say that a relation R on X is **reflexive** if and only if $(x, x) \in R$ for all elements $x \in X$; **symmetric** if and only if $(x, y) \in R$ implies that $(y, x) \in R$; **antisymmetric** if and only if $(x, y) \in R$ and $(y, x) \in R$ implies that $x = y$; and **transitive** if and only if $(x, y) \in R$ and $(y, z) \in R$ together imply that $(x, z) \in R$. **Equivalence relations** are precisely the reflexive, symmetric, and transitive relations; **partial orders** are precisely the reflexive, antisymmetric, and transitive relations. Every equivalence relation E on X induces a partition of E via the **equivalence classes** of elements of X. Explicitly, we say that two elements $x, y \in X$ are **equivalent modulo** E if and only if $(x, y) \in E$, in which case we write that $x \; E \; y$; the equivalence class of an element $x \in X$ is the collection of elements $y \in X$ that are equivalent to x modulo E, i.e., the equivalence class of x is simply the set $[x] = \{y \in X \mid y \; E \; x\} = \{y \in X \mid (y, x) \in E\}$. Every element of a nonempty set X belongs to exactly one equivalence class of X modulo an equivalence relation E, hence X is partitioned by the set of distinct equivalence classes modulo E (see Proposition 0.10.2 and Corollary 0.10.3).

We define a **function** $f : X \to Y$ with **domain** X and **codomain** Y by declaring for each element $x \in X$ a unique (but not necessarily distinct) element $f(x) \in Y$. Every function $f : X \to Y$ induces a subset $f(V) = \{f(v) \mid v \in V\}$ of Y for every subset $V \subseteq X$ called the **direct image** of V (in Y) under f. Given any subset $W \subseteq Y$, the **inverse image** of W (in X) with respect to f is $f^{-1}(W) = \{x \in X \mid f(x) \in W\}$. We say that $f : X \to Y$ is **injective** if it holds that $f(x_1) = f(x_2)$ implies that $x_1 = x_2$ for any pair of elements $x_1, x_2 \in X$. On the other hand, if for every element $y \in Y$, there exists an element $x \in X$ such that $y = f(x)$, then $f : X \to Y$ is **surjective**. We say that a function $f : X \to Y$ is **bijective** provided that it is both injective and surjective.

Given any functions $f : X \to Y$ and $g : Y \to Z$, we may define a function $g \circ f : X \to Z$ called the **composite function** of f under g by declaring that $(g \circ f)(x) = g(f(x))$ for every element $x \in X$; the process of creating a composition function as **function composition**. Composition of functions is **associative**, i.e., $h \circ (g \circ f) = (h \circ g) \circ f$ whenever all composite functions are **well-defined**. Composition of functions preserves the property that two functions are injective or surjective. Every function $f : X \to Y$ is a relation from X to Y, hence there exists an inverse relation f^{-1} from Y to X; this inverse relation f^{-1} is a function if and only if f is bijective. Crucially, the **inverse function** $f^{-1} : Y \to X$ of a bijective function $f : X \to Y$ is the unique function satisfying that $f^{-1} \circ f = \mathrm{id}_X$ and $f \circ f^{-1} = \mathrm{id}_Y$ for the **identity function** $\mathrm{id}_X : X \to X$ defined by $\mathrm{id}_X(x) = x$.

Quite generally, if $f : X \to Y$ is an injective function, then the function $F : X \to \mathrm{range}(f)$ defined by $F(x) = f(x)$ is bijective. Consequently, there exists a function $F^{-1} : \mathrm{range}(f) \to X$ defined by $F^{-1}(y) = x$ for every element $y = f(x)$. Computing the inverse function F^{-1} corresponding to the induced function F amounts to solving the equation $y = F(x)$ in terms of x; the solution has the form $F^{-1}(y) = x$, and it is precisely this function F^{-1} that is the desired inverse function of F.

