MA291: Introduction to Higher Mathematics

Dylan C. Beck

Acknowledgements

Primarily, the contents of this document were created in the Fall 2022 and Spring 2023 semesters at Baker University with some mild revisions and reorganization efforts unfolding during the Spring 2024. I express my sincere gratitude to Vance Gaffar and all of my students in MA291 (Introduction to Higher Mathematics), but I am especially indebted to those who assisted in the enhancement of these notes with comments and suggestions, including Kyler Kosanke and April Thomas.

Contents

Chapter 1

Sets, Relations, 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.

1.1 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 $|$ 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 1.1.1. 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 1.1.2. Consider the finite set $T = \{-7, -5, -3, \ldots, 11, 13\}$. 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 1.1.3. 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 $U = \{0, 1, 4, 9, 16, \ldots\}$. 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 1.1.4. 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 $|D| = 0$. ◇

1.2 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 1.2.1 (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 1.2.2. 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 1.2.3. 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 1.14.1).

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.

1.3 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 1.3.1. Construct a **Venn diagram** to visualize the sets X, Y, $X \setminus Y$, and $X \cap Y$.

Example 1.3.2. Consider the sets $A = \{1, 2, 3, \ldots, 10\}$, $B = \{1, 4, 9\}$, 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 1.3.3 (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 1.3.4. 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 1.3.5. Consider the sets A, B, and C of Example 1.3.2. 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 1.3.6 (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 1.3.7. Consider the sets D, E, and F of Example 1.3.4. 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 1.3.8. 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 $\times$ corresponding respectively to whether an element of U is included or excluded, as follows.

$\times\times\times$: $\varnothing$ — ✓✓$\times$: $\{-1, 0\}$
✓$\times\times$: $\{-1\}$ — ✓$\times$✓: $\{-1, 1\}$
$\times$✓$\times$: $\{0\}$ — $\times$✓✓: $\{0, 1\}$
$\times\times$✓: $\{1\}$ — ✓✓✓: $\{-1, 0, 1\}$

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 1.3.9 (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 1.3.10. 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 $P(Y) = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$.

1.4 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

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

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 $X_1, X_2, \ldots, X_n$ using the index set $[n] = \{1, 2, \ldots, n\}$.

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

Example 1.4.1. Consider the sets $A_1 = \{1, 2\}, A_2 = \{2, 3\}, \ldots, A_{10} = \{10, 11\}$. Crucially, we may define $A_i = \{i, i + 1\}$ for each integer $1 \leq i \leq 10$. Using the index set $[10] = \{1, 2, \ldots, 10\}$ yields

$$\bigcap_{i=1}^{10} A_i = \{a \mid a \in A_i \text{ for each integer } 1 \leq i \leq 10\} = \varnothing,$$

$$\bigcap_{i=j}^{j+1} A_i = \{a \mid a \in A_j \text{ and } a \in A_{j+1}\} = \{j + 1\}, \text{ and}$$

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

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

$$\bigcup_{i=1}^{10} A_i = \{a \mid a \in A_i \text{ for some integer } 1 \leq i \leq 10\} = \{1, 2, \ldots, 11\},$$

$$\bigcup_{i=3}^{7} A_i = \{a \mid a \in A_i \text{ for some integer } 3 \leq i \leq 7\} = \{3, 4, \ldots, 8\}, \text{ and}$$

$$\bigcup_{i=j}^{k} A_i = \{a \mid a \in A_i \text{ for some integer } 1 \leq j \leq k \leq 10\} = \{j, j + 1, \ldots, k + 1\}.$$

Example 1.4.2. Consider the index set $L = \{a, b, c, \ldots, z\}$ 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 1.4.3. 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 1.4.4. 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 1.3.6). 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

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

1.5 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 1.5.1. 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 1.5.2. 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 sets 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 1.5.3. 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 1.5.4. Consider the sets $A = \{1, 4, 7\}$, $B = \{2, 5, 8\}$, and $C = \{3, 6, 9\}$ of Example 1.5.1. Considering that the sets A, B, and C are pairwise disjoint and $A \cup B \cup C = \{1, 2, \ldots, 9\} = [9]$, it follows that the set $\mathcal{P} = \{A, B, C\}$ constitutes a partition of the finite set $[9] = \{1, 2, \ldots, 9\}$.

Conversely, even though the nonempty sets $D = \{1, 3, 5, 7\}$, $E = \{2, 4, 6, 8\}$, and $F = \{3, 5, 7, 9\}$ of Example 1.5.2 satisfy $[9] = D \cup E \cup F$, they are not pairwise disjoint and do not partition $[9]$.

Example 1.5.5. 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 1.5.6. 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

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

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 $\{D_q\}_{q=1}^{\infty}$.

1.6 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 1.6.1. Consider the sets $X = \{-1, 1\}$ and $Y = \{1, 2, 3\}$. We have that

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

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

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

Example 1.6.2. 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 1.6.3. Given any real univariate 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 1.6.4 (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. □

1.7 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 1.7.1. Consider the sets $X = \{-1, 1\}$ and $Y = \{1, 2, 3\}$ of Example 1.6.1. 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

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

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

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

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 1.7.2. Consider the relation $R = \{(1, 1), (1, 2), (1, 3)\}$ from $X = \{-1, 1\}$ to $Y = \{1, 2, 3\}$ of Example 1.7.1. We have that $\text{dom}(R) = \{1\}$ and $\text{range}(R) = \{1, 2, 3\} = Y$.

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 1.7.3. Consider the relation $R = \{(1, 1), (1, 2), (1, 3)\}$ from $\{-1, 1\}$ to $\{1, 2, 3\}$ of Example 1.7.1. We have that $R^{-1} = \{(1, 1), (2, 1), (3, 1)\}$, $\text{dom}(R^{-1}) = \{1, 2, 3\}$, and $\text{range}(R^{-1}) = \{1\}$.

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

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

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.

1.8 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 1.8.1. Consider the following relations on the set $X = \{x, y, z\}$.

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

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 1.8.2. 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 1.8.3. 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 1.8.2, 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 1.8.4. 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.

1.9 Equivalence Relations

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

Example 1.9.1. Consider any set X. We may define a relation R on X by declaring that $x\ R\ y$ 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 1.9.2. *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 1.9.1. □

Example 1.9.3. 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 1.9.4. Consider the relation R defined on the set $\mathbb{Z}$ of integers such that $x\ R\ y$ 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 $x\ R\ y$ and $y\ R\ z$ together yield that $x\ R\ z$. We conclude that R is an equivalence relation on $\mathbb{Z}$.

Example 1.9.5. 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

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

Example 1.9.6. 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 1.9.7. Consider the equivalence relation R defined on the set $\mathcal{C}^1(\mathbb{R})$ of Example 1.9.3. 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 1.9.8. Consider the equivalence relation R of Example 1.9.4. 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 1.9.9. Consider the equivalence relation R of Example 1.9.5. 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 $[2] = \{2, 4\} = [4]$.

1.10 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 $[x] = \{y \in X \mid (y, x) \in E\}$. 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 1.10.1 (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 1.10.2 (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 1.10.1, we conclude that $[x] = [y]$, hence the proposed result is in fact established. □

Corollary 1.10.3 (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 1.10.2, 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 1.10.4. Consider the equivalence relation R of Example 1.9.5. By Corollary 1.10.3, the collection of distinct equivalence classes of $[5]$ modulo R provides a partition of $[5]$. By Example 1.9.9, 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 1.10.5. 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 1.10.3, 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 1.10.1, 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.

1.11 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 typically 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 $b - a$ or $n \mid (b - a)$. Conventionally, if a and b are congruent modulo n, we write $b \equiv a \pmod{n}$.

Example 1.11.1. We have that $7 \equiv 3 \pmod{4}$ because $7 - 3 = 4$ is divisible by 4.

Example 1.11.2. We have that $5 \equiv 21 \pmod{4}$ because $5 - 21 = -16 = 4(-4)$ is divisible by 4.

Example 1.11.3. We have that $11 \not\equiv 8 \pmod{4}$ because $11 - 8 = 3$ is not divisible by 4.

Given any nonzero integer n, we note that congruence modulo n induces a relation R_n on the set $\mathbb{Z}$ of integers: indeed, for any integers a and b, we have that $(a, b) \in R_n$ if and only if $a \; R_n \; b$ if and only if $b \equiv a \pmod{n}$ if and only if n divides $b - a$. Even more, the following proposition and Proposition 1.11.5 guarantee that the relation of congruence modulo n admits "nice" properties.

Proposition 1.11.4 (Properties of Congruence Modulo n)**.** *Consider any nonzero integer n and any integers a, b, and c. Each of the following properties of congruence modulo n holds.*

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

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

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

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

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

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

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

8.) (**Exponentiation Property**) *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 the definition and Identity 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 Multiplicative Property, if $b \equiv a \pmod{n}$, 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 Transitive 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 establishes the result. □

Given any nonzero integer n, the relation R_n defined on the set $\mathbb{Z}$ of integers such that $a \; R_n \; b$ if and only if $b \equiv a \pmod{n}$ is commonly referred to as **congruence modulo** n. By the third, fourth, and fifth properties of Proposition 1.11.4, congruence modulo n is an equivalence relation on $\mathbb{Z}$.

Proposition 1.11.5. *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 1.11.6. Congruence modulo 1 is an equivalence relation on $\mathbb{Z}$, hence we may seek to determine the equivalence classes of the integers modulo 1. Considering that every integer is divisible by 1, it follows that every pair of integers are related by congruence modulo 1: indeed, for any pair of integers a and b, we have that $b - a = 1 \cdot (b - a)$, hence a and b are congruent modulo 1. But this implies that every integer is congruent to 0 modulo 1, hence there is only one equivalence class of integers modulo 1. Explicitly, we have that $[0] = \{1q + 0 \mid q \in \mathbb{Z}\} = \{q \mid q \in \mathbb{Z}\} = \mathbb{Z}$.

Example 1.11.7. Congruence modulo 2 is an equivalence relation on $\mathbb{Z}$ with equivalence classes

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

By Proposition 1.10.2, these are all of the distinct equivalence classes of $\mathbb{Z}$ modulo 2. Even more, by Proposition 1.10.3, we obtain a partition of $\mathbb{Z}$ into distinct equivalence classes modulo 2

$$\mathbb{Z} = [0] \cup [1] = \{\ldots, -4, -2, 0, 2, 4, \ldots\} \cup \{\ldots, -3, -1, 1, 3, 5, \ldots\}.$$

Example 1.11.8. Congruence modulo 3 is an equivalence relation on $\mathbb{Z}$ with equivalence classes

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

By Proposition 1.10.2, these are all of the distinct equivalence classes of $\mathbb{Z}$ modulo 3. Even more, by Proposition 1.10.3, we obtain a partition of $\mathbb{Z}$ into distinct equivalence classes modulo 3

$$\mathbb{Z} = [0] \cup [1] \cup [2] = \{\ldots, -6, -3, 0, 3, 6, \ldots\} \cup \{\ldots, -5, -2, 1, 4, 7, \ldots\} \cup \{\ldots, -4, -1, 2, 5, 8, \ldots\}.$$

Each of the preceding examples is illustrative of the general structure of the equivalence classes of the integers modulo a nonzero integer n. Concretely, for any nonzero integer n, there are n distinct equivalence classes of the integers modulo n, and each class consists of sums of integer multiples of n and a non-negative integer that is strictly smaller than n. We remark that the proof of this fact follows by the Division Algorithm, hence we will not endeavor to provide such justification at the moment; however, the reader should consider how the result makes sense intuitively according to the process of integer division, quotients, remainders, and the definition of congruence modulo n.

Proposition 1.11.9. *Given any nonzero integer n, there are exactly n distinct equivalence classes of $\mathbb{Z}$ modulo n defined by $[r] = \{nq + r \mid q \in \mathbb{Z}\}$ for each integer $0 \leq r \leq n-1$. Consequently, every nonzero integer n induces a partition of the integers into distinct equivalence classes modulo n*

$$\mathbb{Z} = \bigcup_{r=0}^{n-1} \{nq + r \mid q \in \mathbb{Z}\}.$$

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

Example 1.11.10. Consider the relation R defined on the set $\mathbb{Z}$ of integers such that $a \; R \; b$ if and only if $5b \equiv 2a \pmod 3$ for any integers a and b. We claim that R is an equivalence relation.

1.) We must first establish that $a \; R \; a$ for all integers a. By definition of R, 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 of R, 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 $a \; R \; b$ and $b \; R \; c$, 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 $a \; R \; c$.

By definition of an equivalence class of $\mathbb{Z}$ modulo R, the equivalence class of a modulo R is simply

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

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

1.12 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 will distinguish a relation f from X to Y as a **function** if and only if every element of X is the first coordinate 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 $\text{range}(f) = \{f(x) \mid x \in X\}$ 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 direct image $y \in Y$ of the element $x \in X$.

Example 1.12.1. Consider the relation $f = \{(-1, 1), (1, -1)\}$ defined on the set $X = \{-1, 1\}$. Each of the elements of X is the first coordinate 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 1.12.2. Consider the relation $g = \{(x, x-1) \mid x \in \mathbb{R}\}$ on the collection $\mathbb{R}$ of real numbers. Every real number is the first coordinate 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 1.12.3. 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, in general, a function depends intimately 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, all 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 1.12.4. Consider the equivalence relation R defined on the set $\{1, 2, 3, 4, 5\}$ as in Example 1.9.5. Crucially, we note that R is not a function since the ordered pairs $(1, 1)$ and $(1, 3)$ lie in R. 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 admits an **identity function** $\text{id}_X : X \to X$ defined by $\text{id}_X(x) = x$. 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 while 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 the direct 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; this establishes the following important fact about functions.

Proposition 1.12.5 (Equality of Functions)**.** *Given any sets X and Y, the 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 $x \in X$.*

Every 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 $W \subseteq Y$, 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 1.12.6. 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 $\text{range}(f) = \{1, 2, 3\}$, but it is just as true that $\text{range}(f) = f(\{u, v, w\}) = f(\{u, x, y\}) = f(\{x, y, z\})$. 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 1.12.7. 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 $g([-1, 1]) = [0, 1]$. Likewise, if $x^2 = g(x) \geq 4$, then $x \geq 2$ or $x \leq -2$ so that $g^{-1}([4, \infty)) = (-\infty, -2] \cup [2, \infty)$.

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 1.12.8. Consider the sets $X = \{u, v, w, x, y, z\}$ and $Y = [6] = \{1, 2, 3, 4, 5, 6\}$ of Example 1.12.6. We have that $|X| = 6 = |Y|$, hence there are $|Y|^{|X|} = 6^6$ possible functions $f : X \to Y$.

1.13 One-to-One and Onto Functions

We introduce in this section 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 $x_1, x_2 \in X$ induces 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 1.13.1. 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 1.13.2. 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 1.13.3. 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 1.13.4. 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 1.13.3.

Example 1.13.5. 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 $f(w) = 3 = f(x)$.

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 $Y = \text{range}(f) = \{f(x) \mid x \in X\}$.

Example 1.13.6. 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 1.13.7. Consider the real function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 4$ of Example 1.13.2. We claim that f is onto, hence for any real number y, we require a real number x such that $y = f(x) = 3x + 4$. By solving for x in $y = 3x + 4$, we find 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 1.13.8. 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 1.13.9. 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 $Y = [6] = \{1, 2, 3, 4, 5, 6\}$ as in Example 1.13.5. 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 1.13.10. 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 $x \in X$.

1.14 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 rather than a "bijection" from X to Y.

Proposition 1.14.1. *Consider any pair of arbitrary finite sets X and Y.*

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

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

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

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

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

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

7.) *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 also for notational convenience that $X = \{x_1, x_2, \ldots, x_m\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$. We turn our attention to each claim in turn.

(1.) We will assume that there exists an injective function $f : X \to Y$. Consequently, every element $y \in Y$ is obtained from 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 $|X| \leq |Y|$.

(2.) 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 $n \geq m$.

(3.) 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 $|X| \geq |Y|$.

(4.) 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 $y_i = f(x_i)$.

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

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

(7.) 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 with 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 $y = f(x)$ must be 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.

Example 1.14.2. Consider the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = -x$. Cancelling a minus sign, we conclude that 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 1.14.3. Consider the rational 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 identity 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.

Example 1.14.4. Consider the equivalence relation R_6 of congruence modulo 6 defined on the set $\mathbb{Z}$ of integers. Conventionally, the collection of equivalence classes of $\mathbb{Z}$ modulo 6 is denoted $\mathbb{Z}/6\mathbb{Z}$. Every element of $\mathbb{Z}/6\mathbb{Z}$ is the equivalence class $[a] = \{b \in \mathbb{Z} \mid b \equiv a \pmod 6\}$ of an integer a modulo 6, hence there are six distinct elements of $\mathbb{Z}/6\mathbb{Z}$ by Proposition 1.11.5. We will demonstrate in this example how to define a function from $\mathbb{Z}/6\mathbb{Z}$ to itself. We may define a relation $f : \mathbb{Z}/6\mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$ by declaring that $f([x]) = [5x+3]$. By definition, in order to establish that f is a function, we must verify that if $[5x+3]$ and $[5y+3]$ are distinct, then $[x]$ and $[y]$ are distinct. Concretely, this ensures that the function f passes the Vertical Line Test. Consequently, we may assume that $[x] = [y]$ and derive $[5x+3] = [5y+3]$. (Why?) By Equality of Equivalence Classes, it follows that $y \equiv x \pmod 6$ so that 6 divides $y - x$ by the Properties of Congruence Modulo n. By definition of divides, there exists an integer k such that $y - x = 6k$; in turn, this yields the divisibility relation

$$(5y + 3) - (5x + 3) = 5(y - x) = 6(5k).$$

Considering that $5k$ is an integer, we conclude that 6 divides the integer $(5y+3) - (5x+3)$ so that $5y+3 \equiv 5x+3 \pmod 6$. Once again, by Proposition 1.10.1, we conclude that $[5x+3] = [5y+3]$. We say in this case that the relation f is a **well-defined** function. Quite to our delight, it happens that f is a bijection: indeed, we have that $[0] = f([3])$, $[1] = f([2])$, $[2] = f([1])$, $[3] = f([0])$, $[4] = f([5])$, and $[5] = f([4])$ so that f is both injective and surjective. (One can prove this algebraically.)

1.15 Composition of Functions

Every pair of functions $f : X \to Y$ and $g : Y \to Z$ from any three sets X, Y, and Z give rise to a third 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 1.15.1. Consider the sets $X = \{-1, 1\}$, $Y = \{x, y, z\}$, and $Z = \{1, 2, 3\}$. We may define a pair of 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 1.15.2. 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

$$(g \circ f)(a) = g(f(a)) = g(b) = c, \qquad (g \circ f)(c) = g(f(c)) = g(d) = e, \text{ and}$$
$$(g \circ f)(b) = g(f(b)) = g(c) = d, \qquad (g \circ f)(d) = g(f(d)) = g(e) = f.$$

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

Example 1.15.3. Composition of functions ought to be a familiar concept from calculus: indeed, 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

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

Crucially, the latter holds because $e^x > 0$ for all real numbers x, hence it follows that $g \circ f = f$ as real functions. On the other hand, for the real identity function $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ defined by $\mathrm{id}_{\mathbb{R}}(x) = x$, we note that $\mathrm{id}_{\mathbb{R}} \circ f = f$ since $(\mathrm{id}_{\mathbb{R}} \circ f)(x) = \mathrm{id}_{\mathbb{R}}(f(x)) = f(x)$ for all real numbers x. Comparing the two identities derived in this example yields that $g \circ f = \mathrm{id}_{\mathbb{R}} \circ f$; however, it is not the case that $g = \mathrm{id}_{\mathbb{R}}$ because $g(-1) = 1 \neq -1 = \mathrm{id}_{\mathbb{R}}(-1)$. Consequently, we obtain the following important fact.