One of the most useful tools in mathematics is the **Principle of Mathematical Induction**. Collectively, the Principle of Mathematical Induction contains three equivalent statements: the Principle of Ordinary Induction, the Principle of Complete Induction, and the Well-Ordering Principle. Explicitly, the Principle of Ordinary Induction asserts that if $P(n)$ is any open sentence defined for all integers $n \geq n_0$ for some integer n_0 satisfying the properties that

1.) $P(n_0)$ is a true statement and

2.) $P(n+1)$ is a true statement whenever $P(n)$ is a true statement for some integer $n \geq n_0$,

then $P(n)$ is a true statement for all integers $n \geq n_0$. The Principle of Complete Induction asserts that if $P(n)$ is an open sentence defined for all integers $n \geq n_0$ for some integer n_0 and

1.) $P(n_0)$ is a true statement and

2.) $P(n+1)$ is true whenever $P(k)$ is true for all integers $n_0 \leq k \leq n$,

then $P(n)$ is a true statement for all non-negative integers n. One of the benefits of using complete induction is that its stronger hypotheses allow us more information with which to conveniently write proofs that might otherwise be awkward with ordinary induction. Even more, the Principle of Mathematical Induction appears also in the guise of the Well-Ordering Principle for the non-negative integers; this powerful tool guarantees that every nonempty set of non-negative integers admits a smallest element with respect to the total order $\leq$. Put another way, if $S \subseteq \mathbb{Z}_{\geq 0}$ is any nonempty set, then there exists an element $s_0 \in S$ such that $s_0 \leq s$ for all elements $s \in S$.

Chapter 1

Essential Properties of Real Numbers

Quite unsurprisingly, to undertake the study of real analysis, it is imperative to acquire a deeper understanding and knowledge of the real number system. Certainly, the reader is already acquainted with the real numbers in the context of computational mathematics such as calculus, linear algebra, and differential equations; however, many of the familiar properties of the reals are commonly taken for granted. Consequently, our aim throughout this chapter is to develop an axiomatic approach to working with real numbers. We will ultimately come to appreciate the real numbers as a complete totally ordered field that is also a metric space with respect to the usual absolute value function.

1.1 Order and Arithmetic of the Real Numbers

We will henceforth denote by the blackboard boldface symbol $\mathbb{R}$ the collection of all **real numbers**. We recall that the real number system consists of 0, all whole numbers, all negative whole numbers, all rational numbers (i.e., all signed fractions of whole numbers with nonzero denominator), and all irrational numbers. Concretely, the integers -1 and 17 are real numbers; the rational numbers $-\frac{1}{2}$ and $\frac{21}{29}$ are real numbers; and the irrational numbers e and π are real numbers; however, both $\sqrt{-1}$ and $\log(-1)$ are non-real **complex numbers**. Complex numbers can be constructed as a real vector space of dimension two; however, we will refrain from any further discussion here.

We begin our investigation of the real numbers in terms of their form and function. Crucially, the real numbers form an algebraic structure known as a **field** satisfying the following properties.

Definition 1.1.1 (Field Axioms for the Real Numbers)**.** Consider the familiar **binary operations** of addition $+$ and multiplication $\cdot$ of real numbers. Each of the following properties holds.

(1.) Given any real numbers a and b, we have that $a + b = b + a$ as real numbers.

(2.) Given any real numbers a, b, and c, we have that $(a + b) + c = a + (b + c)$ as real numbers.

(3.) We have that $0 + a = a$ for all real numbers a.

(4.) Given any real number a, there exists a real number $-a$ such that $a + (-a) = 0$.

(5.) Given any real numbers a and b, we have that $a \cdot b = b \cdot a$ as real numbers.

(6.) Given any real numbers a, b, and c, we have that $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ as real numbers.

(7.) We have that $1 \cdot a = a$ for all real numbers a.

(8.) Given any nonzero real number a, there exists a real number a^{-1} such that $a \cdot a^{-1} = 1$.

(9.) Given any real numbers a, b, and c, we have $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$.

We refer to properties (1.) and (5.) as the **commutative** properties of addition and multiplication; properties (2.) and (6.) are known as the **associative** properties of addition and multiplication; properties (3.) and (7.) illustrate the additive and multiplicative **identity elements** of the reals; properties (4.) and (8.) exhibit additive and multiplicative **inverses** of real numbers; and property (9.) provides a compatibility between addition and multiplication called the **distributive** property.

We will typically not use the dot symbol for multiplication unless its omission results in ambiguity; rather, we will simply concatenate real numbers so that the symbols $a \cdot b$ and ab both represent the product of the real numbers a and b. We demonstrate next that many (if not all) of the familiar properties of the real numbers can be obtained from the Field Axioms for the Real Numbers alone.

Theorem 1.1.2 (Common Properties of Real Numbers)**.** *Consider any real numbers a and b.*

1.) *If $a + b = b$, then $a = 0$. Consequently, the additive identity 0 is unique.*

2.) *If $ab = b$ and b is nonzero, then $a = 1$. Consequently, the multiplicative identity 1 is unique.*

3.) *If $ab = 1$ and b is nonzero, then $a = b^{-1}$. Consequently, multiplicative inverses are unique.*

4.) *If $ab = 0$, then $a = 0$ or $b = 0$. Consequently, the* **Zero Product Property** *holds.*

5.) *We have that $0a = 0$. Consequently, the* **Multiplication by Zero Property** *holds.*

Like usual, we may view the operations of **subtraction** and **division** in terms of addition and multiplication according to the formalisms $a - b = a + (-b)$ and $\frac{a}{b} = ab^{-1}$. Even more, we define **exponentiation** of real numbers by a positive exponent according to repeated multiplication: if a is any real number and n is a positive whole number, then $a^n = a \cdot a \cdots a$ is the n-fold product of a. Conversely, we establish the conventions that $a^0 = 1$ and $a^{-n} = (a^{-1})^n$ if a is nonzero.

We recall that any signed whole number is called an **integer**. Given any integers a and b such that b is nonzero, the fraction $\frac{a}{b}$ is a **rational number** that can be viewed as a real number ab^{-1}. Each integer a admits a notion of **parity** according to whether or not it is **divisible** by the integer 2: explicitly, we say that a is **even** if $\frac{a}{2}$ is an integer; otherwise, we say that a is **odd**. Consequently, an integer a is even if and only if there exists an integer k such that $a = 2k$. Likewise, an integer a is odd if and only if there exists an integer ℓ such that $a = 2\ell + 1$. (Why?) Parity of an integer is unique in the sense that every integer is either even or odd, and no integer is both even and odd.

Crucially, we may view the collection $\mathbb{Z}_{\geq 0}$ of non-negative whole numbers as a subset of the real numbers $\mathbb{R}$ by identifying each non-negative whole number n with the n-fold sum of the multiplicative identity 1, i.e., we have that $n = 1 + 1 + \cdots + 1$ with n summands. Consequently, under this convention, it follows that 0 is the empty sum (i.e., the sum with no summands). Considering that each non-negative integer n is a real number, the Field Axioms for the Real Numbers ensure the existence of a real number $-n$ such that $n + (-n) = 0$; thus, we define the collection of integers

$$\mathbb{Z} = \{n : n \in \mathbb{Z}_{\geq 0}\} \cup \{-n : n \in \mathbb{Z}_{\geq 0}\}.$$

Last, we obtain the rational numbers $\mathbb{Q}$ as quotients of integers whose denominator is nonzero

$$\mathbb{Q} = \left\{\frac{a}{b} : a, b \in \mathbb{Z} \text{ and } b \text{ is nonzero}\right\}.$$

Even though it is not immediate that there exist real numbers that are not rational, one can prove (e.g., by contradiction) that $\sqrt{2}$ is not rational: indeed, by definition, the square of the real number $\sqrt{2}$ is 2. Consequently, if $\sqrt{2}$ were a rational number, then there would exist relatively prime nonzero integers a and b such that $a^2 = 2b^2$. Checking the parity of a^2 and $2b^2$ shows that a and b are even — a contradiction. We conclude that $\mathbb{R} \setminus \mathbb{Q}$ is a nonempty subset of $\mathbb{R}$ called the **irrational numbers**.

Continuing our formal exposition on the real numbers, we turn our attention to the notion of **positivity** that may in turn be used to order the real numbers according to their "relative size."

Definition 1.1.3 (Positive Real Numbers)**.** Consider the set of real numbers $\mathbb{R}$. We define the set $\mathbb{R}_{>0}$ of **positive real numbers** according to the following three properties.

(1.) Given any real numbers $a, b \in \mathbb{R}_{>0}$, we have that $a + b \in \mathbb{R}_{>0}$.