Proposition 1.15.4 (Function Composition Is Not Cancellative)**.** *Given any quadruple of functions $f : X \to Y$, $g : X \to Y$, $h : Y \to Z$, and $j : Y \to Z$ such that $h \circ f = j \circ f$ and $h \circ f = h \circ g$, we cannot conclude that either $h = j$ or $f = g$. Put another way, function composition is not cancellative.*

Proof. We leave it to the reader to adapt the approach of Example 1.15.3 to determine sets X, Y, and Z and distinct functions $f : X \to Y$, $g : X \to Y$, and $h : Y \to Z$ such that $h \circ f = h \circ g$. □

Even though function composition is not cancellative, we will soon come to find that it satisfies several important properties that make it an indispensable operation in the theory of functions.

Proposition 1.15.5 (Composition of Functions Preserves Injectivity and Surjectivity)**.** *Consider any pair of functions $f : X \to Y$ and $g : Y \to Z$.*

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

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

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

Proof. (1.) 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.

(2.) We must prove that for every element $z \in Z$, we have that $z = (g \circ f)(x)$ for some element $x \in X$. 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 1.15.6 (Composition of Bijective Functions Is Bijective)**.** *Given any bijective functions $f : X \to Y$ and $g : Y \to Z$, the composite function $g \circ f : X \to Z$ is bijective.*

Proof. Both f and g are injective and surjective, so $g \circ f$ is injective and surjective. □

Proposition 1.15.7 (Composition of Functions Is Associative)**.** *Consider any triple of functions $f : W \to X$, $g : X \to Y$, and $h : Y \to Z$. We have that $h \circ (g \circ f) = (h \circ g) \circ f$.*

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 1.12.5. 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 1.15.8. 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 contains the codomain of f. Generally, the composite function $g \circ f$ is well-defined for any function $g : W \to Z$ so long as $W \supseteq \text{range}(f)$. Consider to this end the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$, we have that $\text{range}(f) = \{f(x) \mid x \in \mathbb{R}\} = \{x^2 \mid x \in \mathbb{R}\} = \mathbb{R}_{\geq 0}$, hence for any function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$, the composite function $g \circ f$ 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.

Proposition 1.15.9 (Composition of Functions Is Not Commutative)**.** *Given any pair of functions $f : X \to Y$ and $g : Y \to Z$, we cannot conclude that $g \circ f = f \circ g$.*

Proof. By the preceding remark, we must have that $Y \supseteq \text{range}(f)$, so if this is not the case, then $g \circ f$ is not well-defined. We may assume therefore that $Y \supseteq \text{range}(f)$ and $X \supseteq \text{range}(g)$ so that $g \circ f$ and $f \circ g$ are both well-defined. Consider the real functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x + 1$ and $g(x) = 2x - 1$. Certainly, the reader can verify that f and g are both bijective functions (indeed, the graphs of f and g are lines of slope 2), hence the condition that the domain of g contains the codomain of f and vice-versa hold. Likewise, it is simple to check that

$$(f \circ g)(x) = f(g(x)) = f(2x - 1) = 2(2x - 1) + 1 = 4x - 1 \text{ and}$$
$$(g \circ f)(x) = g(f(x)) = g(2x + 1) = 2(2x + 1) - 1 = 4x + 1.$$

Considering that $(f \circ g)(0) = -1$ and $(g \circ f)(0) = 1$ are not equal, $f \circ g$ and $g \circ f$ are not equal. □

1.16 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. Generally, the answer is no.

Proposition 1.16.1 (The Inverse Relation of a Function Is Not Necessarily a Function)**.** *Given any function $f : X \to Y$, the inverse relation $f^{-1} : Y \to X$ is not necessarily a function.*

Proof. Consider the relation $f = \{(-1, 1), (1, 1)\}$ on the set $X = \{-1, 1\}$. We leave it to the reader to verify that f is a function with inverse relation $f^{-1} = \{(1, -1), (1, 1)\}$. We note that f^{-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 would appear 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 $f(x) \in \text{range}(f)$ 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, it must be the case 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 therefore lead to the following result.

Theorem 1.16.2 (Existence of an Inverse Function)**.** *Given any function $f : X \to Y$, the inverse relation f^{-1} is a function if and only if f is bijective. Even more, if f^{-1} is a function, it is bijective.*

Proof. Observe that if f is a bijective function, then for every element $y \in Y$, there exists a unique element $x \in X$ such that $y = f(x)$. Consequently, the inverse relation $f^{-1} : Y \to X$ of f defined by $(y, x) \in f^{-1}$ if and only if $y = f(x)$ is a function because for every element $y \in Y$, there exists one and only one element $x \in X$ such that $y = f(x)$ by hypothesis that f is a bijection. Concretely, for any pair of elements $(y, x_1), (y, x_2) \in f^{-1}$, we have that $f(x_1) = y = f(x_2)$ so that $x_1 = x_2$ since f is injective, and every element of Y is mapped onto an element of X by f^{-1} since f is surjective.

Conversely, suppose that the inverse relation $f^{-1} = \{(y, x) \mid (x, y) \in f\}$ of f is a function. By definition of a function, for every element $y \in Y$, there exists an element $x \in X$ such that $(y, x) \in f^{-1}$ so that $y = f(x)$. But this implies that f is surjective since every element of Y is the image of some element of X. Even more, for every element $y \in Y$, there exists a unique element $x \in X$ such that $(y, x) \in f^{-1}$ or $y = f(x)$; thus, 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.

Last, we will demonstrate that if f^{-1} is a function, then f^{-1} is bijective. We will assume first that $f^{-1}(y_1) = f^{-1}(y_2)$. By the previous paragraph, if f^{-1} is a function, then f is surjective, hence there exist elements $x_1, x_2 \in X$ such that $y_1 = f(x_1)$ and $y_2 = f(x_2)$. By definition of the inverse relation, if $f^{-1}(y_1) = f^{-1}(y_2)$, then $x_1 = x_2$ so that $y_1 = f(x_1) = f(x_2) = y_2$. Even more, for every element $x \in X$, there exists one and only one element $y \in Y$ such that $y = f(x)$ since f is bijective. Consequently, for every element $x \in X$, there exists an element $y \in Y$ such that $x = f^{-1}(y)$. □

Remark 1.16.3 (Construction of Inverse Functions)**.** By the previous theorem, the inverse relation f^{-1} of a function $f : X \to Y$ is a function if and only if f is bijective. We demonstrate next that if f is injective but not surjective, it is possible to construct an inverse function related to f. Crucially,

every function $f : X \to Y$ **restricts** to a surjective function $F : X \to f(X)$ defined by $F(x) = f(x)$ with the same domain as f but whose codomain is the range of f. Consequently, if $f : X \to Y$ is injective, then $F : X \to f(X)$ is bijective, hence the inverse relation $F^{-1} : f(X) \to X$ is a function. Conversely, even if $f : X \to Y$ is not injective, we may modify the domain of f to obtain a bijection. Consider the set X_f of $x \in X$ such that for every pair of elements $x_1, x_2 \in X_f$, we have that $x_1 = x_2$ if $f(x_1) = f(x_2)$. One can verify that $\tilde{f} : X_f \to f(X_f)$ defined by $\tilde{f}(x) = f(x)$ is bijective.

We illustrate these concepts in the following examples. Consider the real function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ defined by $f(x) = \sqrt{x}$. Observe that $\sqrt{x} \geq 0$ for all real numbers $x \geq 0$, hence f is not surjective: indeed, we have that $f(\mathbb{R}_{\geq 0}) = \mathbb{R}_{\geq 0}$. Consequently, the induced function $F : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by $F(x) = \sqrt{x}$ is a bijection. Likewise, the real function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2$ is neither injective nor surjective since $g(-1) = 1 = g(1)$ and $x^2 \geq 0$ implies that $g(\mathbb{R}) = \mathbb{R}_{\geq 0}$. On the other hand, the induced function $\tilde{g} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by $\tilde{g}(x) = x^2$ is a bijection: indeed, for any real numbers $x_1^2 = x_2^2$ such that $x_1, x_2 \geq 0$, we must have that $x_1 = x_2$ since $-x_1, -x_2 \leq 0$.

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 1.16.4 (Construction and Uniqueness of Inverse Functions)**.** *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 $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$, then we must have that $g = f^{-1}$. Put another way, the inverse function $f^{-1} : Y \to X$ corresponding to any bijective function $f : X \to Y$ is the unique function $g : Y \to X$ satisfying that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.*

Proof. Given any bijective function $f : X \to Y$, the inverse relation $f^{-1} : Y \to X$ is a function by Theorem 1.16.2. By definition of f^{-1}, we have that $f^{-1}(f(x)) = x = \mathrm{id}_X(x)$ for every element $x \in X$ so that $f^{-1} \circ f = \mathrm{id}_X$ by Proposition 1.12.5. Likewise, we have that $f(f^{-1}(y)) = y = \mathrm{id}_Y(y)$ for every element $y \in Y$ so that $f \circ f^{-1} = \mathrm{id}_Y$. We will assume next that $g : Y \to X$ is any function such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. By Proposition 1.15.7, we have that

$$g(y) = (g \circ \mathrm{id}_Y)(y) = [g \circ (f \circ f^{-1})](y) = [(g \circ f) \circ f^{-1}](y) = (\mathrm{id}_X \circ f^{-1})(y) = f^{-1}(y)$$

for every element $y \in Y$. We leave it to the reader to verify that if $f \circ g = \mathrm{id}_Y$, then $g = f^{-1}$. □

Example 1.16.5. We proved in Example 1.14.2 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 1.16.6. We proved in Example 1.14.3 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 1.16.7. Generally, Proposition 1.16.4 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 1.16.8. Consider the rational function $f : \mathbb{R} \setminus \{2\} \to \mathbb{R} \setminus \{1\}$ defined by

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

We may solve the equation $y = f(x)$ to find a function $x = f^{-1}(y)$ that is the inverse of f.

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

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

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

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

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

Consequently, we obtain a rational 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 1.16.9, we will conclude that f^{-1} is the inverse of f.

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

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

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 a bijection. 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 1.16.9. *Given any function $f : X \to Y$ such that there exists a function $g : Y \to X$ for which $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$, it follows that f and g are bijections satisfying that $g = f^{-1}$.*

Proof. We will prove only that f is bijective. By Propositions 1.16.2 and 1.16.4, the result follows. Consider any elements $x_1, x_2 \in X$ such that $f(x_1) = f(x_2)$. By hypothesis, we have that

$$x_1 = \mathrm{id}_X(x_1) = (g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2) = \mathrm{id}_X(x_2) = x_2,$$

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

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

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

1.17 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 $x_1, x_2, \ldots, x_n$, then we may describe the set X using the **explicit notation** $X = \{x_1, x_2, \ldots, x_n\}$. 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 $Y \setminus X = \{w \mid w \in Y \text{ and } w \notin X\}$.

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 $Y \subsetneq X$. By the Going-Down Property of Set Intersection or the Going-Up Property of Set Union, we have that Y is a subset of X if and only if $X \cap Y = Y$ if and only if $X \cup Y = X$. If $Y \subseteq X$ and $X \subseteq Y$, then $X = Y$; otherwise, the sets X and Y are not equal. 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 work with large collections of sets by introducing an **index set** I. Concretely, we may 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 a **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 any set, a **relation on** X is a subset R of the Cartesian product $X \times X$. Every set X admits a relation $\Delta_X = \{(x, x) \mid x \in X\}$ called the **diagonal**. 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 those relations that are reflexive, symmetric, and transitive; **partial orders** are precisely those relations that are reflexive, antisymmetric, and transitive. 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 $x \; E \; y$; thus, 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

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

Every element of a nonempty set X belongs to one and only one equivalence class of X modulo an equivalence relation E, hence the distinct equivalence classes of X modulo E partition X.

We may 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 all $x \in X$; the construction of composite functions is an operation known as **function composition**. Composition of functions is **associative** so that $h \circ (g \circ f) = (h \circ g) \circ f$ whenever each of these composite functions is **well-defined**; however, function composition is neither cancellative nor commutative. Concretely, we cannot conclude that $f = g$ simply because $h \circ f = h \circ g$; we cannot conclude that $h = j$ simply because $h \circ g = j \circ g$; and it is not necessarily the case that $f \circ g = g \circ f$. Composition of

functions preserves the injectivity and surjectivity of functions, so it preserves bijections, as well. 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 $\mathrm{id}_X(x) = x$.

Generally, if $f : X \to Y$ is an injective function, then the induced function $F : X \to f(X)$ defined by $F(x) = f(x)$ is bijective. Consequently, there exists a function $F^{-1} : f(X) \to X$ 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 the function F^{-1} that is the inverse function of F.

1.18 Chapter 1 Exercises

Exercise 1.18.1. Express each of the following sets in set-builder notation.

(a.) $S = \{1, 4, 7, 10\}$

(b.) $T = \{-5, -4, -3, 3, 4, 5\}$

(c.) $U = \{-19, -18, \ldots, -4, 4, 5, \ldots, 19\}$

(d.) $V = \{2, 3, 5, 7, 11, 13, 17, 19, \ldots\}$

(e.) $W = \{\ldots, -3, -1, 1, 3, \ldots\}$

(f.) $X = \left\{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots\right\}$

(g.) $Y = \left\{\frac{1}{9}, -\frac{1}{3}, 1, -3, 9, \ldots\right\}$

(h.) $Z = \{\ldots, -2\pi, -\pi, 0, \pi, 2\pi, \ldots\}$

Exercise 1.18.2. Express each of the following sets in explicit notation.

(a.) $S = \left\{s \in \mathbb{R} \mid s^2 + \frac{4}{3}s + \frac{1}{3} = 0\right\}$

(b.) $T = \{t \in \mathbb{R} \mid \tan(t) = 0\}$

(c.) $U = \left\{u \in \mathbb{R} : \frac{d}{du}\sqrt{u^2 + 1} = 0\right\}$

(d.) $V = \{v \in \mathbb{N} \mid v^2 + 1 = 26\}$

(e.) $W = \{w \in \mathbb{Z} : w \text{ is odd and } |w| < 10\}$

(f.) $X = \{x \in \mathbb{R} \mid x^3 - 6x^2 + 11x - 6 = 0\}$

(g.) $Y = \{y \in \mathbb{R} \mid y^4 + 3 = 0\}$

(h.) $Z = \left\{z \in \mathbb{R} : \lim_{x \to z} \frac{x^2}{x^4 - 2x^2 + 1} = \infty\right\}$

Exercise 1.18.3. Consider the set $U = \{1, 2, 3, 4, 5\}$ with subsets A and B such that

(i.) $|A| = |B| = 3$;

(ii.) 1 lies in A but does not lie in B;

(iii.) 2 lies in B but does not lie in A;

(iv.) 3 lies in either A or B but not both;

(v.) 4 lies in either A or B but not both; and

(vi.) 5 lies in either A or B but not both.

List all possibilities for A in curly brace notation; then, determine the corresponding sets B.

Exercise 1.18.4. Consider the following sets.

$$W = \{1, 2, 3, \ldots, 10\} \qquad \mathbb{E} = \{n \mid n \text{ is an even integer}\}$$
$$X = \{1, 3, 5, 7, 9\} \qquad \mathbb{O} = \{n \mid n \text{ is an odd integer}\}$$
$$Y = \{2, 4, 6, 8, 10\} \qquad \mathbb{Z} = \{n \mid n \text{ is an integer}\}$$

Use the set operations $\subseteq$, $\cup$, $\cap$, and $\setminus$ to describe as many relations among these sets as possible.

Exercise 1.18.5. Let $W, X, Y, \mathbb{E}, \mathbb{O}$, and $\mathbb{Z}$ be the sets defined in Exercise 1.18.4.

(a.) Compute the number of elements of $X \times Y$.

(b.) List at least three distinct elements of $\mathbb{O} \times \mathbb{E}$.

(c.) List all elements of the diagonal Δ_X of X.

(d.) Every odd integer can be written as $2k + 1$ for some integer k, and every even integer can be written as 2ℓ for some integer ℓ. Express the sets $\mathbb{O}$ and $\mathbb{E}$ in set-builder notation accordingly.

(e.) Convince yourself that $\mathbb{O}$ and $\mathbb{E}$ have "essentially the same" number of elements; then, find a function $f : \mathbb{O} \to \mathbb{E}$ such that f is injective and f is surjective. Observe that this gives a rigorous justification of the fact that $\mathbb{O}$ and $\mathbb{E}$ have "essentially the same" number of elements.

(f.) Convince yourself that $\mathbb{O}$ and $\mathbb{Z}$ have "essentially the same" number of elements; then, find a function $f : \mathbb{O} \to \mathbb{Z}$ such that f is injective and f is surjective. Conclude from this exercise and the previous one that there are "as many" odd (or even) integers as there are integers.

Exercise 1.18.6. Let $\mathbb{Z}$ denote the set of integers.

(a.) Provide a partition of $\mathbb{Z}$ into three sets.

(**Hint:** What are the possible remainders of an integer modulo 3?)

(b.) Provide a partition of $\mathbb{Z}$ into four sets.

(c.) Provide a partition of $\mathbb{Z}$ into n sets for any positive integer n.

Exercise 1.18.7. Consider the set W consisting of all words in the English language.

(a.) Prove that $R = \{(v, w) \in W \times W \mid v \text{ and } w \text{ begin with the same letter}\}$ is an equivalence relation on W; then, determine the number of distinct equivalence classes of W modulo R.

(b.) Prove that $R = \{(v, w) \in W \times W \mid v \text{ and } w \text{ have the same number of letters}\}$ is an equivalence relation on W; then, describe the equivalence class of the word "awesome."

Exercise 1.18.8. Let $\mathbb{Z}$ be the set of integers. Prove that $(a, b)\ R\ (c, d)$ if and only if $ad = bc$ on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ is an equivalence relation. Describe the collection of distinct equivalence classes.

(**Hint:** For the second part of the problem, try replacing the notation (a, b) with a/b, instead.)

Exercise 1.18.9. Let X be an arbitrary set. Consider the collection $S = \{Y \mid Y \subseteq X\}$. Prove that the inclusion $\subseteq$ defines a partial order P on S such that $(Y_1, Y_2) \in P$ if and only if $Y_1 \subseteq Y_2$; then, either prove that P is a total order on S, or provide a counterexample to show that it is not.

Exercise 1.18.10. List the maximal elements of the subset $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$ of the set $\mathbb{Z}_{\geq 0}$ of non-negative integers with respect to the partial order D of divisibility.

(**Hint:** List as many pairs of comparable elements of S as necessary to compute the chains in S with three or four elements; then, use this information deduce the maximal elements of S.)

Exercise 1.18.11. Complete the following using modular arithmetic.

(a.) If $a \equiv 1 \pmod 6$, find the least positive x for which $5a + 4 \equiv x \pmod 6$.

(b.) If $a \equiv 4 \pmod 7$ and $b \equiv 5 \pmod 7$, find the least positive x for which $6a - 3b \equiv x \pmod 7$.

(c.) (Modular Exponentiation) Use the fact that $2^{2023} \equiv 8 \pmod{10}$ to find $2022^{2023} \pmod{10}$.

Exercise 1.18.12. Consider any nonzero integer n and any integers a and b. If $ab \equiv 0 \pmod n$, then must it be true that $a \equiv 0 \pmod n$ or $b \equiv 0 \pmod n$? Explain.

Exercise 1.18.13. Let p be any prime number. Prove that if a and b are any integers such that $ab \equiv 0 \pmod p$, then either $a \equiv 0 \pmod p$ or $b \equiv 0 \pmod p$.

Exercise 1.18.14. Let X and Y be arbitrary sets.

(a.) Prove that if there exists a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$, then f is injective.

(b.) Prove that if there exists a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$, then f is surjective.

Chapter 2

Logic and Truth Tables

Generally, the purpose of mathematics is to describe the universe quantitatively in a manner that is consistent, replicable, and unambiguous. Combined with the language of set theory, the calculus of logic provides the basis for mathematical communication: if sets, relations, and functions constitute the skeleton of some structure or organism that can be modelled mathematically, then the connective tissue is represented by (mathematical) statements, logical quantifiers, and truth tables. We introduce in this chapter several axioms and symbols that are commonplace in modern logic.

2.1 Statements

We have thus far garnered a working knowledge of set theory — including the theory of relations and functions — and we have seen examples of mathematical proofs. We turn our attention next to fleshing out some details regarding the calculus of logic that will soon assist us with writing original proofs. We will assume throughout this section that the symbols P and Q are **statements**, i.e., P and Q are complete sentences that assert a property that is unambiguously true (T) or false (F).