(2.) Given any real numbers $a, b \in \mathbb{R}_{>0}$, we have that $ab \in \mathbb{R}_{>0}$.

(3.) (**Trichotomy Property**) Given any real number a, we have that $a \in \mathbb{R}_{>0}$ or $-a \in \mathbb{R}_{>0}$ or $a = 0$ and no pair of inclusions among these holds. Put another way, we have that

$$\mathbb{R} = \{a : a \in \mathbb{R}_{>0}\} \cup \{a : -a \in \mathbb{R}_{>0}\} \cup \{0\} \text{ is a partition of the real numbers.}$$

We will write that $a > 0$ if $a \in \mathbb{R}_{>0}$, and we will say in this case that a is a **positive** real number. Conversely, we will write that $a < 0$ if $-a \in \mathbb{R}_{>0}$, and we will say in this case that a is **negative**. We define the **non-negative** real numbers $\mathbb{R}_{\geq 0} = \mathbb{R}_{>0} \cup \{0\}$ as well as the **non-positive** real numbers $\mathbb{R}_{\leq 0} = \{a \in \mathbb{R} : a < 0\} \cup \{0\}$, and we extend the notation to include $a \geq 0$ or $a \leq 0$, respectively.

Definition 1.1.4 (Greater-Than and Less-Than)**.** Given any real numbers a and b, we say that a is **greater than** b (or that b is **less than** a) if $a - b > 0$ (or $a - b < 0$), and we write that $a > b$ (or $a < b$). Likewise, we say that a is **greater than or equal to** b (or that b is **less than or equal to** a) if $a - b \geq 0$ (or $a - b \leq 0$), and we write that $a \geq b$ (or $a \leq b$). Equality of real numbers then follows from the Trichotomy Property: indeed, we have that $a = b$ if and only if $a \geq b$ and $a \leq b$.

Each of the axioms of the positive real numbers can be used to formally prove the following.

Theorem 1.1.5 (Basic Properties of Real Inequalities)**.** *Consider any real numbers a, b, and c.*

1.) *If $a > b$ and $b > c$, then $a > c$. Consequently, greater-than is transitive.*

2.) *If $a > b$, then $a + c > b + c$. Consequently, greater-than is preserved under addition.*

3.) *If $a > b$ and $c > 0$, then $ca > cb$. Conversely, if $a > b$ and $c < 0$, then $ca < cb$. Consequently, greater-than is order-reversing with respect to sign under multiplication by a real number.*

4.) *We have that $a^2 > 0$. Consequently, it follows that $1 > 0$ and $n > 0$ for all $n \in \mathbb{Z}_{>0}$.*

Likewise, each of the inequalities in properties (1.), (2.), and (3.) extends analogously to less-than.

We provide next what is undoubtedly one of the most important features of the real numbers. Considering the utility of the below argument in future exposition, we provide a short proof.

Theorem 1.1.6. *Given any real number $a \geq 0$ such that $a < \varepsilon$ for all real numbers $\varepsilon > 0$, we have that $a = 0$. Consequently, there is no smallest positive real number.*

Proof. On the contrary, suppose that the real number $a > 0$ satisfies that $a < \varepsilon$ for all real numbers $\varepsilon > 0$. Consider the real number $\varepsilon_0 = \frac{1}{2}a$. We claim that $\varepsilon_0 > 0$ so that $a < \varepsilon_0$ by our assumption and $\varepsilon_0 < a$ by virtue of the fact that $a - \varepsilon_0 = \varepsilon_0 > 0$ — a contradiction. We must first note that $\frac{1}{2} > 0$ by the Trichotomy Property: indeed, we cannot have that $\frac{1}{2} < 0$ since this would imply that $-\frac{1}{2} > 0$ so that $-1 = -\frac{1}{2} + \left(-\frac{1}{2}\right) > 0 + \left(-\frac{1}{2}\right) > 0 + 0 = 0$ and $1 < 0$ — a contradiction. Likewise, we cannot have that $\frac{1}{2} = 0$ since this would imply that $1 = \frac{1}{2} + \frac{1}{2} = 0 + 0 = 0$ — a contradiction. We conclude that $\frac{1}{2} > 0$ so that $\varepsilon_0 = \frac{1}{2}a > 0$ by the third property listed under the Basic Properties of Real Inequalities. Our proof is therefore complete in view of our previous exposition. □

We conclude this section with one final familiar property of real inequalities that could best be summarized as asserting that greater-than and less-than are sign-determining under products.