Example 2.1.1. "Every positive whole number is an integer" is an example of a true statement.

Example 2.1.2. "The integer 10 is divisible by 3" is an example of a false statement.

Example 2.1.3. "The weather in Kansas City is lovely this time of year" is not a statement because some might think so, but others might not: its truth value is ambiguous. Generally, any sentence that is exclamatory (e.g., any observation), imperative (e.g., any command), or interrogative (e.g., any question) is not a statement because these types of sentences have no inherent truth value.

Exclamatory: "What a story, Mark!"

Imperative: "Don't forget to mow the lawn."

Interrogative: "How about those Chiefs?"

We will henceforth refer to the verity of a statement as its **truth value**. Our ability to determine the truth value of a sentence does not preclude the possibility that the sentence is a valid statement; indeed, there are many unsolved statements throughout mathematics. Generally, a statement whose truth value is undetermined is called a **conjecture**. Common examples of mathematical statements with undetermined veracity include those that involve a potentially unknown or variable quantity x. We have encountered statements of these kinds throughout many of our mathematics courses.

Example 2.1.4. "The real number x is irrational" is an example of a valid statement; it is neither true nor false, but rather, its truth value depends explicitly on the value of the real number x.

Conventionally, any declarative statement of the form $P(x)$ for some variable quantity x is called an **open sentence**; the set of all possible values that x can assume is called the **domain** of x; and the truth value of $P(x)$ depends explicitly upon the determination of the variable x.

Example 2.1.5. Observe that the statement $P(x)$ that "the real number x is irrational" is an open sentence; the domain of x is the set of real numbers; and $P(x)$ is true if and only if $x \in \mathbb{R} \setminus \mathbb{Q}$.

We will typically represent an open sentence in the variable x by the symbol $P(x)$, and we will separate $P(x)$ from the open sentence it represents with a colon, as in the following example.

Example 2.1.6. Consider the following open sentence.

$$P(x)\colon \text{ We have that } x^2 - 1 = 0.$$

By solving for the unknown quantity x, we find that $P(x)$ is a true statement if and only if $x = \pm 1$, hence the natural domain for the statement $P(x)$ is the set $\mathbb{Z}$ of integers.

Example 2.1.7. Consider the following open sentence.

$$P(x)\colon \text{ We have that } x^2 + 1 = 0$$

By solving for the unknown value x, we find that $P(x)$ is a true statement if and only if $x = \pm\sqrt{-1}$, hence $P(x)$ is false if the domain of x is any subset of $\mathbb{R}$ because $\sqrt{-1}$ is not a real number.

Example 2.1.8. Consider the following open sentence.

$$P(x, y)\colon \text{ We have that } x^2 + y^2 \geq 0$$

Considering that $x^2 + y^2 \geq 0 + 0 = 0$ for any pair of real numbers x and y, it follows that $P(x, y)$ is a true statement if the domain of x and y is any subset of the set $\mathbb{R}$ of real numbers; however, if the domains of x and y are both the set $\mathbb{C}$ of complex numbers, then we can determine values of x and y such that $P(x, y)$ is false. Concretely, we note that $P(i, i)$ is false since $i^2 + i^2 = -1 - 1 = -2 < 0$.

Example 2.1.9. Consider the following open sentence.

$$P(x, y)\colon \text{ We have that } x + y \text{ is a positive prime number.}$$

Let us assume throughout this example that the domain of x is $X = \{1, 2, 3, 4\}$ and the domain of y is $Y = \{-1, -2, -3, -4\}$. Calculating the sum $x + y$ for each of the sixteen elements of $X \times Y$, we find that $P(x, y)$ is true if and only if $(x, y) \in \{(3, -1), (4, -1), (4, -2)\}$; otherwise, $P(x, y)$ is false.

Often, it is convenient to collect the truth values of some finitely many statements $P_1, P_2, \ldots, P_n$ in a **truth table**. Each column of a truth table contains one statement followed by all of its possible truth values relative to the other statements. Concretely, the first row of a truth table contains the symbols that represent the statements, and the subsequent rows contain the possible truth values of each statement relative to the other. Considering that any statement attains one and only truth value, a truth table for the n statements $P_1, P_2, \ldots, P_n$ admits n columns and $2^n + 1$ rows as follows.

P
T
F

P	Q
T	T
T	F
F	T
F	F

P	Q	R
T	T	T
T	T	F
T	F	T
T	F	F
F	T	T
F	T	F
F	F	T
F	F	F

Table 2.1: the truth tables for one, two, and three statements

2.2 Conjunction, Disjunction, and Negation

We examine next the myriad ways to construct new statements from any finite number of existing statements. We concern ourselves immediately with a statement P. We refer to the statement "not P" (precisely, "It is not the case that P") as the **negation** of P; symbolically, the negation of any statement P is denoted by $\neg P$. Often, it is possible to represent the negation $\neg P$ of a statement P in a less clunky way than simply by, "It is not the case that P," as the following examples illustrate.

Example 2.2.1. Consider the following statement.

$$P\colon \text{ The integer 2 is even.}$$

By definition, the negation $\neg P$ of the given statement P is the following statement.

$$\neg P\colon \text{ It is not the case that the integer 2 is even.}$$

Considering that any integer is either even or odd, we can rephrase $\neg P$ as follows.

$$\neg P\colon \text{ The integer 2 is odd.}$$

Crucially, we note that P is a true statement, and its negation $\neg P$ is a false statement.

Example 2.2.2. Consider the following statement.

$$P\colon \text{ The integer 111 is prime.}$$

By definition, the negation $\neg P$ of the given statement P is the following statement.

$$\neg P\colon \text{ It is not the case that the integer 111 is prime.}$$

Even less clunky than the above is the following representation of $\neg P$.

$$\neg P\colon \text{ The integer 111 is not prime.}$$

Better yet, since every integer is either prime or composite, we can rephrase $\neg P$ as follows.

$$\neg P\colon \text{ The integer 111 is composite.}$$

Observe that in this case, the statement P is false, and its negation $\neg P$ is a true statement.

Generally, it ought to be clear to the reader that the statements P and $\neg P$ have opposite truth values: if P is true, then $\neg P$ must be false; however, if P is false, then $\neg P$ must be true.

P	$\neg P$
T	F
F	T

Table 2.2: the truth table for the negation $\neg P$

Even more, we will soon see for any statement P, it must be the case that either P is true or $\neg P$ is true. Before we arrive at this conclusion, we must discuss other ways to create new statements from a pair of statements P and Q. One way to do this is to consider the case that either P is true or Q is true. Put into symbols, the **disjunction** $P \vee Q$ of the statements P and Q is the statement, "Either it is the case that P or it is the case that Q," for which the upside-down wedge $\vee$ denotes the connective "or." Compare the similarities between the disjunction $\vee$ and the set union $\cup$.

Example 2.2.3. Consider the following pair of statements.

P: Topeka is the capital of Kansas.

Q: The real number $\sqrt{2}$ is a root of $x^2 - 2$.

We may construct the disjunction $P \vee Q$ by placing the connective "or" between the statements.

$P \vee Q$: Either Topeka is the capital of Kansas or the real number $\sqrt{2}$ is a root of $x^2 - 2$.

Both of the statements P and Q are in fact true, hence the disjunction $P \vee Q$ is true.

Example 2.2.4. Consider the following pair of statements.

P: Kansas City is the capital of Missouri.

Q: The real number π is transcendental.

We may construct the disjunction $P \vee Q$ by placing the connective "or" between the statements.

$P \vee Q$: Either Kansas City is the capital of Missouri or the real number π is transcendental.

Even though the statement P is false (since the capital of Missouri is Jefferson City), the disjunction $P \vee Q$ is true because π is a transcendental number (this fact is non-trivial but well-known).

Example 2.2.5. Consider the following pair of statements.

P: The square root of -1 is a real number

Q: The integer 11 is composite.

We may construct the disjunction $P \vee Q$ by placing the connective "or" between the statements.

$P \vee Q$: Either the square root of -1 is a real number or the integer 11 is composite.

Both of these statements P and Q are false: $\sqrt{-1}$ is a non-real complex number, and 11 is prime. Consequently, the disjunction $P \vee Q$ is a false statement because neither P nor Q is a true statement.

Crucially, we note that if either of the statements P or Q is true, then the disjunction $P \vee Q$ must also be true; however, if neither of the statements P or Q is true, then $P \vee Q$ must be false.

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Table 2.3: the truth table for the disjunction $P \vee Q$

We may also consider the case that both of the statements P and Q are true simultaneously. Put another way, we may form the statement, "It is the case that P and it is the case that Q." We refer to this statement as the **conjunction** $P \wedge Q$ of P and Q, and we use the wedge $\wedge$ as the connective "and." Compare the similarities between the conjunction $\wedge$ and the set intersection $\cap$.

Example 2.2.6. Consider the following pair of statements.

P: Paris is the capital of France.

Q: The real number 1 is less than the real number $\sqrt{2}$.

We may construct the conjunction $P \wedge Q$ by placing the connective "and" between the statements.

$P \wedge Q$: Paris is the capital of France, and the real number 1 is less than the real number $\sqrt{2}$.

Both of the statements P and Q are in fact true, hence the conjunction $P \wedge Q$ is true.

Example 2.2.7. Consider the following pair of statements.

P: Leticia is the capital of France.

Q: The identity function on a set is injective.

We may construct the conjunction $P \wedge Q$ by placing the connective "and" between the statements.

$P \wedge Q$: Leticia is the capital of France, and the identity function on a set is injective.

Considering that the statement P is false (since we know that Paris is the capital of France), the conjunction $P \wedge Q$ is false. Explicitly, it is not true that both P and Q are true, so $P \wedge Q$ is false.

Example 2.2.8. Consider the following pair of statements.

P: We have that $\cos(k\pi) = 0$ for all integers k.

Q: The integer 8 is a perfect square.

We may construct the conjunction $P \wedge Q$ by placing the connective "and" between the statements.

$P \wedge Q$: We have that $\cos(k\pi) = 0$ for all integers k, and integer 8 is a perfect square.

Both of these statements are false: indeed, $\cos(k\pi) = (-1)^k$ for all integers k, and $\sqrt{8} = 2\sqrt{2}$ is not an integer. Consequently, the conjunction $P \wedge Q$ is false because neither P nor Q is true.

We note that the conjunction $P \wedge Q$ of statements P and Q is true if and only if both P and Q are true. Consequently, if either of the statements P or Q is false, then $P \wedge Q$ is false. Be careful not to confuse the upside-down wedge $\vee$ (meaning "or") with the wedge $\wedge$ (meaning "and").

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Table 2.4: the truth table for conjunction $P \wedge Q$

We are now in a position to state and prove two fundamental principles in the calculus of logic.

Theorem 2.2.9 (Law of the Excluded Middle)**.** *We have that $P \vee \neg P$ is true for any statement P.*

Proof. Given any statement P, consider the disjunction $P \vee \neg P$ of P and $\neg P$. Observe that if P is true, then $P \vee \neg P$ is true. Conversely, if P is false, then $\neg P$ is true, hence $P \vee \neg P$ is true. □

P	$\neg P$	$P \vee \neg P$
T	F	T
F	T	T

Table 2.5: the Law of the Excluded Middle

Theorem 2.2.10 (Law of Non-Contradiction)**.** *We have that $P \wedge \neg P$ is false for any statement P.*

Proof. Given any statement P, consider the conjunction $P \vee \neg P$ of P and $\neg P$. Observe that if P is true, then $\neg P$ is false, hence $P \wedge \neg P$ is false. Conversely, if P is false, then $P \wedge \neg P$ is false. □

P	$\neg P$	$P \wedge \neg P$
T	F	F
F	T	F

Table 2.6: the Law of Non-Contradiction

2.3 Conditional and Biconditional Statements

Going forward, we will be interested primarily in statements of the form $P \Rightarrow Q$ in which the two-tailed right arrow $\Rightarrow$ reads "implies." Under this convention, the entire statement $P \Rightarrow Q$ can be read either as "P implies Q" or "If P, then Q." Unsurprisingly, a statement of this form is called an **implication** or a **conditional statement**. We refer to the statement P in this construction as the **antecedent**; the statement Q is called the **consequent**. Observe that the statement $P \Rightarrow Q$ is false if and only if Q is false and P is true; otherwise, the implication $P \Rightarrow Q$ is true.

Example 2.3.1. Consider the following pairs of statements.

P: Madrid is the capital of Spain.

Q: The integer 3 is odd.

We may construct the implication $P \Rightarrow Q$ as follows.

$P \Rightarrow Q$: If Madrid is the capital of Spain, then the integer 3 is odd.

Considering that both P and Q are true statements, it follows that $P \Rightarrow Q$ is true.

Example 2.3.2. Consider the following pairs of statements.

P: The integer 3 divides the integer 243.

Q: The integer 3 is even.

We may construct the implication $P \Rightarrow Q$ as follows.

$P \Rightarrow Q$: If the integer 3 divides the integer 243, then the integer 3 is even.

Observe that $243 = 81 \cdot 3$, hence 3 divides 243; however, we know well that 3 is not an even integer. Consequently, the conditional statement $P \Rightarrow Q$ is false: indeed, we are lying here.

Below is the truth table for the conditional statement $P \Rightarrow Q$, as indicated above.

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Table 2.7: the truth table for the implication $P \Rightarrow Q$

Crucially, if the statement P is false, then according to Table 2.7, the implication $P \Rightarrow Q$ is true regardless of the truth value of Q; in this case, the conditional statement $P \Rightarrow Q$ is called a **vacuous truth**, or equivalently, we say that $P \Rightarrow Q$ is **vacuously** true. Concretely, the idea is that the antecedent P cannot be satisfied because it is false, so the implication must be true.

Example 2.3.3. Consider the following pairs of statements.

P: The integer 17 is negative.

Q: Dr. Beck is a multi-instrumentalist.

We may construct the implication $P \Rightarrow Q$ as follows.

$P \Rightarrow Q$: If the integer 17 is negative, then Dr. Beck is a multi-instrumentalist.

Considering that the antecedent P is false (since its negation "$\neg P$: The integer 17 is positive." is in fact the true statement), it follows that the conditional statement $P \Rightarrow Q$ is vacuously true.

One way to justify this result (as promised by Table 2.7) is that no lies were told: Dr. Beck is a multi-instrumentalist, so there was no harm in (falsely) assuming that 17 is a negative integer.

Example 2.3.4. Consider the following pairs of statements.

P: The integer 17 is negative.

Q: Dr. Beck is a multi-millionaire.

We may construct the implication $P \Rightarrow Q$ as follows.

$P \Rightarrow Q$: If the integer 17 is negative, then Dr. Beck is a multi-millionaire.

Considering that the antecedent P is false (since its negation "$\neg P$: The integer 17 is positive." is in fact the true statement), it follows that the conditional statement $P \Rightarrow Q$ is vacuously true. (Unfortunately for Dr. Beck, this makes no difference for his situation: the integer 17 is positive.)

One way to verify this result is that no lies were told: Dr. Beck is in fact not a multi-millionaire, but on the other hand, there was nothing guaranteed unless 17 were in fact a negative integer.

We will typically say that "P implies Q" or "If P, then Q" if the conditional statement $P \Rightarrow Q$ is true. Conventionally, if P implies Q, then we will say that P is **sufficient** for Q. One can rephrase this by saying that P is sufficient for Q if Q is true provided the statement P. Crucially, as Table 2.7 illustrates, the statement P may be either true or false; it does not actually matter. Equivalently, we may say that "P only if Q" if the conditional statement $P \Rightarrow Q$ is true. We declare in this case that Q is **necessary** for P. Consequently, each of the following statements is equivalent.

(a.) $P \Rightarrow Q$

(b.) If P, then Q.

(c.) Q if P.

(d.) P only if Q.

(e.) P is sufficient for Q.

(f.) Q is necessary for P.

We will fix our attention throughout the rest of the course primarily on conditional statements in which P and Q are open sentences. Consider the following examples along these lines.

Example 2.3.5. Consider the following pairs of statements about a positive integer n.

$$P(n)\text{: The integer } n^4+1 \text{ is prime.}$$
$$Q(n)\text{: The integer } n^2+1 \text{ is prime.}$$

By plugging in different values of the integer $n \geq 1$, we obtain explicit statements $P(n)$ and $Q(n)$.

$P(1)$: The integer 2 is prime.
$P(2)$: The integer 17 is prime.
$P(3)$: The integer 82 is prime.
$P(4)$: The integer 257 is prime.
$P(5)$: The integer 626 is prime.

$Q(1)$: The integer 2 is prime.
$Q(2)$: The integer 5 is prime.
$Q(3)$: The integer 10 is prime.
$Q(4)$: The integer 17 is prime.
$Q(5)$: The integer 26 is prime.

Consider the conditional statement $P(n) \Rightarrow Q(n)$ defined as follows.

$$P(n) \Rightarrow Q(n)\text{: If the integer } n^4+1 \text{ is prime, then the integer } n^2+1 \text{ is prime.}$$

By Table 2.7, we know that $P(n) \Rightarrow Q(n)$ is false if and only if $P(n)$ is true and $Q(n)$ is false. Consequently, the statement $P(n) \Rightarrow Q(n)$ is true for all integers $1 \leq n \leq 5$. Quite astonishingly, this statement is in fact true for all integers $1 \leq n \leq 27$; however, we have that $28^4+1 = 614657$ is prime and $28^2+1 = 785$ is not prime, hence the statement $P(28) \Rightarrow Q(28)$ is false.

Example 2.3.6. Consider the following pairs of statements about a positive integer n.

$$P(n)\text{: The integer } n^2+1 \text{ is prime.}$$
$$Q(n)\text{: The integer } n^4-1 \text{ is prime.}$$

By definition, the conditional statement $P(n) \Rightarrow Q(n)$ is given as follows.

$$P(n) \Rightarrow Q(n)\colon \text{ If the integer } n^2+1 \text{ is prime, then the integer } n^4-1 \text{ is prime.}$$

By Table 2.7, we know that $P(n) \Rightarrow Q(n)$ is false if and only if $P(n)$ is true and $Q(n)$ is false. Considering that $n^4-1=(n^2-1)(n^2+1)$ is divisible by n^2+1 for all integers n, it follows that n^4-1 is composite for all integers n, hence the open sentence $Q(n)$ is false for every integer n. We conclude that the conditional statement $P(n) \Rightarrow Q(n)$ is false for all integers $n \geq 1$.

Example 2.3.7. Consider the following pairs of statements about a pair of real numbers x and y.

$$P(x,y)\colon \text{ We have that } x+y=1.$$
$$Q(x,y)\colon \text{ We have that } x^2+y^2=1.$$

By definition, the conditional statement $P(x,y) \Rightarrow Q(x,y)$ is given as follows.

$$P(x,y) \Rightarrow Q(x,y)\colon \text{ If } x+y=1, \text{ then } x^2+y^2=1.$$

By Table 2.7, we know that $P(x,y) \Rightarrow Q(x,y)$ is false if and only if $P(x,y)$ is true and $Q(x,y)$ is false. Observe that if $x+y=1$, then $y=1-x$ so that $y^2=x^2-2x+1$. Consequently, it follows that $x^2+y^2=2x^2-2x+1$; thus, the open sentence $Q(x,y)$ is true if and only if $2x^2-2x+1=1$ if and only if $2x^2-2x=0$ if and only if $2x(x-1)=0$ if and only if $x=0$ or $x=1$. We conclude that the statement $P(x,y) \Rightarrow Q(x,y)$ is true if and only if $x=0$ and $y=1$ or $x=1$ and $y=0$.

Given any pair of statements P and Q, the conditional statement $Q \Rightarrow P$ formed by swapping the antecedent and the consequent is called the **converse** of the conditional statement $P \Rightarrow Q$. Like the implication $P \Rightarrow Q$, its converse $Q \Rightarrow P$ can be understood in different ways.

(a.) $Q \Rightarrow P$

(b.) If Q, then P.

(c.) P if Q.

(d.) Q only if P.

(e.) Q is sufficient for P.

(f.) P is necessary for Q.