Theorem 1.1.7. *Given any real numbers a and b such that $ab > 0$, it follows that a and b possess the same sign. Concretely, we have that $a > 0$ and $b > 0$ or $a < 0$ and $b < 0$. Conversely, if $ab < 0$, then a and b must have the opposite sign so that $a > 0$ and $b < 0$ or $a < 0$ and $b > 0$.*

1.2 Geometry of the Real Numbers

We have seen in the previous section that the real number system can be defined abstractly in terms of the Field Axioms for the Real Numbers. Carrying out formal manipulation of real numbers with respect to these axioms allows for the development of arithmetic of real numbers. We assume further the existence of positive real numbers $\mathbb{R}_{>0}$ in order obtain inequalities of real numbers.

Our aim throughout this section is to extend this abstraction to further axiomatize the properties of real numbers we have heretofore taken for granted. We begin with a discussion of the indispensable **absolute value** function. By the Trichotomy Property, every real number a satisfies that $a \geq 0$ or $-a \geq 0$. Essentially, the absolute value function always returns a non-negative real number.

Definition 1.2.1 (Absolute Value Function)**.** We define the **absolute value** of a real number a as

$$|a| = \begin{cases} a & \text{if } a \geq 0 \text{ and} \\ -a & \text{if } a < 0. \end{cases}$$

Consequently, the absolute value of a real number is a non-negative real number.

Theorem 1.2.2 (Basic Properties of the Absolute Value)**.** *Consider any real numbers a and b.*

1.) *We have that $|a| = 0$ if and only if $a = 0$. Consequently, absolute value is* **non-degenerate**.

2.) *We have that $|-a| = |a|$. Consequently, absolute value is not affected by sign.*

3.) *We have that $|ab| = |a||b|$. Consequently, absolute value distributes over multiplication.*

4.) *We have that* $|a|^2 = a^2$.

5.) *Given that* $b \geq 0$, *we have that* $|a| \leq b$ *if and only if* $-b \leq a \leq b$.

6.) *We have that* $-|a| \leq a \leq |a|$.

7.) (**Triangle Inequality**) *We have that* $|a + b| \leq |a| + |b|$.

Proof. (1.) Observe that if $a = 0$, then $|a| = 0$, so there is nothing to prove in this case. Conversely, if a is nonzero, then we must have that $|a| = a$ or $|a| = -a$ so that $|a|$ is likewise nonzero.

(2.) Certainly, the claim holds for $a = 0$ since it follows that $a = -a$ in this case. On the other hand, if $a > 0$, then $-a < 0$ so that $|-a| = -(-a) = a$. Likewise, if we assume that $a < 0$, then $-a > 0$ so that $|-a| = -a = |a|$ by definition of the absolute value function.

(3.) We leave the proof (by cases) as an exercise for the reader.

(4.) By the third property above, we have that $|a|^2 = |a||a| = |a^2| = a^2$ by Theorem 1.1.5(4.).

(5.) We note that if $-b \leq a \leq b$, then $a \leq b$ and $-a \leq b$ yield that $|a| \leq b$. Conversely, if $|a| \leq b$, then we must have that $a \leq b$ or $-a \leq b$. Concretely, if $a \geq 0$, then $-b \leq 0 \leq a \leq b$, as desired. Likewise, if we assume that $a < 0$, then $-a \leq b$ yields that $-b \leq a < 0 \leq b$.

(6.) We may appeal to the fifth property above by setting $b = |a|$.

(7.) By the sixth property above, we have that $-|a| \leq a \leq |a|$ and $-|b| \leq b \leq |b|$. By adding these inequalities, we find that $-(|a|+|b|) \leq a+b \leq |a|+|b|$. Last, use the fifth property above. □

We will soon visualize the real number system geometrically as a line in two dimensions, but for now, we make the following abstract definition of the distance between two real numbers.

Definition 1.2.3 (Distance Between Real Numbers)**.** Given any real numbers a and b, the **distance** between a and b is the real number $|a-b|$. Consequently, the real number $|a|$ represents the distance from the real number a to the real number 0. We will henceforth refer to 0 as the **origin**.