P	Q	$Q \Rightarrow P$
T	T	T
T	F	T
F	T	F
F	F	T

Table 2.8: the truth table for the converse $Q \Rightarrow P$ of the implication $P \Rightarrow Q$

Example 2.3.8. Consider the following pairs of statements.

$$P\colon \text{ The integer 17 is negative.}$$
$$Q\colon \text{ Dr. Beck is a multi-instrumentalist.}$$

We may construct the converse $Q \Rightarrow P$ of the implication $P \Rightarrow Q$ as follows.

$$Q \Rightarrow P\colon \text{ If Dr. Beck is a multi-instrumentalist, then the integer 17 is negative.}$$

Unlike the implication $P \Rightarrow Q$ (which is vacuously true since P is false), the converse $Q \Rightarrow P$ is false: Dr. Beck is a multi-instrumentalist, but the integer 17 is not negative.

Example 2.3.9. Consider the following pairs of statements.

P: The integer 17 is negative.

Q: Dr. Beck is a multi-millionaire.

We may construct the converse $Q \Rightarrow P$ of the implication $P \Rightarrow Q$ as follows.

$Q \Rightarrow P$: If Dr. Beck is a multi-millionaire, then the integer 17 is negative.

Considering that the antecedent Q is false, the conditional statement $Q \Rightarrow P$ is vacuously true. One way to verify this result is that no lies were told: Dr. Beck is not a multi-millionaire.

Bearing in mind Examples 2.3.3 and 2.3.8, even if P is sufficient for Q, it may not be true that P is necessary for Q; however, if P is both necessary and sufficient for Q, then both of the conditional statements $P \Rightarrow Q$ and $Q \Rightarrow P$ are true. We say in this case that P holds if and only if Q holds, and we represent this relationship symbolically by $P \Leftrightarrow Q$. We will typically say that the statements P and Q are **(materially) equivalent** if P is true if and only if Q is true. Put another way, the material equivalence $P \Leftrightarrow Q$ is simply the conjunction $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$. Each of the following statements concerning the **biconditional statement** $P \Leftrightarrow Q$ is equivalent.

(a.) $P \Leftrightarrow Q$

(b.) P if and only if Q.

(c.) P is (materially) equivalent to Q.

(d.) P is necessary and sufficient for Q.

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$	$P \Leftrightarrow Q$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	T	F	F	F
F	F	T	T	T	T

Table 2.9: the truth table for the biconditional $P \Leftrightarrow Q$

Put another way, we have that $P \Leftrightarrow Q$ is true if and only if P and Q have the same truth value.

Example 2.3.10. Consider the following statements.

P: The integer 3 divides the integer 243.

Q: The integer 3 is even.

R: The integer 17 is negative.

S: The integer 2027 is prime.

Considering that P and S are true statements, but Q and R are false statements, it follows that $P \Leftrightarrow S$ and $Q \Leftrightarrow R$ are both true statements and $P \Leftrightarrow Q$, $P \Leftrightarrow R$, $Q \Leftrightarrow S$, and $R \Leftrightarrow S$ are all false statements. Examples of these statements in words are provided below.

$P \Leftrightarrow Q$: The integer 3 divides the integer 243 if and only if 3 is even.

$P \Leftrightarrow S$: The integer 3 divides the integer 243 if and only if the integer 2027 is prime.

$Q \Leftrightarrow R$: The integer 3 is even if and only if the integer 17 is negative.

Example 2.3.11. Consider the following statements about an integer n.

$$P(n)\colon \text{ The integer } n \text{ is even.}$$
$$Q(n)\colon \text{ The integer } n^2 \text{ is even.}$$

We construct the biconditional statement $P(n) \Leftrightarrow Q(n)$ as follows.

$$P(n) \Leftrightarrow Q(n)\colon \text{ The integer } n \text{ is even if and only if the integer } n^2 \text{ is even.}$$

By definition, an integer n is even if and only if $n = 2k$ for some integer k. Consequently, if n is even, then $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$ is even. Conversely, if n^2 is even, then there exists an integer k such that $n^2 = 2k$. Considering that 2 is prime, we must have that 2 divides n, hence n is even. We conclude therefore that the statement $P(n) \Leftrightarrow Q(n)$ is true for all integers n.

Example 2.3.12. Consider the following statements about an integer n.

$$P(n)\colon \text{ The integer } n \text{ is odd.}$$
$$Q(n)\colon \text{ The integer } n^2 \text{ is odd.}$$

We construct the biconditional statement $P(n) \Leftrightarrow Q(n)$ as follows.

$$P(n) \Leftrightarrow Q(n)\colon \text{ For the integer } n \text{ to be odd, it is necessary and sufficient that } n^2 \text{ is odd.}$$

Often, this construction is more awkward than the more natural "if and only if" statement.

$$P(n) \Leftrightarrow Q(n)\colon \text{ The integer } n \text{ is odd if and only if the integer } n^2 \text{ is odd.}$$

We leave it as an exercise for the reader to prove that $P(n) \Leftrightarrow Q(n)$ is true for all integers n.

Example 2.3.13. Consider the following pairs of statements about a pair of real numbers x and y.

$$P(x,y)\colon \text{ We have that } x^2 + y^2 = 1.$$
$$Q(x,y)\colon \text{ We have that } (x,y) \text{ lies on a circle of radius 1 centered at } (0,0).$$

By definition, the statements $P(x,y) \Rightarrow Q(x,y)$ and $Q(x,y) \Rightarrow P(x,y)$ are as follows.

$$P(x,y) \Rightarrow Q(x,y)\colon \text{ If } x^2 + y^2 = 1, \text{ then } (x,y) \text{ lies on a circle of radius 1 centered at } (0,0).$$
$$Q(x,y) \Rightarrow P(x,y)\colon \text{ If } (x,y) \text{ lies on a circle of radius 1 centered at } (0,0), \text{ then } x^2 + y^2 = 1.$$

Recall that the equation of a circle of radius r centered at (h,k) is given by

$$(x-h)^2 + (y-k)^2 = r^2.$$

Consequently, the conditional statement $Q(x,y) \Rightarrow P(x,y)$ is true by definition. Conversely, if $x^2 + y^2 = 1$, then $(x-0)^2 + (y-0)^2 = 1^2$ implies that the point $(x,y) \in \mathbb{R} \times \mathbb{R}$ lies on a circle of radius 1 centered at $(0,0)$. Put another way, we have that $P(x,y) \Rightarrow Q(x,y)$ is true. Ultimately, these observations together yield that the biconditional statement $P(x,y) \Leftrightarrow Q(x,y)$ is true.

2.4 Tautologies and Contradictions

By the Law of the Excluded Middle, the statement $P \vee \neg P$ ("P or not P") is always true; it is a **tautology**. Generally, a tautology is any statement that is true for all possible truth inputs.

Example 2.4.1. Given any statements P and Q, the disjunction $(\neg Q) \vee (P \Rightarrow Q)$ is a tautology. We can convince ourselves of this by realizing that $P \Rightarrow Q$ is true if either P is false or P and Q are both true. Consequently, the statement $(\neg Q) \vee (P \Rightarrow Q)$ is true in the case that P is false or P and Q are both true. But if Q is false, then $\neg Q$ is true, hence $(\neg Q) \vee (P \Rightarrow Q)$ is true.

P	Q	$\neg Q$	$P \Rightarrow Q$	$(\neg Q) \vee (P \Rightarrow Q)$
T	T	F	T	T
T	F	T	F	T
F	T	F	T	T
F	F	T	T	T

Table 2.10: the truth table for $(\neg Q) \vee (P \Rightarrow Q)$

Example 2.4.2. Given any statements P and Q, the implication $[(P \vee Q) \wedge (\neg Q)] \Rightarrow P$ is a tautology: indeed, it suffices to check that all of its values in the following truth table are T.

P	Q	$\neg Q$	$P \vee Q$	$(P \vee Q) \wedge (\neg Q)]$	$[(P \vee Q) \wedge (\neg Q)] \Rightarrow P$
T	T	F	T	F	T
T	F	T	T	T	T
F	T	F	T	F	T
F	F	T	F	F	T

Table 2.11: the truth table for $[(P \vee Q) \wedge (\neg Q)] \Rightarrow P$

Consequently, when Beyoncé says, "I break the internet: top two, and I ain't number two" on the track "Top Off" by DJ Khaled, it means that she is number one.

By the Law of Non-Contradiction, the statement $P \wedge \neg P$ ("P and not P") is always false; it is a **contradiction**. Generally, a contradiction is a statement that is false for all possible truth inputs.

Example 2.4.3. Given any statements P and Q, the conjunction $P \wedge [P \Rightarrow (Q \wedge \neg Q)]$ is a contradiction. We can verify this by convincing ourselves (by the Law of Non-Contradiction) that $Q \wedge \neg Q$ is false; therefore, the conditional statement $P \Rightarrow (Q \wedge \neg Q)$ is false if P is true. On the other hand, the implication $P \Rightarrow (Q \wedge \neg Q)$ is true if P is false. Combined, these two observations yield that P and $P \Rightarrow (Q \wedge \neg Q)$ take opposite truth values, so their conjunction is false.

P	Q	$\neg Q$	$Q \wedge \neg Q$	$P \Rightarrow (Q \wedge \neg Q)$	$P \wedge [P \Rightarrow (Q \wedge \neg Q)]$
T	T	F	F	F	F
T	F	T	F	F	F
F	T	F	F	T	F
F	F	T	F	T	F

Table 2.12: the truth table for $P \wedge [P \Rightarrow (Q \wedge \neg Q)]$

Example 2.4.4. Given any statements P and Q, the conjunction $(P \wedge Q) \wedge [Q \Rightarrow \neg P]$ is a contradiction. We can verify this by constructing the corresponding truth table as follows.

P	Q	$\neg P$	$P \wedge Q$	$Q \Rightarrow \neg P$	$(P \wedge Q) \wedge (Q \Rightarrow \neg P)$
T	T	F	T	F	F
T	F	F	F	T	F
F	T	T	F	T	F
F	F	T	F	T	F

Table 2.13: the truth table for $(P \wedge Q) \wedge (Q \Rightarrow \neg P)$

2.5 Logical Equivalence

Given any statements P and Q, recall from Table 2.7 that the conditional statement $P \Rightarrow Q$ is vacuously true if P is false; therefore, in order to determine the truth value of $P \Rightarrow Q$, it suffices to consider the case that P is true. Unfortunately, in some situations, it is difficult to establish the verity of Q even if P is known to be true. Under these circumstances, it is not possible to determine if the statement $P \Rightarrow Q$ is true or false because this depends entirely on whether Q is true or false; however, it is possible in some cases to extract a statement $S(P, Q)$ that depends on both P and Q that is **logically equivalent** to the implication $P \Rightarrow Q$. We say that two statements S_1 and S_2 are logically equivalent if and only if their values in a truth table are equal; if this is the case, then we write $S_1 \equiv S_2$ to assert symbolically that S_1 and S_2 are logically equivalent. Consequently, if we demonstrate that the statement $S(P, Q)$ is true, then $P \Rightarrow Q$ must be true, as well.

We will concern ourselves primarily with the interplay between the conjunction, disjunction, implication, and negation. We seek to construct a glossary of statements that are logically equivalent to the implication $P \Rightarrow Q$. Conventionally, if the statement P is false, then the implication $P \Rightarrow Q$ is vacuously true. Even more, if the statement Q is true, then the implication $P \Rightarrow Q$ is trivially true regardless of the truth value of P. Consequently, we may deduce that the statements $P \Rightarrow Q$ and $\neg P \vee Q$ are logically equivalent, as the following truth table illustrates.

P	Q	$\neg P$	$P \Rightarrow Q$	$\neg P \vee Q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Table 2.14: the truth table for the implication $P \Rightarrow Q$ and the disjunction $\neg P \vee Q$

Proposition 2.5.1. *Given any statements P and Q, we have that $(P \Rightarrow Q) \equiv (\neg P \vee Q)$.*

Consider the statement $\neg Q \Rightarrow \neg P$ called the **contrapositive** of the implication $P \Rightarrow Q$. Observe that if Q is true, then $\neg Q$ is false, hence the statement $\neg Q \Rightarrow \neg P$ is vacuously true. Likewise, if Q is true, then the statement $P \Rightarrow Q$ is true regardless of the verity of P. Conversely, if Q is false, then $\neg Q$ is true, hence $\neg Q \Rightarrow \neg P$ is true if and only if $\neg P$ is true if and only if P is false. Consequently, we are lead to the following truth table and the subsequent proposition.

P	Q	$\neg P$	$\neg Q$	$P \Rightarrow Q$	$\neg Q \Rightarrow \neg P$
T	T	F	F	T	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

Table 2.15: the truth table for the contrapositive $\neg Q \Rightarrow \neg P$ of the implication $P \Rightarrow Q$

Proposition 2.5.2. *Given any statements P and Q, we have that $(P \Rightarrow Q) \equiv (\neg Q \Rightarrow \neg P)$.*

Example 2.5.3. Consider the following statements.

P: Bernard earns an A on his final exam in MA291.

Q: Bernard earns an A as his final grade in MA291.

Let us assume that if Bernard earns an A on his final exam in MA291, then Bernard earns an A as his final grade in MA291. Consider the following statements regarding Bernard's grade in MA291.

R: Either Bernard does not earn an A on his final exam or Bernard earns an A in MA291.

S: If Bernard does not earn an A in MA291, then Bernard did not earn an A on his final exam.

Observe that the statement R is true: indeed, if Bernard does not earn an A on his final exam, then there is no promise as to what his final grade in MA291 will be, so no lies have been told regardless of the outcome. On the other hand, if Bernard earns an A as his final grade, then it does not matter what he earned on his final exam in MA291 because he will surely be happy with his grade. Likewise, the statement S is true: indeed, if Bernard does not earn an A as his final grade, then he must not have earned an A on his final exam because that would have guaranteed him an A in the course. We have corroborated the logical equivalence of the statements $P \Rightarrow Q$, $\neg P \vee Q$, and $\neg Q \Rightarrow \neg P$ for the example at hand, as guaranteed by Propositions 2.5.1 and 2.5.2.

Example 2.5.4. Consider the following statements.

P: It is overcast in Kansas City.

Q: Bernard brings an umbrella to work.

Let us assume as before that if it is overcast in Kansas City, then Bernard brings an umbrella to work. Observe that if Bernard does not bring an umbrella to work, then it must not be overcast in Kansas City; otherwise, if it were overcast in Kansas City, then Bernard would have brought an umbrella to work. Even more, it is either sunny in Kansas City or Bernard brings an umbrella to work: indeed, if Bernard does not bring an umbrella to work, then it must be sunny in Kansas City. Our exposition here bears out the logical equivalence of $P \Rightarrow Q$, $\neg Q \Rightarrow \neg P$, and $\neg P \vee Q$.

Often, it is useful to determine when the conditional statement $P \Rightarrow Q$ is false (i.e., P does not provide sufficient information from which to deduce Q). By Table 2.7, we have that $P \Rightarrow Q$ is false if and only if P is true and Q is false if and only if $P \wedge \neg Q$ is true, hence the statements $\neg(P \Rightarrow Q)$ and $P \wedge \neg Q$ are logically equivalent, as the following truth table illustrates.

P	Q	$\neg Q$	$P \Rightarrow Q$	$\neg(P \Rightarrow Q)$	$P \wedge \neg Q$
T	T	F	T	F	F
T	F	T	F	T	T
F	T	F	T	F	F
F	F	T	T	F	F

Table 2.16: the truth table for the negated implication $\neg(P \Rightarrow Q)$ and the disjunction $P \wedge \neg Q$

Proposition 2.5.5. *Given any statements P and Q, we have that $\neg(P \Rightarrow Q) \equiv (P \wedge \neg Q)$.*

Example 2.5.6. Consider the following statements.

P: Bernard earns an A on his final exam in MA291.

Q: Bernard earns an A as his final grade in MA291.

Observe that if Bernard earns an A on his final exam in MA291 but Bernard does not earn an A as his final grade, then it is a lie to say that if Bernard earns an A on his final exam in MA291, then Bernard earns an A as his final grade in MA291; this illustrates the result of Proposition 2.5.5.

By Table 2.3, if $P \vee Q$ is false, then neither P nor Q is true. Likewise, by Table 2.4, if $P \wedge Q$ is false, then either P is false or Q is false. Combined, these observations form **De Morgan's Laws**.

P	Q	$\neg P$	$\neg Q$	$P \vee Q$	$\neg(P \vee Q)$	$\neg P \wedge \neg Q$	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P \vee \neg Q$
T	T	F	F	T	F	F	T	F	F
T	F	F	T	T	F	F	F	T	T
F	T	T	F	T	F	F	F	T	T
F	F	T	T	F	T	T	F	T	T

Table 2.17: the truth table for $\neg(P \vee Q)$ and $\neg(P \wedge Q)$

Theorem 2.5.7 (De Morgan's Laws)**.** *Consider any statements P and Q.*

(a.) *We have that $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$, i.e., $\neg(P \vee Q)$ is logically equivalent to $\neg P \wedge \neg Q$.*

(b.) *We have that $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$, i.e., $\neg(P \wedge Q)$ is logically equivalent to $\neg P \vee \neg Q$.*

Example 2.5.8. Consider the following statements.

P: It is overcast in Kansas City.

Q: Bernard brings an umbrella to work.

Observe that if it is not the case that either it is overcast in Kansas City or Bernard brings an umbrella to work, then it must be the case that neither it is overcast in Kansas City nor Bernard brings an umbrella to work. Likewise, if it is not the case that it is overcast in Kansas City and Bernard brings an umbrella to work, then either it is not overcast in Kansas City or Bernard does not bring an umbrella to work. We have thus verified De Morgan's Laws for the given statements.

2.6 Quantified Statements

Often, we seek to determine the verity of an open sentence for all possible values in its domain. Explicitly, if $P(x)$ is any open sentence that depends on a variable x with domain S, then for each element $s \in S$, the truth value of the statement $P(s)$ is well-defined and can be determined.

Example 2.6.1. Consider the following statement about an integer n.

$$P(n)\colon \text{ The integer } n \text{ is even.}$$

We can plainly see that the verity of $P(n)$ depends entirely on the value of n. Each of the following statements in the left-hand column is true, but each statement in the right-hand column is false.

$P(0)\colon$ The integer 0 is even.	$P(1)\colon$ The integer 1 is even.
$P(2)\colon$ The integer 2 is even.	$P(3)\colon$ The integer 3 is even.
$P(4)\colon$ The integer 4 is even.	$P(5)\colon$ The integer 5 is even.

Quantification is another process of converting an open sentence $P(x)$ in the variable x into a statement whose truth value can be determined. **Quantified statements** are expressed using **logical quantifiers**. Primarily, we will study three logical quantifies throughout this course.

We use the **universal quantifier** $\forall$ to symbolically represent the phrases "for all," "for every," or "for each." Consequently, the statement $\forall x \in S$, $P(x)$ can be understood in words as, "For all elements $x \in S$, we have that $P(x)$." Observe that the quantified statement $\forall x \in S$, $P(x)$ is true if $P(x)$ is true for all elements $x \in S$; otherwise, this statement is false. Put another way, if the statement $P(x_0)$ is false for some element $x_0 \in S$, then the statement $\forall x \in S$, $P(x)$ is false.

Summary 2.6.2. Given any open sentence $P(x)$ with domain S, the quantified statement

$$\forall x \in S,\ P(x)\colon \text{ For every element } x \in S, \text{ we have that } P(x).$$

is true if and only if $P(x)$ is true for all elements $x \in S$. Conversely, this quantified statement is false if and only if there exists an element $x_0 \in S$ such that $P(x_0)$ is false.

Example 2.6.3. Consider the following statement about an integer n.

$$P(n)\colon \text{ The integer } n \text{ is even.}$$

By using the universal quantifier $\forall$ ("for all"), we obtain the following quantified statement.

$$\forall n \in \mathbb{Z},\ P(n)\colon \text{ For every integer } n, \text{ we have that } n \text{ is even.}$$

By Example 2.6.1 and Summary 2.6.2, the above quantified statement is false because $P(1)$ is false.

Example 2.6.4. Consider the following statements about an integer n.

$$P(n)\colon \text{ The integer } n \text{ is even.}$$
$$Q(n)\colon \text{ The integer } n^2 \text{ is even.}$$

By using the universal quantifier $\forall$ ("for all"), we obtain the following quantified statement.

$$\forall n \in \mathbb{Z},\ [P(n) \Leftrightarrow Q(n)]\colon \text{ For every integer } n, \text{ we have that } n \text{ is even if and only if } n^2 \text{ is even.}$$

By Example 2.3.11, this statement is true because $P(n) \Leftrightarrow Q(n)$ is true for all integers n.

Example 2.6.5. Consider the following statement about a pair of real numbers x and y.

$$P(x, y)\colon \text{ The real number } x^2 + y^2 \text{ is non-negative.}$$

By using the universal quantifier $\forall$ ("for all"), we obtain the following quantified statement.

$$\forall x, y \in \mathbb{R},\ P(x, y)\colon \text{ For every pair of real numbers } x \text{ and } y, \text{ we have that } x^2 + y^2 \geq 0.$$

Considering that $x^2 \geq 0$ for every real number x, it follows that $x^2 + y^2 \geq 0$ for every pair of real numbers x and y. Consequently, the quantified statement $\forall x, y \in \mathbb{R}, P(x, y)$ is true.

Example 2.6.6. Consider the following statements about a real number x.

$$P(x)\colon \text{ The real number } x^2 + 4 \text{ satisfies that } x^2 + 4 \geq 4.$$
$$Q(x)\colon \text{ The real number } x^2 + 4 \text{ satisfies that } x^2 + 4 \leq 4.$$

By using the universal quantifier $\forall$ ("for all"), we obtain the following quantified statements.

$$\forall x \in \mathbb{R},\ P(x)\colon \text{ For all real numbers } x, \text{ we have that } x^2 + 4 \geq 4.$$
$$\forall x \in \mathbb{R},\ Q(x)\colon \text{ For all real numbers } x, \text{ we have that } x^2 + 4 \leq 4.$$

Considering that $x^2 \geq 0$ for every real number x, it follows that $x^2 + 4 \geq 4$ for every real number x. Consequently, the quantified statement $\forall x \in \mathbb{R},\ P(x)$ is true; however, $Q(1)$ is false because the real number $5 = 1^2 + 4$ does not satisfy that $5 \leq 4$. We conclude that $\forall x \in \mathbb{R},\ Q(x)$ is false.

One other indispensable way to view the universally quantified statement $\forall x \in S,\ P(x)$ in words is, "If x is an element of S, then we have that $P(x)$" or "If $x \in S$, then $P(x)$." Observe that in this manner, any statement involving the universal quantifier is simply a conditional statement. Consequently, Proposition 2.5.1 entails the logical equivalence of the universally quantified statement $\forall x \in S,\ P(x)$ and the disjunction $(x \notin S) \vee P(x)$. By De Morgan's Laws, the negation of $\forall x \in S,\ P(x)$ is logically equivalent to the negation of $(x \notin S) \vee P(x)$ — namely, $(x \in S) \wedge \neg P(x)$.

Summary 2.6.7. Given any open sentence $P(x)$ with domain S, the following are equivalent.

(a.) $\forall x \in S,\ P(x)\colon$ For every element $x \in S,$ we have that $P(x)$.

(b.) $(x \notin S) \vee P(x)\colon$ Either x is not an element of S or we have that $P(x)$.

Better yet, the negation of a quantified statement is itself a quantified statement. Explicitly, we use the **existential quantifier** $\exists$ to express the phrases "there exists," "for at least one," or "for some." Consequently, the quantified statement $\exists x \in S,\ P(x)$ can be understood in words as, "There exists an element $x \in S$ such that $P(x)$." Observe that the quantified statement $\exists x \in S,\ P(x)$ is true if $P(x_0)$ is true for some element $x_0 \in S$; otherwise, this statement is false. Put another way, if $P(x)$ is false for every element $x \in S$, then the quantified statement $\exists x \in S,\ P(x)$ is false.

Summary 2.6.8. Given any open sentence $P(x)$ with domain S, the quantified statement

$$\exists x \in S,\ P(x)\colon \text{ There exists an element } x \in S \text{ such that } P(x).$$

is true if and only if $P(x_0)$ is true for some element $x_0 \in S$. Conversely, this quantified statement is false if and only if the statement $P(x)$ is false for all elements $x \in S$.

Example 2.6.9. Consider the following statement about an integer n.

$$P(n)\colon \text{ The integer } n \text{ is even.}$$

By using the existential quantifier $\exists$ ("there exists"), we obtain the following quantified statement.

$$\exists n \in \mathbb{Z},\ P(n)\colon \text{ There exists an integer } n \text{ such that } n \text{ is even.}$$

Certainly, the above quantified statement is true because $P(2)$ is true.

Example 2.6.10. Consider the following statement about an integer n.

$$P(n)\colon \text{ The integer } n^4 + 1 \text{ is prime.}$$

By using the existential quantifier $\exists$ ("there exists"), we obtain the following quantified statement.

$$\exists n \in \mathbb{Z},\ P(n)\colon \text{ There exists an integer } n \text{ such that } n^4 + 1 \text{ is prime.}$$

Considering that $2 = 1^4 + 1$ is prime, $P(1)$ is true, hence the above quantified statement is true.

Example 2.6.11. Consider the following statement about a pair of real numbers x and y.

$$P(x, y)\colon \text{ The real numbers } x \text{ and } y \text{ satisfy that } x^2 + y^2 = 4.$$

By using the existential quantifier $\exists$ ("there exists"), we obtain the following quantified statement.

$$\exists x, y \in \mathbb{R},\ P(x, y)\colon \text{ There exist real numbers } x \text{ and } y \text{ such that } x^2 + y^2 = 4.$$

Considering that the set of ordered pairs (x, y) of real numbers satisfying that $x^2 + y^2 = 4$ is the graph of a circle of radius 2 centered at the origin in the Cartesian plane, it follows that the above quantified statement is true: indeed, both of the statements $P(2, 0)$ and $P(0, 2)$ are true.

Example 2.6.12. Consider the following statements about a real number x.

$$\begin{aligned} P(x)&\colon \text{ The real number } x \text{ satisfies that } x^2 - 2x - 3 = 0.\\ Q(x)&\colon \text{ The real number } x^3 \text{ satisfies that } x^3 \geq 8. \end{aligned}$$

By using the existential quantifier $\exists$ ("there exists"), we obtain the following quantified statements.

$$\begin{aligned} \exists x \in \mathbb{R},\ [P(x) \Rightarrow Q(x)]&\colon \text{ There exists a real number } x \text{ such that } x^3 \geq 8 \text{ if } x^2 - 2x - 3 = 0.\\ \exists x \in \mathbb{R},\ [P(x) \wedge \neg Q(x)]&\colon \text{ There exists a real number } x \text{ such that } x^2 - 2x - 3 = 0 \text{ and } x^3 < 8. \end{aligned}$$

Observe that if $P(x)$ is false, then the conditional statement $P(x) \Rightarrow Q(x)$ is vacuously true. Consequently, the first quantified statement above is true for any real number x such that $x^2 - 2x - 3$ is nonzero (e.g., suppose that $x = 0$ or $x = 1$). On the other hand, we can determine explicitly the values of x such that $P(x)$ is true since $(x-3)(x+1) = x^2 - 2x - 3 = 0$ if and only if $x = 3$ or $x = -1$. Consequently, we have that $P(3) \Rightarrow Q(3)$ is true. Likewise, the second quantified statement above is true because the real number $x = -1$ satisfies that $(-1)^2 - 2(-1) - 3 = 0$ and $(-1)^3 = 1 < 8$. Put another way, we have that $P(-1)$ is true and $Q(-1)$ is false.

We provide next the crucial theorem that relates the universal and existential quantifiers.

Theorem 2.6.13 (Negation of Quantified Statements)**.** *Consider any open sentence $P(x)$ over S.*

1.) *We have that* $\neg[\forall x \in S,\, P(x)] \equiv [\exists x \in S,\, \neg P(x)]$.

2.) *We have that* $\neg[\exists x \in S,\, P(x)] \equiv [\forall x \in S,\, \neg P(x)]$.

Last, if $P(x)$ is any open sentence whose domain is any nonempty set S, then we say that an element $x_0 \in S$ is the **unique** element of S **satisfying the statement** $P(x_0)$ if and only if

(a.) the statement $P(x_0)$ is true and

(b.) for every element $x \in S$, if $P(x)$ is true, then we must have that $x = x_0$.

We use the **uniqueness quantifier** ! to represent the phrase "unique." Explicitly, we will write $\exists! x \in S,\, P(x)$ to signify that "there exists a unique element $x \in S$ such that $P(x)$."

Example 2.6.14. Consider the following statement about an integer n.

$$P(n)\colon \text{ The integer } n \text{ satisfies that } 3n - 4 = 5.$$

Observe that $P(n)$ is true if and only if $3n - 4 = 5$ if and only if $n = 3$, hence the statement $P(n)$ admits a unique element $n_0 \in \mathbb{Z}$ satisfying that $P(n_0)$ is true: namely, it is the integer $n_0 = 3$. Put another way, the following quantified statement involving the uniqueness quantifier is true.

$$\exists! n \in \mathbb{Z},\, P(n)\colon \text{ There exists a unique integer } n \text{ such that } 3n - 4 = 9.$$

Example 2.6.15. Consider the following statement about a real number x.

$$P(x)\colon \text{ The real number } x \text{ satisfies that } x - 5 + \frac{25}{x+5} = \frac{4x+5}{x+5}.$$

By solving the rational equation that defines $P(x)$, we find that $P(x)$ is true if and only if $x = 1$.

$$\frac{(x-5)(x+5)+25}{x+5} = \frac{4x+5}{x+5}$$

$$(x-5)(x+5)+25 = 4x+5$$

$$x^2 - 25 + 25 = 4x + 5$$

$$x^2 - 4x - 5 = 0$$

$$(x-1)(x+5) = 0$$

Considering that $x + 5$ cannot equal 0, the Zero Product Property yields that $x - 1 = 0$ so that $x = 1$. Put another way, the following statement involving the uniqueness quantifier is true.

$$\exists! x \in \mathbb{R},\, P(x)\colon \text{ There exists a unique real number } x \text{ such that } x - 5 + \frac{25}{x+5} = \frac{4x+5}{x+5}.$$

2.7 Chapter 2 Overview

We say that a complete sentence P is a **statement** if it asserts something that can be unambiguously measured as true or false. Examples of statements include, "The integer 3 is an odd" and "The integer 17 is negative." We note that the first statement is true, but the second statement is false. Using logical connectives, we can form new statements from given statements P and Q. Explicitly, the **implication** $P \Rightarrow Q$ is the statement, "P implies Q" (or equivalently, "If P, then Q"); the implication is false if and only if P is true and Q is false. Regardless of the verity of the statement Q, if the statement P is false, then the implication $P \Rightarrow Q$ must be **vacuously** true. We define the **disjunction** $P \vee Q$ ("P or Q"), the **conjunction** $P \wedge Q$ ("P and Q"), and the **negation** $\neg P$ ("not P"). Observe that the disjunction $P \vee Q$ is true if and only if P is true or Q is true; the conjunction $P \vee Q$ is true if and only if P is true and Q is true; and the negation $\neg P$ is true if and only if P is false. Given any statement P, the disjunction $P \vee \neg P$ is true by the Law of the Excluded Middle, and the conjunction $P \wedge \neg P$ is false by the Law of Non-Contradiction.

We use **truth tables** to deduce the verity of a statement $S(P, Q)$ that depends on two statements P and Q. One can construct a truth table for $S(P, Q)$ by writing all possible **truth values** of P in one column; all possible truth values of Q in a subsequent column; and the resultant truth values of the statement $S(P, Q)$ in a third column. Considering that the statements P and Q could themselves depend upon other statements $P_1, \ldots, P_n$, a truth table grows arbitrarily large as the number of statements increases. Generally, we need $2^n + 1$ rows and $n + 1$ columns to construct the truth table of a statement $S(P_1, \ldots, P_n)$ defined for n distinct statements $P_1, \ldots, P_n$.

We say that two statements P and W are **logically equivalent** if and only if they induce the same truth table; in particular, if P and Q are equivalent statements, the truth values of P are the same as the truth values of Q for all possible truth inputs, hence the verity of the statement P is exactly the same as the verity of the statement Q. Even more, if the truth values of P are all true, then P is a **tautology**; if the truth values for P are all false, then P is a **contradiction**.

De Morgan's Laws are two rules of inference that relate the conjunction, disjunction, and negation. Concretely, De Morgan's Laws assert the logical equivalence of the following statements.

(a.) $\neg(P \vee Q)$: It is not the case that either P or Q.

(b.) $\neg P \wedge \neg Q$: It is neither the case that P nor the case that Q.

Likewise, De Morgan's Laws for the negation of a conjunction assert the equivalence of the following.

(c.) $\neg(P \wedge Q)$: It is not the case that both P and Q.

(d.) $\neg P \vee \neg Q$: It is either not the case that P or not the case that Q.

Logical quantifiers allow us to symbolically handle statements involving quantities. We use the **universal quantifier** $\forall$ to express that an open sentence $P(x)$ is true "for all" possible values of x in its domain, and we use the **existential quantifier** $\exists$ to express that "there exists" a value of x in the domain of $P(x)$ such that $P(x)$ is true. We say that an element x_0 in the domain of the open sentence $P(x)$ is **unique** if it is the only value in the domain of $P(x)$ such that $P(x_0)$ is true. We use the **uniqueness quantifier** $\exists!$ to express the existence ($\exists$) and uniqueness ($!$) of x_0.

2.8 Chapter 2 Exercises

Exercise 2.8.1. Explain whether each of the following is a statement. Provide the negation of each statement; identify tautologies and contradictions; and write the contrapositive of each implication.

(a.) I yam what I yam.

(b.) If you know, then you know.

(c.) Where there is a will, there is a way.

(d.) Jacob, keep your head down!

(e.) Every four years, there is a Leap Year.

(f.) Does it come in a pint?

(g.) Not all who wander are lost.

(h.) I was and I was not.

(i.) Either it is freezing or Sam wears shorts.

(j.) There exists an irrational number.

Exercise 2.8.2. Consider the following statements.

P: The sun is shining in Kansas City.
Q: Bernard rides his bike to work.

Use the symbols P and Q and logical connectives such as the biconditional $\Leftrightarrow$ conjunction $\wedge$, disjunction $\vee$, implication $\Rightarrow$ and negation $\neg$ to convert each of the following statements into symbols; then, identify all logically equivalent statements, tautologies, and contradictions.

(a.) If the sun is shining in Kansas City, then Bernard rides his bike to work.

(b.) Bernard rides his bike to work only if the sun is shining in Kansas City.

(c.) Either the sun is not shining in Kansas City or Bernard rides his bike to work.

(d.) The sun is shining in Kansas City, and Bernard does not ride his bike to work.

(e.) If the sun is not shining in Kansas City, then Bernard does not ride his bike to work.

(f.) If Bernard does not ride his bike to work, then the sun is not shining in Kansas City.

(g.) Neither the sun is shining in Kansas City nor Bernard rides his bike to work.

(h.) Either the sun is not shining in Kansas City or Bernard does not ride his bike to work.

(i.) The sun is not shining in Kansas City, and Bernard does not ride his bike to work.

(j.) Either Bernard rides his bike to work or Bernard does not ride his bike to work.

(k.) The sun is shining in Kansas City, and the sun is not shining in Kansas City.

(l.) Bernard rides his bike to work if and only if the sun is shining in Kansas City.

(m.) The sun is not shining in Kansas City if and only if Bernard does not ride his bike to work.

Exercise 2.8.3. Let P, Q, and R be any statements. Construct an appropriate truth table to prove that the statements "If P, then Q or R" and "If P and not Q, then R" are logically equivalent.

Chapter 3

Basic Proof Techniques

Generally, mathematical research and problem solving are carried out in two steps: first, one must conduct extensive experimentation to determine some underlying pattern; then, the most significant effort is exerted to establish the veracity of the observed phenomenon in general. Concretely, this is achieved using set theory and the calculus of logic to construct a mathematical proof. Put simply, a mathematical proof is nothing more than a convincing argument that is replicable and unambiguous. We demonstrate in this chapter how to employ the basic axioms and general principles of certain mathematical structures to write mathematical proofs. We devote our attention to the three most common types of proofs: direct proof, proof by contrapositive, and proof by contradiction.

3.1 Direct Proof

Our primary focus throughout this chapter is to use the foundations of the calculus of logic presented in Chapter 2 to inform and develop the writing of mathematical proofs: indeed, the overwhelming impetus of contemporary mathematics lies in proving statements of the form "if P, then Q" for some statements (or open sentences) P and Q. Consequently, our attention will be by-and-large fixed on conditional statements of the form $P \Rightarrow Q$. Considering the truth table 2.7 for the implication, if either the statement Q is true or the statement P is false, then the conditional statement $P \Rightarrow Q$ is true. Proofs that are carried out by showing that Q is true are called **trivial proofs**. Conversely, any proof that demonstrates that P is false is called a **vacuous proof**. We begin our discussion of direct proofs with this low-hanging fruit, as illustrated in the following typical examples.

Example 3.1.1. Prove that if n is an even integer, then $n^2 + 4 \geq 3$.

Solution. Consider the following statements involving an integer n.

$$\begin{aligned} P(n)&: \text{ The integer } n \text{ is even.} \\ Q(n)&: \text{ The integer } n \text{ satisfies that } n^2 + 4 \geq 3. \end{aligned}$$

We seek to prove that $\forall n \in \mathbb{Z}$, $[P(n) \Rightarrow Q(n)]$ is a true statement. Considering that $n^2 \geq 0$ for any real number (and hence any integer) n, it follows that $n^2 + 4 \geq 4$. Consequently, the statement $Q(n)$ is true for all integers n, hence the statement $\forall n \in \mathbb{Z}$, $[P(n) \Rightarrow Q(n)]$ is trivially true. $\diamond$

Our above work is merely a suggestion of a proof of the statement in Example 3.1.1. Below, we provide an example of how a proof of this statement might look "in the wild." Crucially, observe that in the following proof, there is no need to provide any symbols for the statements.

Proof. (Example 3.1.1) Considering that $n^2 \geq 0$ for any real number n, it follows that $n^2 + 4 \geq 4$. Consequently, we have that $n^2 + 4 \geq 3$ for every integer n, so the claim holds trivially. □

We point out at this juncture two important features of a mathematical proof. First, it is vitally important for the writer to indicate the beginning of a proof with an italicized "Proof" and a period. Equally as important is the ending of the proof. We will use in this course an empty box □ to signal the conclusion of a proof; however, the reader may alternatively use the acronym "QED" (Latin for "quod erat demonstrandum" or "what was to be shown") depending upon their preference.

Example 3.1.2. Prove that if a real number x satisfies that $x^2 - 2 = 0$, then 7 is an odd integer.

Solution. Like the previous example, the hypothesis that x is a real number satisfying that $x^2 - 2 = 0$ has no bearing on the truth value of the conclusion that 7 is an odd integer: indeed, 7 is an odd integer, so regardless of what hypotheses we make, the if-then statement remains true. ◇

Proof. (Example 3.1.2) Considering that 7 is an odd integer, the statement is trivially true. □

Example 3.1.3. Prove that if the Riemann Hypothesis holds, then $\frac{d}{dx}e^x = e^x$.

Proof. By elementary calculus, it holds that $\frac{d}{dx}e^x = e^x$, hence the statement is trivially true. □

Example 3.1.4. Prove that if -1 is an even integer, then the Riemann Hypothesis holds.

Solution. We are now in the opposite case of a trivial proof: indeed, the hypotheses of the statement are false because -1 is not an even integer, hence the statement is true vacuously. ◇

Proof. (Example 3.1.4) Considering that -1 is an odd integer, the statement is vacuously true. □

Example 3.1.5. Prove that if there exist a pair of real numbers x and y such that $x^2 + y^2 = -4$, then only finitely many positive integers are prime.