One immediate advantage of the previous definition is that it allows us to rigorously define the notion of "closeness" of two real numbers. Concretely, we would like to say that two real numbers a and b are "close" to one another provided that the distance $|a - b|$ between a and b is "small." Considering the utility of this notion in future exposition, we provide the following definition.

Definition 1.2.4 (ε-Neighborhood of a Point)**.** Given any real numbers a and $\varepsilon > 0$, the collection of real numbers x such that the distance $|x - a|$ between x and a does not exceed ε is referred to as the ε-**neighborhood** of the real number a. Explicitly, the ε-neighborhood of a is defined as

$$V_\varepsilon(a) = \{x \in \mathbb{R} : |x - a| < \varepsilon\} = \{x \in \mathbb{R} : a - \varepsilon < x < a + \varepsilon\}.$$

Crucially, the only real number that belongs to every ε-neighborhood of a real number a is a itself.

Theorem 1.2.5. *Consider any real number* a. *If* x *is any real number that lies in the* ε*-neighborhood* $V_\varepsilon(a)$ *for every real number* $\varepsilon > 0$, *then we must have that* $x = a$.

Proof. By Theorem 1.1.6, we conclude that $|x-a| = 0$: indeed, we have that $|x-a| < \varepsilon$ for any real number $\varepsilon > 0$ by assumption; then, the Basic Properties of the Absolute Value yield $x - a = 0$. □

Example 1.2.6. Consider the set $U = \{x \in \mathbb{R} : -1 < x < 1\}$ of real numbers whose absolute value does not exceed 1. Given any real number $a \in U$, we have that $-1 < a < 1$ so that $-1 < -a < 1$ and $0 < 1 - a < 2$. We conclude that $0 < \frac{1}{2}(1-a) < 1$. Likewise, we note that $0 < \frac{1}{2}(1+a) < 1$. Choosing $\varepsilon = \min\left\{\frac{1}{2}(1-a), \frac{1}{2}(1+a)\right\}$ yields an ε-neighborhood $V_\varepsilon(a)$ of a that is contained in U.

Example 1.2.7. Consider the set $U = \{x \in \mathbb{R} : -1 < x < 1\}$ from Example 1.2.6. We note that for any real number $\varepsilon > 0$, the ε-neighborhood of $a = 1$ must contain real numbers that do not lie in U. Explicitly, for any real number $\varepsilon > 0$, we have that $\frac{1}{2}\varepsilon > 0$ so that $1 + \frac{1}{2}\varepsilon > 1$. But this implies that $y = 1 + \frac{1}{2}\varepsilon$ lies in $V_\varepsilon(1) \setminus U$ since $y - 1 = \frac{1}{2}\varepsilon > 0$ yields that $|y-1| = \frac{1}{2}\varepsilon < \varepsilon$ and $y \in V_\varepsilon(1)$.

Example 1.2.8. Given any real numbers a, b, x, y, and $\varepsilon > 0$ such that $x \in V_\varepsilon(a)$ and $y \in V_\varepsilon(b)$, there exists a real number δ such that $x + y \in V_\delta(a+b)$. Explicitly, if we choose $\delta = 2\varepsilon$, then

$$|(x+y) - (a+b)| = |(x-a) + (y-b)| \leq |x-a| + |y-b| < \varepsilon + \varepsilon = \delta$$

by the Triangle Inequality. Essentially, this example illustrates that the sum of two elements of two ε-neighborhoods of some real numbers a and b lies in a (potentially larger) neighborhood of $a + b$.

References

[BS11] Robert G. Bartle and Donald R. Sherbert. *Introduction to Real Analysis*. 4th ed. John Wiley & Sons, Inc., 2011.

[DW00] J.P. D'Angelo and D.B. West. *Mathematical Thinking: Problem Solving and Proofs*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[RA15] Jon Rogawski and Colin Adams. *Calculus: Early Transcendentals*. 3rd ed. W.H. Freeman, 2015.

[Ste07] James Stewart. *Essential Calculus*. Brooks/Cole, Cengage Learning, 2007.

[Wad10] William R. Wade. *An Introduction to Analysis*. 4th ed. Prentice Hall, 2010.