Proof. Given any real number x, we have that $x^2 \geq 0$. Consequently, we find that $x^2 + y^2 \geq 0$ for all real numbers x and y. Bearing this in mind, it follows that the statement is vacuously true. □

Example 3.1.6. Prove that if $\{(x, y) \mid x, y \in \mathbb{R} \text{ and } y^2 = x\}$ is a function, then $\frac{1}{0} = 1$.

Proof. Observe that if $x = 1$, then the real numbers $y = 1$ and $y = -1$ both satisfy that $y^2 = x$. Consequently, the ordered pairs $(1, 1)$ and $(1, -1)$ both belong to $\{(x, y) \mid x, y \in \mathbb{R} \text{ and } y^2 = x\}$, hence this relation is not a function. We conclude that the statement is vacuously true. □

Often, it will not be the case that we will encounter a statement that can be proved by a trivial or vacuous proof; rather, we will typically assume that the hypotheses of the statement are true in the first place, and we will subsequently perform some algebraic analysis or arithmetic manipulation in order to rigorously justify that the conclusion of the statement holds. We refer to this process as a **direct proof**. Explicitly, a direct proof of a conditional statement $P \Rightarrow Q$ usually begins

with the phrase, "Suppose that P is true" and ends with the phrase, "We conclude that Q is true." Between these two points, the writer is left to fill in the details — how ever complicated they are.

Crucially, the validity of a direct proof relies on the law of inference called **modus ponens** that asserts that the conditional statement $[(P \Rightarrow Q) \wedge P] \Rightarrow Q$ is a tautology. Eliminating the trivial or vacuous cases, in order to establish the verity of a conditional statement $P \Rightarrow Q$, we need only assume that P is true and deduce from this that $P \Rightarrow Q$ is true (because if P is false, then $P \Rightarrow Q$ is true vacuously). Let us construct a truth table to verify the law of modus ponens.

P	Q	$P \Rightarrow Q$	$(P \Rightarrow Q) \wedge P$	$[(P \Rightarrow Q) \wedge P] \Rightarrow Q$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

Table 3.1: the truth table for modus ponens $[(P \Rightarrow Q) \wedge P] \Rightarrow Q$

We conclude this section with several examples of direct proofs that require a bit more work than trivial or vacuous proofs. Before this, we recall the definitions of an even integer versus an odd integer. Explicitly, an integer n is **even** if and only if there exists an integer k such that $n = 2k$. Conversely, an integer n is **odd** if and only if there exists an integer ℓ such that $n = 2\ell + 1$.

Example 3.1.7. Prove that if n is an even integer, then $4n + 7$ is an odd integer.

Proof. By definition, if n is an even integer, then there exists an integer k satisfying that $n = 2k$. Consequently, we have that $4n+7 = 4(2k)+7 = 8k+7 = (8k+6)+1 = 2(4k+3)+1$. Considering that $4k + 3$ is also an integer, it follows that $4n + 7$ is an odd integer, as desired. □

Example 3.1.8. Prove that if n is an odd integer, then $3n - 1$ is an even integer.

Proof. By definition, if n is an odd integer, then there exists an integer k satisfying that $n = 2k+1$. Consequently, we have that $3n - 1 = 3(2k + 1) - 1 = 6k + 2 = 2(3k + 1)$. Considering that $3k + 1$ is also an integer, it follows that $3n - 1$ is an even integer, as desired. □

Example 3.1.9. Prove that if n is an even integer, then $3n^2 + 5n - 3$ is an odd integer.

Proof. By definition, if n is an even integer, then $n = 2k$ for some integer k. Consequently, we have

$$3n^2 + 5n - 3 = 3(2k)^2 + 5(2k) - 3 = 12k^2 + 10k - 3 = (12k^2 + 10k - 4) + 1 = 2(6k^2 + 5k - 2) + 1.$$

Considering that $6k^2 + 5k - 2$ is an integer, it follows that $3n^2 + 5n - 3$ is an odd integer. □

Example 3.1.10. Prove that if a, b, c are integers, then $ab + ac + bc$ is even if a and b are even.

Proof. We will assume that a, b, and c are integers such that a and b are even. By definition, there exist integers k and ℓ such that $a = 2k$ and $b = 2\ell$. Consequently, we have that

$$ab + ac + bc = (2k)(2\ell) + (2k)c + (2\ell)c = 4k\ell + 2(ck) + 2(c\ell) = 2(ck + c\ell + 2k\ell).$$

Considering that $ck + c\ell + 2k\ell$ is an integer, it follows that $ab + ac + bc$ is an even integer. □

3.2 Proof by Contrapositive

Consider any pair of statements P and Q. Recall from Section 2.5 that the **contrapositive** of the conditional statement $P \Rightarrow Q$ is the conditional statement $\neg Q \Rightarrow \neg P$. By the result of Table 2.15 and Proposition 2.5.2, any conditional statement is logically equivalent to its contrapositive. Consequently, the **proof by contrapositive** is a proof technique that exploits this logical equivalence. Explicitly, a proof by contrapositive is used to establish the verity of a conditional statement $P \Rightarrow Q$ by instead demonstrating the truth of its contrapositive statement $\neg Q \Rightarrow \neg P$ and using the logical equivalence of the two statements to conclude the truth of the original implication $P \Rightarrow Q$. Bearing this in mind, a typical proof by contrapositive ought to begin with the phrase, "Suppose that $\neg Q$ is true" and end with the phrase, "We conclude that $\neg P$ is true."

Before we proceed with an illustration of the technique of proof by contrapositive, we turn our attention to the law of inference called **modus tollens** that is closely related to the law of modus ponens and asserts that the conditional statement $[\neg Q \wedge (P \Rightarrow Q)] \Rightarrow \neg P$ is a tautology.

P	Q	$\neg P$	$\neg Q$	$P \Rightarrow Q$	$\neg Q \wedge (P \Rightarrow Q)$	$[\neg Q \wedge (P \Rightarrow Q)] \Rightarrow \neg P$
T	T	F	F	T	F	T
T	F	F	T	F	F	T
F	T	T	F	T	F	T
F	F	T	T	T	T	T

Table 3.2: the truth table for modus tollens $[\neg Q \wedge (P \Rightarrow Q)] \Rightarrow \neg P$

Proof by contrapositive is a powerful technique that is most useful when either the verity of Q is difficult to deduce from the verity of P or $\neg Q$ is a stronger hypothesis than P itself. We illustrate the importance and usefulness of the proof by contrapositive in the following examples. Be sure to make note of where a direct proof might falter or what difficulties arise from weak assumptions.

Example 3.2.1. Prove that if n is an integer, then n is even if and only if n^2 is even.

Solution. Consider the following statements involving an integer n.

$$P(n)\text{: The integer } n \text{ is even.}$$
$$Q(n)\text{: The integer } n^2 \text{ is even.}$$

We seek to establish the veracity of the biconditional statement $P(n) \Leftrightarrow Q(n)$ for each integer n. Consequently, we must establish that both the implication $P(n) \Rightarrow Q(n)$ and its converse $Q(n) \Rightarrow P(n)$ are true for each integer n. One direction is fairly straightforward: if the integer n is even, then there exists an integer k such that $n = 2k$. By squaring both sides of this equation, we conclude that $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$ is even because $2k^2$ is an integer. Conversely, if we assume that n is an integer such that n^2 is even, then there exists an integer k such that $n^2 = 2k$. Unfortunately, this assumption does not afford us much deductive power: it is unclear to the author (and likely to the reader) at this point why the equation $n^2 = 2k$ entails that n must be even. (Later, we will learn about division by prime numbers, but for now, we make no assumption that the reader is familiar with this technique.) Consequently, the hypothesis of $Q(n)$ is relatively "weak."

We may therefore seek to prove the conditional statement $Q(n) \Rightarrow P(n)$ by contrapositive: indeed, we fare immediately better using this proof technique because the assumption $\neg P(n)$ that n is an odd integer is "stronger" than the assumption that n^2 is an even integer. Concretely, if n is an odd integer, then $n = 2k + 1$ for some integer k. By squaring both sides of this equation, we find that $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. Considering that $2k^2 + 2k$ is an integer, we conclude that n^2 is an odd integer; thus, our proof by contrapositive is complete. ◇

Proof. (Example 3.2.1) We will assume first that n is an even integer. By definition, there exists an integer k satisfying that $n = 2k$. Consequently, by squaring both sides of this equation, we find that $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$. Considering that $2k^2$ is an integer, it follows that n^2 is even.

Conversely, we will prove the converse by contrapositive. We must assume to this end that n is an odd integer. By definition of an odd integer, there exists an integer k such that $n = 2k + 1$. By squaring both sides of this equation, we find that $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. Considering that $2k^2 + 2k$ is an integer, it follows that n^2 is an odd integer, as desired. □

Example 3.2.2. Prove that if n is an integer such that $7n + 6$ is even, then n is even.

Solution. We might first attempt a direct proof. Observe that if $7n+6$ is even, then $7n+6 = 2k$ for some integer k. By subtracting 6 from both sides, we find that $7n = 2k-6 = 2(k-3)$; however, it is here that things become unclear without a solid understanding of how prime numbers behave with respect to divisibility. Consequently, a direct proof is unsatisfactory; on the other hand, we might fare better with a proof by contrapositive. Observe that if n is odd, then there exists an integer k such that $n = 2k + 1$ and $7n + 6 = 7(2k + 1) + 6 = 14k + 13 = (14k + 12) + 1 = 2(7k + 6) + 1$. We conclude therefore that if n is odd, then $7n + 6$ is odd, hence the contrapositive is true. ◇

Proof. (Example 3.2.2) We will prove the contrapositive of the statement. We must assume to this end that n is an odd integer. By definition of an odd integer, there exists an integer k such that $n = 2k + 1$. Observe that $7n + 6 = 7(2k + 1) + 6 = 14k + 13 = (14k + 12) + 1 = 2(7k + 6) + 1$. Considering that $7k + 6$ is an integer, it follows that $7n + 6$ is an odd integer, as desired. □

Example 3.2.3. Prove that if n is an integer such that $7n - 3$ is odd, then $11n + 6$ is even.

Solution. We will first attempt a direct proof. We will assume along these lines that $7n-3 = 2k+1$ for some integer k. By adding $4n + 9$ to both sides of this equation, we find that

$$11n + 6 = (7n - 3) + (4n + 9) = (2k + 1) + (4n + 9) = 2k + 4n + 10 = 2(k + 2n + 5)$$

so that $11n+6$ is an even integer because $k+2n+5$ is an integer. But perhaps it seems miraculous to the reader that we were able to add $4n+9$ to both sides of the equation to obtain a direct proof. Bearing this in mind, we might seek a proof by contrapositive; this would entail that $11n+6 = 2k+1$ for some integer k so that $11n = 2k - 5$. We are at this point stuck because it is not clear how to extract any meaning from this equation. Our intuition might suggest that if $7n - 3$ is odd, then n must be even: indeed, an odd integer times an odd integer is an odd integer, and the difference of two odd integers is an odd integer, so n cannot (ostensibly) be odd. We are therefore brought to the potential midpoint in the present problem to prove that if $7n - 3$ is odd, then n is even. ◇

Often, the proof of an assertion could benefit from (or potentially even requires) some more powerful observation. Conventionally, such a helping proposition is referred to as a **lemma**. Let us state and prove a lemma that will make the proof of the previous example follow more efficiently.

Lemma 3.2.4. *If n is an integer such that $7n-3$ is an odd integer, then n is even.*

Solution. We might first attempt a direct proof: indeed, suppose that $7n-3=2k+1$ for some integer k. We have that $7n=2k+4=2(k+2)$. But again, without knowledge of divisibility of prime numbers, this equation is rather useless; we will therefore attempt a proof by contrapositive for this lemma. Observe that if n is odd, then there exists an integer k satisfying that $n=2k+1$. Consequently, we have that $7n-3=7(2k+1)-3=14k+4=2(7k+2)$ is even, as desired. ◇

Proof. (Lemma 3.2.4) We will prove the contrapositive of the statement of the lemma. We must assume to this end that n is an odd integer. By definition of an odd integer, there exists an integer k such that $n=2k+1$. Observe that $7n-3=7(2k+1)-3=14k+4=2(7k+2)$. Considering that $7k+2$ is an integer, it follows that $7n+6$ is an even integer, as desired. □

Proof. (Example 3.2.3) By Lemma 3.2.4, if n is an integer such that $7n-3$ is odd, then n is even. Consequently, there exists an integer k such that $n=2k$. Even more, we have that

$$11n+6=11(2k)+6=22k+6=2(11k+3).$$

Considering that $11k+3$ is an integer, we conclude that $11n+6$ is an even integer. □

Example 3.2.5. Prove that if n is any integer, then $2n^2+n$ is odd if and only if $\cos\left(\frac{n\pi}{2}\right)=0$.

Solution. Glancing at this proposition, it might seem quite unwieldy — after all, we are comparing the parity of an integer $2n^2+n$ with the roots of the cosine function — but if one takes a moment to recognize the values this cosine sequences takes, the proof strategy becomes clear: indeed, computing $\cos\left(\frac{n\pi}{2}\right)=0$ for some integers n, the reader will have a much better handle of the situation.

$$\cos(0)=1 \qquad \cos\left(\frac{\pi}{2}\right)=0 \qquad \cos(\pi)=-1 \qquad \cos\left(\frac{3\pi}{2}\right)=0$$

Consequently, we deduce that $\cos\left(\frac{n\pi}{2}\right)=0$ if and only if n is odd. We are lead to the following. ◇

Lemma 3.2.6. *If n is an integer, then $\cos\left(\frac{n\pi}{2}\right)=0$ if and only if n is odd.*

Proof. By elementary trigonometry, we have that $\cos\left(\frac{n\pi}{2}\right)=0$ if and only if $\frac{n\pi}{2}=\frac{(2k+1)\pi}{2}$ for some integer k if and only if $n=2k+1$ for some integer k if and only if n is an odd integer. □

Proof. (Example 3.2.5) By Lemma 3.2.6, it suffices to prove that $2n^2+n$ is odd if and only if n is odd. We will assume first that n is an odd integer. By definition of an odd integer, there exists an integer k such that $n=2k+1$. Consequently, we have that

$$2n^2+n=2(2k+1)^2+(2k+1)=2(4k^2+4k+1)+(2k+1)=2(4k^2+5k+1)+1.$$

Considering that $4k^2+5k+1$ is an integer, it follows that $2n^2+n$ is odd.

Conversely, we will prove the contrapositive of the converse. We must assume to this end that n is an even integer. By definition of an even integer, we have that $n=2k$ for some integer k. Consequently, we find that $2n^2+n=2(2k)^2+(2k)=8k^2+2k=2(4k^2+k)$. Considering that $4k^2+k$ is an integer, it follows that $2n^2+n$ is an even integer, as desired. □

Example 3.2.7. Prove that if x and y are real numbers such that $x^3 + xy^2 \leq y^3 + x^2y$, then $x \leq y$.

Proof. We will prove the contrapositive statement. We must assume to this end that x and y are real numbers such that $x > y$. By multiplying this inequality by the non-negative real number y^2, we find that $xy^2 \geq y^3$. Likewise, by multiplying this inequality by the non-negative real number x^2, we find that $x^3 \geq x^2y$. By adding these two inequalities, we conclude that $x^3 + xy^2 \geq y^3 + x^2y$. Considering that $x > y$, one of the real numbers x or y must be nonzero, hence one of the inequalities $xy^2 \geq y^3$ or $x^3 \geq x^2y$ must be strict. Consequently, we conclude that $x^3 + xy^2 > y^3 + x^2y$, as desired. □

3.3 Proof by Cases

Consider any open sentence $P(x_1, \ldots, x_n)$ involving the n variables $x_1, \ldots, x_n$ with domain S. **Proof by cases** is an exhaustive proof technique that exploits some "finiteness property" of the set S. Often, this "finiteness property" of S can be realized as one of the following situations.

(a.) We have that S is finite and it is possible to prove the statement $P(x)$ for each element $x \in S$.

(b.) We have that S admits a finite partition $S = S_1 \cup S_2 \cup \cdots \cup S_n$ and it is possible to prove the statement $P(x)$ for each element $x \in S_i$ for each integer $1 \leq i \leq n$.

Concretely, we will illustrate the proof by cases by completing the following typical examples.

Example 3.3.1. Consider the finite subset $S = \left\{1, \sqrt{2}, 2\sqrt{2}\right\}$ of $\mathbb{R}$. Prove that for every element $x \in S$, there exists an element $y \in S$ such that $x - y \leq 0$ and $x^2 + y^2$ is a perfect square.

Proof. We may consider the following three cases.

1.) If $x = 1$, then observe that for $y = 2\sqrt{2}$, we have that $x < y$ so that $x - y \leq 0$ and

$$x^2 + y^2 = (1)^2 + \left(2\sqrt{2}\right)^2 = 1 + 8 = 9 = 3^2.$$

2.) If $x = \sqrt{2}$, then observe that for $y = \sqrt{2}$, we have that $x = y$ so that $x - y \leq 0$ and

$$x^2 + y^2 = \left(\sqrt{2}\right)^2 + \left(2\sqrt{2}\right)^2 = 2 + 2 = 4 = 2^2.$$

3.) If $x = 2\sqrt{2}$, then observe that for $y = 2\sqrt{2}$, we have that $x = y$ so that $x - y \leq 0$ and

$$x^2 + y^2 = \left(2\sqrt{2}\right)^2 + \left(2\sqrt{2}\right)^2 = 8 + 8 = 16 = 4^2.$$

We have exhausted all possibilities for an element $x \in S$, hence our proof is complete. □

Example 3.3.2. Consider the finite subset $S = \{2, 3, 4\}$ of $\mathbb{N}$. Prove that for every element $x \in S$ such that $x^2(x-1)^2/4$ is even, we have that $x^2(x+1)^2/4$ is even.

Proof. We may consider the following three cases.

1.) If $x = 2$, then $x^2(x-1)^2/4 = 2^2(2-1)^2/4 = 1$ is not even, so we proceed to the next case.

2.) If $x = 3$, then $x^2(x-1)^2/4 = 3^2(3-1)^2/4 = 9$ is not even, so we proceed to the next case.

3.) If $x = 4$, then each of $x^2(x-1)^2/4$ and $x^2(x+1)^2/4$ have a factor of 4, so they are even.

We have exhausted all possibilities for an element $x \in S$, hence our proof is complete. □

Essentially, a proof by cases for an open sentence $P(x)$ with finite domain S amounts to verifying $P(x)$ for each element $x \in S$. Consequently, there are at most $|S|$ cases in this situation.

We turn our attention next to open sentences that involve integers or elements of other infinite sets. Recall that an integer is either even or odd but not both; the quality that an integer is even or odd is called the **parity** of the integer. Consequently, if we encounter a statement involving an integer, then it is possible to construct a proof by cases by inspecting the situation when n is even and when n is odd separately. We illustrate this idea concretely in the following three examples.

Example 3.3.3. Prove that for every integer n, we have that $n^2 + 3n - 4$ is even.

Proof. We may consider the following two cases.

1.) By definition, if n is even, then there exists an integer k such that $n = 2k$. Consequently, we have that $n^2 + 3n - 4 = (2k)^2 + 3(2k) - 4 = 4k^2 + 6k - 4 = 2(2k^2 + 3k - 2)$. Considering that $2k^2 + 3k - 2$ is an integer, we conclude that $n^2 + 3n - 4$ is an even integer.

2.) By definition, if n is odd, then there exists an integer k such that $n = 2k+1$. Consequently, we have that $n^2+3n-4 = (2k+1)^2+3(2k+1)-4 = (4k^2+4k+1)+(6k+3)-4 = 2(2k^2+5k)$. Considering that $2k^2 + 10k$ is an integer, we conclude that $n^2 + 3n - 4$ is an even integer.

We have exhausted all possibilities for the parity of the integer n, hence our proof is complete. □

Example 3.3.4. Prove that any integers x and y have the same parity if and only if $x + y$ is even.

Proof. We will first prove the statement that if x and y are any integers of the same parity, then $x + y$ is even. Consider toward this end the following two cases.

1.) By definition, if the integers x and y are both even, then there exist integers k and ℓ satisfying that $x = 2k$ and $y = 2\ell$. Consequently, we have that $x + y = 2k + 2\ell = 2(k + \ell)$. Considering that $k + \ell$ is an integer, we conclude that $x + y$ is even, as desired.

2.) By definition, if the integers x and y are both odd, then there exist integers k and ℓ such that $x = 2k+1$ and $y = 2\ell+1$. Consequently, we have that $x+y = (2k+1)+(2\ell+1) = 2(k+\ell+1)$. Considering that $k + \ell + 1$ is an integer, we conclude that $x + y$ is even, as desired.

We have exhausted all possibilities for the parity of the integers x and y, hence the statement holds.

Conversely, we will prove the contrapositive of the statement that if $x + y$ is even, then the integers x and y have the same parity. Explicitly, we will demonstrate that if x and y have opposite parity, then the integer $x + y$ is odd. We may assume **without loss of generality** that x is even and y is odd. Consequently, there exist integers k and ℓ such that $x = 2k$ and $y = 2\ell + 1$. Observe that $x + y = 2k + (2\ell + 1) = 2(k + \ell) + 1$. Because $k + \ell$ is an integer, the integer $x + y$ is odd. □

Remark 3.3.5. We reflect here on two important features of the proof of Example 3.3.4.

1.) First, it is important to note that the biconditional ("if and only if") statement was proved by using a proof by cases for one direction of the biconditional (the "only if" direction) and using a proof by contrapositive for the other direction (the "if" direction). Often, we will be required to use multiple proof techniques in tandem to write a satisfactory proof of a proposition.

2.) We have introduced in the body of the proof of Example 3.3.4 an important phrase in the trade of mathematical writing: "without loss of generality." Essentially, what this means is that the author is asserting to the reader that there is no need to distinguish between the two variables x and y in the above proof: indeed, it does not matter if x is even and y is odd or vice-versa; the result would work the same if the names (or roles) of x and y were swapped. One way to think about the phrase "without loss of generality" is that it can be useful to save the author and the reader precious time if the same (or at least a similar) proof could be used for the other cases that would be necessary to consider in the proof by cases; therefore, one might instead use the phrase, "A similar proof can be used to establish the result."

Example 3.3.6. Prove that $3x + 5y + 7z$ is odd if exactly two of the integers x, y, z are even.

Proof. Observe that $3x + 5y + 7z = 2(x + 2y + 3z) + x + y + z$. Consequently, it suffices to prove that $x + y + z$ is odd by Example 3.3.4: indeed, if $x + y + z$ were even, then $3x + 5y + 7z$ would be even. Consequently, we may assume without loss of generality that x and y are even and z is odd. Explicitly, suppose that there exist integers k, ℓ, and m such that $x = 2k$, $y = 2\ell$, and $z = 2m + 1$. We have that $x + y + z = 2k + 2\ell + (2m + 1) = 2(k + \ell + m) + 1$, hence $x + y + z$ is odd. □

Remark 3.3.7. Observe that the proof of Example 3.3.6 is quite clever and drastically reduces the amount of work required to prove the statement. We immediately used the result of Example 3.3.4 to reduce the problem at hand to simply demonstrating that $x + y + z$ is odd whenever exactly two of the integers x, y, and z are even; then, because each of the integers x, y, and z appeared as terms of the sum, there was no need to distinguish between them, so we could appeal to the phrase "without loss of generality" to reduce a proof potentially involving three cases to just one case. Compare this with the amount required to write a proof for Example 3.3.4 with three cases.

Last, a proof by cases can sometimes be used to handle statements involving the union of sets.

Example 3.3.8. Prove that if A, B, and C are sets with $x \in A \cup B$, then $x \in A \cup C$ or $x \in B \cup C$.

Proof. Observe that $x \in A \cup B$ if and only if $x \in A$ or $x \in B$. Consequently, there are two cases.

1.) If $x \in A$, then $x \in A \cup C$ by definition of the set union.

2.) If $x \in B$, then $x \in B \cup C$ by definition of the set union.

Either way, we conclude that $x \in A \cup C$ or $x \in B \cup C$, as desired. □

Remark 3.3.9. We note that in the previous proof, the sets A and B are analogous: indeed, our ultimate objective is to verify a disjunctive statement in the sets $A \cup C$ and $B \cup C$. Consequently, it is possible to use the phrase "without loss of generality" rather than appeal to a proof by cases.

Proof. (Example 3.3.8) Considering that $x \in A \cup B$ if and only if $x \in A$ or $x \in B$, we may assume without loss of generality that $x \in A$. We conclude that $x \in A \cup C$, as desired. □

3.4 Counterexamples

Before we are able to prove a statement, we must first deduce that it is true. Often, this amounts to computing several examples to convince ourselves that the statement is valid. Best case scenario, either this practice reveals the nature of a potential proof or a **counterexample** is revealed to us. By counterexample, we mean an explicit instance for which the statement in question is false.

Example 3.4.1. Consider the following conditional statement.

$$P(n)\colon \text{ If } n \text{ is an integer, then } 5n+4 \text{ is even.}$$

Considering that $5(1)+4=9$ is odd, the conditional statement $P(1)$ is false. Consequently, the integer $n_0=1$ provides a counterexample and illustrates that $P(n)$ is not a true statement.

Example 3.4.2. Consider the following conditional statement.

$$P(x)\colon \text{ If } x \text{ is a real number, then } x-e^x>0.$$

Considering that $0-e^0=0-1=-1\leq 0$, the conditional statement $P(0)$ is false. Consequently, the real number $x_0=0$ provides a counterexample and illustrates that $P(x)$ is not a true statement.

Example 3.4.3. Consider the following conditional statement.

$$P(x)\colon \text{ If } x \text{ is a real number, then } \cot^2(x)+1=\csc^2(x).$$

Considering that $\cot(0)$ and $\csc(0)$ are undefined, the conditional statement $P(0)$ is false. Consequently, the real number $x_0=0$ provides a counterexample and illustrates that $P(x)$ is not true.

Counterexamples can be astonishingly difficult to determine in many cases: in fact, it is a highly active area of mathematical research to find counterexamples to certain statements of particular interest. Explicitly, the desire for a counterexample is illuminated by the following observation. Consider an open sentence $P(x)$ in a variable x with domain S. Recall that the quantified statement "$\forall x\in S,\ P(x)$" is true if and only if $P(x)$ is true for all elements $x\in S$. By Theorem 2.6.13, if we wish to **disprove** the statement "$\forall x\in S,\ P(x)$" (or show that this statement is false), it suffices to exhibit an element $x_0\in S$ such that $P(x_0)$ is false. By name, this element x_0 is a counterexample.

Example 3.4.4. Disprove the following conditional statement.

$$P(x)\colon \text{ If } x \text{ is a real number, then } \frac{x^3+1}{x^3-1}=\frac{x^2-x+1}{x^2+x+1}.$$

Solution. Observe that if $x=1$, then $x^3-1=0$, hence the left-hand fraction in the statement of $P(x)$ is undefined. Consequently, $x=1$ is a counterexample to the conditional statement $P(x)$. ◇

Example 3.4.5. Disprove the following conditional statement.

$$P(x,y)\colon \text{ If } x \text{ and } y \text{ are real numbers, then } x^2-4xy+y^2>0.$$

Solution. Observe that if $x=1$ and $y=1$, then $x^2-4xy+y^2=1-4+1=-2\leq 0$. Consequently, the ordered pair $(x,y)=(1,1)$ is a counterexample to the conditional statement $P(x,y)$. ◇

Example 3.4.6. Disprove the following conditional statement.

$$P(x, y, z)\text{: If } x, \text{ } y, \text{ and } z \text{ are positive real numbers, then } (x^y)(x^z) = x^{yz}.$$

Solution. Observe that if $x = 2$, $y = 1$, and $z = 3$, then $(x^y)(x^z) = (2^1)(2^3) = 16$ and $x^{yz} = 8$, hence the ordered triple $(x, y, z) = (2, 1, 3)$ is a counterexample to the conditional statement $P(x, y, z)$. $\diamond$

Be sure to make note of the form we use when solving a problem that asks us to disprove something: we begin with an italicized "Solution" and a period; we exhibit an explicit counterexample to the statement; and we conclude with an empty diamond $\diamond$ to signify the conclusion of our solution.

3.5 Proof by Contradiction

Last but certainly not least, the **proof by contradiction** (or **reductio ad absurdum**) rounds out the tools that we will most often use in mathematical proofs. Essentially, the proof by contradiction constitutes a valid proof technique by a combination of the Law of the Excluded Middle, the Law of Non-Contradiction, and Table 2.14. We bear out the details in two cases of particular interest. We will first assume toward this end that P is a statement that we wish to prove is true.

1.) By the Law of the Excluded Middle, either P is true or P is false.

2.) By the Law of Non-Contradiction, if P is not false, then P is true.

3.) Consequently, in order to demonstrate that P is true, it suffices to prove that P is not false. We assume toward this end that P is in fact false, i.e., we assume that $\neg P$ is true.

4.) By some properties of $\neg P$, it might be possible to derive a contradiction C, i.e., a statement C that is false with respect to all possible truth inputs. Crucially, the contradiction C could reveal itself as a direct consequence of the assumption $\neg P$ or it might be possible to derive a contradiction C from some other known facts (e.g., definitions, propositions, and theorems).

5.) We conclude that the conditional statement $\neg P \Rightarrow C$ is true. But C is false, so $\neg P$ must be false; therefore, our initial assumption that P is false is untenable, so P must be true.

Often, a proof by contradiction is desirable to prove a conditional statement $P \Rightarrow Q$. We outline next how a proof by contradiction for such a statement could be carried out and why it is valid.

1.) By the Law of the Excluded Middle, either $P \Rightarrow Q$ is true or $P \Rightarrow Q$ is false.

2.) By the Law of Non-Contradiction, if $P \Rightarrow Q$ is not false, then $P \Rightarrow Q$ is true.

3.) Consequently, it suffices to prove that $P \Rightarrow Q$ is not false. By Table 2.7, we must show that if Q is false, then P is false. We assume toward this end that Q is false and P is true.

4.) Like in the case of the proof by contradiction discussed above, it might be possible to derive a contradiction C from some properties of $\neg Q$; this would entail that $\neg Q \Rightarrow C$ is true.

5.) Observe that if C is false and $\neg Q \Rightarrow C$ is true, then $\neg Q$ is false; therefore, our assumption that Q is false is untenable, hence Q must be true so that $P \Rightarrow Q$ is true.

We point out at this time that the writer should always mention in the first line of the proof the proof technique that will be used. Best practice dictates (in the case of a proof by contradiction) that this is achieved using the phrase, "Suppose on the contrary that P is true and Q is false" or, "We will assume toward a contradiction that P is true and Q is false." Even more, the writer should take care to point out exactly what contradiction is derived in a proof by contradiction.

Example 3.5.1. Prove that there is no smallest integer.

Proof. Suppose on the contrary that n is the smallest integer. Observe that $n-1$ is an integer. By adding n to both sides of the inequality $-1 < 0$, we find that $n-1 < n$. But this is a contradiction: if n is the smallest integer, there can be no integer less than n. Our assumption that there exists a smallest integer is therefore untenable, hence we conclude that there is no smallest integer. □

Example 3.5.2. Prove that no integer is both even and odd.

Proof. Suppose on the contrary that n is an even integer that is also odd. By definition of an even integer, there exists an integer k such that $n = 2k$. By definition of an odd integer, there exists an integer ℓ such that $n = 2\ell + 1$. Considering that $n = n$ is a tautology, it follows that $2k = 2\ell + 1$ so that $1 = 2k - 2\ell = 2(k - \ell)$. By dividing both sides of this equation by 2, we find that $k - \ell = \frac{1}{2}$. But this is a contradiction: the difference of two integers is an integer, but the rational number $\frac{1}{2}$ is not an integer. Our assumption that there exists an integer that is both even and odd is therefore untenable, hence we conclude that no integer is both even and odd. □

Example 3.5.3. Prove that no even integer is the sum of three odd integers.

Proof. Suppose on the contrary that n is an even integer that is the sum of three odd integers a, b, and c. By definition of an odd integer, we have that $a = 2k + 1$, $b = 2\ell + 1$, and $c = 2m + 1$ for some integers k, ℓ, and m. Considering that $n = a + b + c$, it follows that

$$n = (2k + 1) + (2\ell + 1) + (2m + 1) = 2(k + \ell + m + 1) + 1,$$

hence n is odd. But this is a contradiction: by Example 3.5.2, we have that no integer is both even and odd. Our assumption that there exists an even integer that is the sum of three odd integers is untenable, hence we conclude that no even integer is the sum of three odd integers. □

Example 3.5.4. Prove that if a, b, and c are integers such that $a^2 + b^2 = c^2$, then a or b is even.

Proof. Suppose on the contrary that a and b are both odd integers. By definition of an odd integer, we have that $a = 2k + 1$ and $b = 2\ell + 1$ for some integers k and ℓ. Consequently, we find that

$$c^2 = a^2 + b^2 = (2k + 1)^2 + (2\ell + 1)^2 = (4k^2 + 4k + 1) + (4\ell^2 + 4\ell + 1) = 2(2k^2 + 2k + 2\ell^2 + 2\ell + 1).$$

Considering that $2k^2 + 2k + 2\ell^2 + 2\ell + 1$ is an integer, we conclude that c^2 is even so that c is even. By definition of an even integer, we have that $c = 2m$ for some integer m so that

$$4m^2 = (2m)^2 = c^2 = 2(2k^2 + 2k + 2\ell^2 + 2\ell + 1).$$

Cancelling one factor of 2 from both sides of this equation yields that

$$2m^2 = 2k^2 + 2k + 2\ell^2 + 2\ell + 1 = 2(k^2 + k + \ell^2 + \ell) + 1.$$

But this is a contradiction: the left-hand side shows an even integer, but the right-hand side shows an odd integer. Our assumption that a and b are odd is untenable, hence a or b is even. □

Example 3.5.5. Prove that if x is even and y is odd, then $x^2 + 2y^2$ is not divisible by 4.

Proof. Suppose on the contrary that x is an even integer and y is an odd integer such that $x^2 + 2y^2$ is divisible by 4. By definition of the parity of an integer, we have that $x = 2k$ and $y = 2\ell + 1$ for some integers k and ℓ. Consequently, we may simplify the expression $x^2 + 2y^2$ to find that

$$x^2 + 2y^2 = (2k)^2 + (2\ell + 1)^2 = 4k^2 + 2(4\ell^2 + 4\ell + 1) = 4(k^2 + 2\ell^2 + 2\ell) + 2.$$

By assumption that $x^2 + 2y^2$ is divisible by 4, there exists an integer m such that $x^2 + 2y^2 = 4m$. Combined with our previous displayed equation, this yields that

$$4m = 4(k^2 + 2\ell^2 + 2\ell) + 2,$$

from which we deduce that $2 = 4m - 4(k^2 + 2\ell^2 + 2\ell) = 4(m - k^2 - 2\ell^2 - 2\ell)$. By cancelling a factor of 2 from both sides, we find that $1 = 2(m - k^2 - 2\ell^2 - 2\ell)$. But this is a contradiction: the integer 1 is odd, so it cannot be divisible by 2 by Example 3.5.2. Our assumption that x is an even integer and y is an odd integer such that $x^2 + 2y^2$ is divisible by 4 is therefore untenable, hence we conclude that if x is an even integer and y is an odd integer, then $x^2 + 2y^2$ is not divisible by 4. □

Example 3.5.6. Prove that $\sqrt{2}$ is irrational.

Proof. Suppose on the contrary that $\sqrt{2}$ is rational. By definition of a rational number, there exist integers a and b such that b is nonzero; a and b possess no common factors other than ± 1; and

$$\sqrt{2} = \frac{a}{b}.$$

By squaring both sides of this equation and clearing the denominator, we find that

$$a^2 = 2b^2.$$

Consequently, the integer a^2 is even. Considering that the square of an integer is even if and only if that integer is even, it follows that a is even so that $a = 2k$ for some integer k. By substituting this identity back into our above displayed equation, we find that

$$4k^2 = (2k)^2 = a^2 = 2b^2.$$

Cancelling a factor of 2 from both sides yields that b^2 is an even integer since

$$b^2 = 2k^2.$$

By the same rationale as before, we conclude that b is even so that $b = 2\ell$ for some integer ℓ. But this is a contradiction: we had originally assumed that a and b possess no common factors other than ± 1, but if a and b are both even, then they have a common factor of 2. Our assumption that $\sqrt{2}$ is rational is therefore untenable, hence we conclude that $\sqrt{2}$ is irrational. □

3.6 Existence Proofs

Complementary to counterexamples, proving the existence of certain mathematical objects or structures with desirable properties is also a foremost concern throughout mathematics. We remind the reader at this point that an existence statement is a quantified statement of the form

$$\exists x \in S,\, P(x)\colon \text{ There exists an element } x \in S \text{ such that } P(x).$$

for some open sentence $P(x)$ in a variable x with domain S. Consequently, in order to determine the verity of an existence statement, it suffices to provide an explicit example of an element $x_0 \in S$ such that $P(x_0)$ is true; if this is possible, then the attendant proof of the existence statement is called a **constructive proof** because the element $x_0 \in S$ is often "constructed" or produced by explicitly performing some algebraic manipulation or computation. We provide some examples below.

Example 3.6.1. Prove that there exists an integer whose cube is equal to its square.

Solution. Before we prove this existence statement, we may find it beneficial to write the statement in symbols. Observe that if n is an integer, then n^3 is its cube and n^2 is its square. Consequently,

$$P(n)\colon \text{ The integer } n \text{ satisfies that } n^3 = n^2.$$

is the open sentence that n is an integer whose cube is equal to its square. Ultimately, we are trying to prove the following existentially quantified statement in the variable n over the domain $\mathbb{Z}$.

$$\exists n \in \mathbb{Z},\, P(n)\colon \text{ There exists an integer } n \text{ such that } n^3 = n^2.$$

Observe that if $n^3 = n^2$, then $n^3 - n^2 = 0$ so that $n^2(n-1) = 0$. By the Zero Product Property, it follows that $n = 0$ or $n = 1$. Either one of these integers provides an explicit solution to the integer equation $n^3 = n^2$, hence we have the ingredients to write a constructive proof for the statement. $\diamond$

Proof. Observe that the integer $n = 1$ satisfies that $n^3 = 1^3 = 1 = 1^2 = n^2$, and the claim holds. $\square$

Example 3.6.2. Prove that there exist real numbers x and y such that $(x+y)^2 = x^2 + y^2$.

Solution. Before we determine a proof of the statement, we note that we seek to establish the verity of the following existential statement in the variables x and y over the domain $\mathbb{R}$.

$$\exists x, y \in \mathbb{R},\, P(x,y)\colon \text{ There exist real numbers } x \text{ and } y \text{ such that } (x+y)^2 = x^2 + y^2.$$

Observe that if $(x+y)^2 = x^2 + y^2$, then $x^2 + 2xy + y^2 = x^2 + y^2$ so that $2xy = 0$. By the Zero Product Property, it follows that $x = 0$ or $y = 0$. Either way, the statement that $(x+y)^2 = x^2 + y^2$ will be true for any value of the variables x and y so long as one of them is zero: indeed, if $y = 0$, then $(x+y)^2 = (x+0)^2 = x^2 = x^2 + 0^2 = x^2 + y^2$. We have the makings of a constructive proof. $\diamond$

Proof. Observe that the real numbers $x = 1$ and $y = 0$ satisfy that

$$(x+y)^2 = (1+0)^2 = 1^2 = 1^2 + 0^2 = x^2 + y^2.$$

Consequently, the statement in question holds, and our proof is complete. $\square$

Example 3.6.3. Prove that $f(x) = x^5 + x^4 + x^3 + x^2 + x + 1$ has at least one real root.

Solution. By definition of a root of a function, the statement we are tasked to prove is as follows.

$$\exists x \in \mathbb{R},\ P(x)\colon\ \text{There exists a real number } x \text{ such that } f(x) = 0.$$

Observe that $f(-1) = (-1)^5 + (-1)^4 + (-1)^3 + (-1)^2 + (-1) + 1 = -3 + 3 = 0$, hence a direct proof is possible because we have found an explicit example of a real root of $f(x)$. ◇

Proof. Observe that the real number $x = -1$ is a root of $f(x)$ since we have that

$$f(-1) = (-1)^5 + (-1)^4 + (-1)^3 + (-1)^2 + (-1) + 1 = -3 + 3 = 0.$$ □

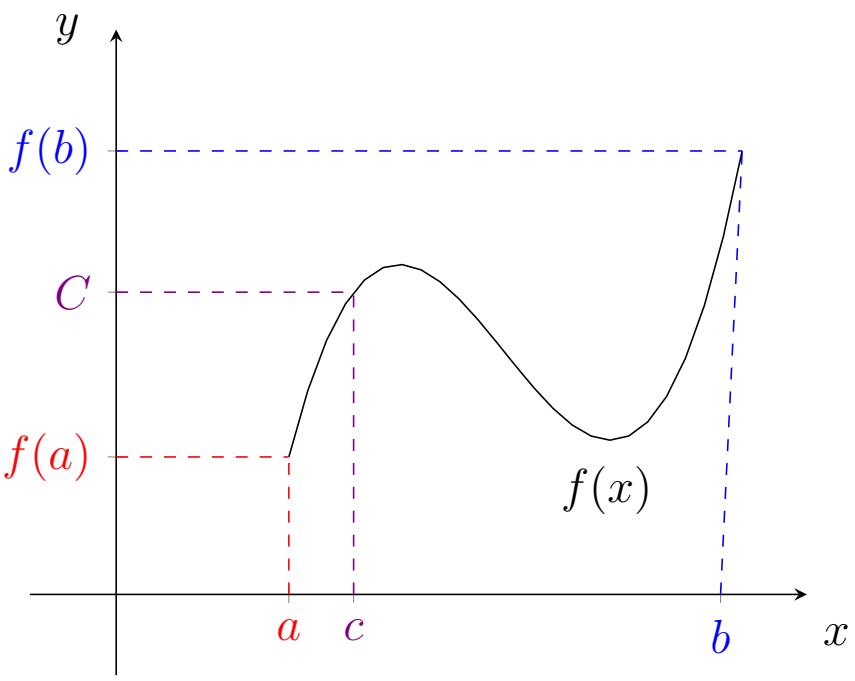
Remark 3.6.4. We make an important and necessary observation about the serendipitous nature of the existence proof provided in Example 3.6.3. Exactly how did we stumble upon the real number $x = -1$, and why did we suspect that it is a root of the polynomial $f(x) = x^5 + x^4 + x^3 + x^2 + x + 1$? Unfortunately, this was simply a lucky coincidence — the product of years of schema and knowing where to look. One natural starting point in the aforementioned example is to begin by plugging in integer values of x near zero. Plugging in $x = 0$ yields that $f(0) = 0$, and plugging in $x = 1$ yields that $f(1) = 6$ — both failures. We were exceedingly lucky that our next guess $x = -1$ worked.

Generally, the roots of a real function $f(x)$ are seriously difficult to compute. By the Quadratic Formula, the roots of any real function of the form $f(x) = ax^2 + bx + c$ with a nonzero are known; there are also the Cubic Formula and the Quartic Formula, but these are typically not taught, and students are not expected to know them (the author freely admits to not knowing them, either). Beyond that, it is a landmark result of Galois Theory that there is no closed form expression for the roots of a real polynomial of degree at least five. Consequently, there is little hope for deducing the roots of a polynomial of degree five or larger — let alone trying to find the roots of a real function that is not a polynomial (other than certain trigonometric, inverse trigonometric, or logarithmic functions) — for students in this course without specialized knowledge (such as the Newton-Raphson Method or other recursive numerical methods for finding roots of differentiable functions).

Even still, using elementary calculus, there is a way to determine existence of roots of continuous functions without ever knowing exactly what those roots are! Before we provide a proof along these lines, we must first recall the following important fact about continuous functions from Calculus I.

Theorem 3.6.5 (Intermediate Value Theorem)**.** *Every real univariate function $f : D_f \to \mathbb{R}$ with domain $D_f \subseteq R$ that is continuous on a closed and bounded interval $[a, b]$ satisfies that for every real number C between $f(a)$ and $f(b)$, there exists a real number c such that $a \leq c \leq b$ and $f(c) = C$.*

Concretely, the Intermediate Value Theorem states any every real function $f(x)$ that is continuous on a closed and bounded interval $[a, b]$ achieves every possible y-value between $f(a)$ and $f(b)$ for some x-value between a and b. Graphically, the intuition is that a continuous function can be represented visually by drawing without lifting one's pencil, hence as the curve $y = f(x)$ is traced out from $x = a$ to $x = b$ along the x-axis, every real number along the y-axis between $f(a)$ and $f(b)$ must correspond to some point on the x-axis. Consider the picture below for an illustration.



Consequently, the upshot is that in order to prove the existence of roots of a continuous function $f(x)$, we may find real numbers a and b such that $f(a)$ and $f(b)$ have opposite sign, i.e., $f(a) < 0$ and $f(b) > 0$ (or vice-versa); then, because $f(x)$ is a continuous function such that $f(a)$ and $f(b)$ have opposite sign, there must exist a real number c such that $a \leq c \leq b$ and $f(c) = 0$. We refer to such a proof of the existence of the roots of a continuous function as a **non-constructive proof**: in fact, we are not explicitly exhibiting the roots of the function. We are instead simply relying on the Intermediate Value Theorem to conclude that some root must exist. Generally, a non-constructive proof relies on some well-known fact, theorem, or definition. Consequently, a non-constructive proof may not be direct. We conclude this section with several examples of non-constructive proofs.

Example 3.6.6. Prove that $f(x) = x^5 + x^4 + x^3 + x^2 + x + 1$ has at least one real root.

Proof. Observe that the polynomial $f(x)$ is a continuous function. Considering that $f(-2) = -21$ and $f(0) = 1$, by the Intermediate Value Theorem, there exists a real number c such that $-2 < c < 0$ and $f(c) = 0$. By definition, the real number c is a root of the polynomial $f(x)$, as desired. □

Example 3.6.7. Prove that $f(x) = e^x - 3x$ has at least one real root.

Proof. Observe that $f(x)$ is a continuous function since it is the difference of the continuous functions e^x and $3x$. Considering that $f(1) = e - 3 < 0$ and $f(0) = 1 > 0$, by the Intermediate Value Theorem, there exists a real number c such that $0 < c < 1$ and $f(c) = 0$, as desired. □

Example 3.6.8. Prove that $\cos(x) - \sin(x) = \frac{1}{2}$ for some real number x such that $0 \leq x \leq \frac{\pi}{4}$.

Proof. Consider the function $f(x) = \cos(x) - \sin(x)$. Observe that $f(x)$ is continuous because it is the difference of the continuous functions $\cos(x)$ and $\sin(x)$. Considering that

$$f(0) = 1 - 0 = 1 \geq 0 = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} = f\left(\frac{\pi}{4}\right),$$

by the Intermediate Value Theorem, there exists a real number $0 \leq c \leq \frac{\pi}{4}$ and $f(c) = \frac{1}{2}$. □

Example 3.6.9. Prove that some digit appears infinitely often in the decimal expansion of π.

Solution. Before we outline a proof strategy, we note the novelty of the example. We cannot use the Intermediate Value Theorem because the statement does not involve a continuous function. We cannot verify directly that some digit appears infinitely often in the decimal expansion of π because we cannot check the infinitely many digits in the decimal expansion of π. Consequently, there is no hope for a constructive proof. Our statement is not conditional, so there is no contrapositive. We require a proof by contradiction! Observe that the negation of the above statement is, "No digit appears infinitely often in the decimal expansion of π," or, "Every digit in the decimal expansion of π appears finitely many times." Considering that π is irrational, this statement cannot be true, so the statement we seek to prove must be true by the Law of the Excluded Middle and the Law of Non-Contradiction. We did not check any digits of π, hence this is a non-constructive proof. $\diamond$

Proof. On the contrary, suppose that every digit in the decimal expansion of π appears finitely many times. Considering that the digits in the decimal expansion of any real number are $0, 1, 2, \ldots, 9$, the decimal expansion of π contains at most ten digits. But this is a contradiction: if the decimal expansion of π is finite, then π is a rational number; however, π is irrational! Consequently, our assumption that every digit in the decimal expansion of π appears finitely many times is untenable, hence we conclude that some digit in the decimal expansion of π appears infinitely many times. $\square$

3.7 Chapter 3 Overview

Often, we seek to prove conditional statements of the form $P \Rightarrow Q$. Generally, a **vacuous proof** amounts to showing that P is false. Conversely, a **trivial proof** follows by showing that Q is true. If neither P is false nor Q is true, then a **direct proof** follows by assuming that P is true and deducing that Q is true. We refer to this rule of inference as **modus ponens**. Conditional statement $P \Rightarrow Q$ is logically equivalent to its **contrapositive** $\neg Q \Rightarrow \neg P$ (see Table 2.15 and the subsequent Proposition 2.5.2). Consequently, a **proof by contrapositive** follows by assuming that $\neg Q$ is true and deducing that $\neg P$ is true. We refer to this rule of inference as **modus tollens**. Concretely, if $\neg P$ can be deduced from $\neg Q$, then we may construct a proof by contrapositive; on the other hand, if Q can be deduced from P, then we may construct a direct proof. Otherwise, we seek a **proof by contradiction** by assuming that P is true and Q is false and deriving a contradiction using any assumption made in the context of the proof or any definition or well-known fact. Proof by contradiction can be deduced from the Law of the Excluded Middle, the Law of Non-Contradiction, and the logical equivalence of the statements $\neg(P \Rightarrow Q)$ and $P \wedge \neg Q$ (see Table 2.14).

3.8 Chapter 3 Exercises

Exercise 3.8.1. Use Example 2.11 to prove that if Bob placed in the top two in a cycling race on Saturday and he did not place second, then Bob must have placed first.

Exercise 3.8.2. Construct a proof by contradiction to demonstrate that if Bob placed in the top two in a cycling race on Saturday and he did not place second, then Bob must have placed first. Cite any theorems or laws of inference by name that you use in your proof.

Chapter 4

Proofs in the Wild

Once we are satisfactorily acquainted with the basic proof strategies outlined in the previous chapter, we may consider examples and write proofs in a variety of familiar mathematical contexts. We aim throughout this chapter to employ the techniques of the previous three chapters as they pertain to the study of combinatorics, elementary number theory, modern algebra, and naïve set theory.

4.1 Principle of Mathematical Induction

Consider any open sentence $P(n)$ defined for a variable n with domain $S \subseteq \mathbb{Z}$. Observe that if S admits a smallest element n_0, then we may denote $n_0 = \min\{n \mid n \in S\}$ since it is the minimum element of S. We have seen in Chapter 3 that it may be possible to prove the quantified statement $\forall n \in S$, $P(n)$ by cases; however, this may be tedious if S is finite and $|S|$ is large, and it may not be clear why the statement $P(n)$ is true even if we assume that n is either even or odd. Consequently, we may require another technique all together to demonstrate that $P(n)$ is true for all $n \in S$.

We turn our attention thus to one of the most useful proof techniques for establishing the verity of universally quantified statements defined forintegers: a **proof by induction** appeals to one of the three forms of the **Principle of Mathematical Induction**. Before we proceed to the definition, let us explore some examples of properties of integers for which a proof by induction is appropriate.

Example 4.1.1. Consider the sum of the first n consecutive odd positive integers.

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

Computing the values 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	5
$o(n)$	1	4	9	16	25

Table 4.1: the sum of first five 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 4.1.2. We have that $o(n) = n^2$ for all integers $n \geq 1$ for $o(n)$ as in Example 4.1.1.

Observe that $o(1) = 1^2$ and $o(n+1) = o(n) + (2n+1)$, hence if we were to assume that $o(n) = n^2$ for some integer $n \geq 1$, then we would conclude that $o(n+1) = n^2 + 2n + 1 = (n+1)^2$. We will soon return to validate this idea as one of the tenants of the Principle of Mathematical Induction!

Example 4.1.3. Consider the sum of the first n consecutive positive integers.

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

Computing the values 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 4.2: the sum of the first five 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 $1 + 2 + 3 + \cdots + n$ both forwards and backwards, adding each column of the sum to determine the value of $2(1 + 2 + 3 + \cdots + n)$.

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

Considering that there are n columns in this table 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.

Conjecture 4.1.4. We have that $c(n) = \frac{n(n+1)}{2}$ for all integers $n \geq 1$ for $c(n)$ as in Example 4.1.3.

Like before, we can readily verify the facts that $c(1) = 1 = \frac{1 \cdot 2}{2}$ and $c(n+1) = c(n) + (n+1)$, hence if we were to assume that $c(n) = \frac{n(n+1)}{2}$ for some integer $n \geq 1$, 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 4.1.5 (Principle of Ordinary Induction)**.** Given any integer n_0, consider any open sentence $P(n)$ defined for all integers $n \geq n_0$. We may define the following criteria.

(a.) We have that $P(n_0)$ is a true statement.

(b.) If $P(n)$ is a true statement for some integer $n \geq n_0$, then $P(n+1)$ is a true statement.

Provided that both of these statements are true, it follows that $P(n)$ is true for all integers $n \geq n_0$.

By the Principle of Ordinary Induction, we can return to prove Conjectures 4.1.2 and 4.1.4.

Proof. (Conjecture 4.1.2) Consider the following open sentence defined for all integers $n \geq 1$.

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

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.

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

(b.) 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 (a.) $P(1)$ is a true statement and (b.) $P(n+1)$ 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 4.1.4) Consider the following open sentence defined for all integers $n \geq 1$.

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

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.

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

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

$$1 + 2 + 3 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2}.$$

Considering that (a.) $P(1)$ is a true statement and (b.) $P(n+1)$ 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 abbreviate the steps in an inductive proof under the assumption that our intended audience is familiar with induction. We illustrate a typical proof by induction.

Example 4.1.6. 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 $5n = 4n + n \geq 4n + 5 \geq 2n + 1$ 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 defined for induction. Currently, we may view induction as a property of falling dominoes: (a.) if the n_0th domino falls and (b.) 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 4.1.7 (Principle of Complete Induction)**.** Given any integer n_0, consider any open sentence $P(n)$ defined for all integers $n \geq n_0$. We may define the following criteria.

(a.) We have that $P(n_0)$ is a true statement.

(b.) If $P(k)$ is a true statement for each integer $n_0 \leq k \leq n$, then $P(n+1)$ is a true statement.

Provided that both of these statements are true, it follows that $P(n)$ is true for all integers $n \geq n_0$.

Even though the criteria of the Principle of Complete Induction ostensibly appear to be much stronger than the Principle of Ordinary Induction, the two principles are in fact materially equivalent (see Exercise 4.8.4). Last, we obtain another crucial tool that is ubiquitous in mathematics.

Theorem 4.1.8 (Well-Ordering Principle)**.** *Every nonempty set of non-negative integers admits a smallest element with respect to the total order $\leq$ on the real numbers. 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 $P(k)$ is true for each integer $1 \leq k \leq n$. By definition of $P(k)$, 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 this theorem is materially equivalent to both ordinary induction and complete induction (see Exercise 4.8.5). 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.

Before we conclude this section, we provide an example using the Well-Ordering Principle.

Example 4.1.9. Prove that every integer is of the form $2^a b$ for some integers $a \geq 0$ and b odd.

Proof. Certainly, every odd integer n is of the form $n = 2^a b$ since we may take $a = 0$ and $b = n$ in this case. Conversely, if n is even, then there exists an integer k such that $n = 2k$. Observe that if k is odd, then our proof is complete since we may take $a = 1$ and $b = k$ in this case. Otherwise, we must have that k is even, hence we may repeat the same argument for k. Continuing in this manner yields a strictly decreasing sequence $|n| > |k| > \cdots > |b|$ of a positive integers that must eventually terminate in some odd integer b by the Well-Ordering Principle. We conclude that $n = 2^a b$. □

Remark 4.1.10. Canonically, any proof with non-negative integers that invokes the Well-Ordering Principle to ensure the termination of a repeating process is considered a proof by infinite descent (or Fermat's Method of Descent). Crucially, the Well-Ordering Principle ensures there is no infinite strictly decreasing sequence of non-negative integers, hence every process involving non-negative integers that ostensibly results in "infinite descent" must eventually terminate. Classically, such a proof was structured as a proof by contradiction, assuming that an infinite process were possible.

4.2 Divisibility Properties of Integers

We say that a nonzero integer a **divides** an integer b if there exists an integer q such that $b = aq$. We will write $a \mid b$ in this case, and we will typically say that b is **divisible by** a. Conversely, the **divisors** of b are the nonzero integers a that divide b. We are already familiar with this notion from Section 1.11, but for illustrative purposes, we note that the integers 1, 2, 3, 4, 6, and 12 divide 12 (i.e., the divisors of 12 are the integers 1, 2, 3, 4, 6, and 12) because $12 = 12 \cdot 1 = 2 \cdot 6 = 3 \cdot 4$. We say that an integer $p \geq 2$ is **prime** if its only positive divisors are 1 and p. Conversely, an integer $n \geq 2$ that admits positive divisors other than 1 and n is **composite**. Quite useful is the following property of the divisors of composite integers that provides a characterization of compositeness.

Theorem 4.2.1 (Factorization Criterion for Composite Integers)**.** *Given any integer $n \geq 2$, we have that n is composite if and only if $n = ab$ for some positive integers a and b such that $2 \leq a \leq n - 1$.*

Proof. Observe that if $n \geq 2$ is composite, then by definition, we may write $n = ab$ for some positive integers a and b such that a is neither 1 nor n. Considering that $b \geq 1$, it follows that $n = ab \geq a$ so that $2 \leq a \leq n - 1$ by hypothesis that a is neither 1 nor n. Conversely, if $n \geq 2$ admits positive integers a and b such that $n = ab$ and $2 \leq a \leq n - 1$, then n must be composite by definition. □

We will soon see that primes form the "building blocks" for all integers. Explicitly, every integer $n \geq 2$ can be written as a product of primes. We refer to such an expression of an integer as a product of its prime factors as the **prime factorization** of the integer. Observe that $12 = 4 \cdot 3 = 2^2 \cdot 3$ is the prime factorization of 12 and $30 = 2 \cdot 15 = 2 \cdot 3 \cdot 5$ is the prime factorization of 30. Before we are able to prove that every integer $n \geq 2$ admits a unique prime factorization, we set out to develop some basic tools for understanding divisibility of integers. Our first task is to verify the following.

Proposition 4.2.2 (Properties of Divisibility of Integers)**.** *Consider any nonzero integers a and b and any integers c and d. Each of the following properties of divisibility of integers holds.*

1.) (**Product Property**) *If $a \mid c$ or $a \mid d$, then $a \mid cd$.*

2.) (**Transitive Property**) *If $a \mid b$ and $b \mid c$, then $a \mid c$.*

3.) (**Homogeneity Property**) *If $a \mid c$ and $b \mid d$, then $ab \mid cd$.*

4.) (**Linearity Property**) *If $a \mid c$ and $a \mid d$, then $a \mid (cx + dy)$ for any integers x and y.*

Proof. We may prove each statement in turn directly by appealing to the definition of divisibility.

1.) We may assume without loss of generality that $a \mid c$ since $cd = dc$. By definition, if a divides c, then there exists an integer q such that $c = aq$. Consequently, we have that $cd = (aq)d = a(dq)$. Considering that dq is an integer because d and q are integers, we conclude that a divides cd.

2.) By definition, if $a \mid b$ and $b \mid c$, then $b = aq$ and $c = br$ for some integers q and r. We conclude by substitution that $c = br = (aq)r = a(qr)$ is divisible by a because qr is an integer.

3.) By definition, if $a \mid c$ and $b \mid d$, then $c = aq$ and $d = br$ for some integers q and r. We conclude by substitution that $cd = (aq)(br) = ab(qr)$ is divisible by ab because qr is an integer.

4.) By definition, if $a \mid c$ and $a \mid d$, then $c = aq$ and $d = ar$ for some integers q and r so that

$$cx + dy = (aq)x + (ar)y = a(qx) + a(ry) = a(qx + ry)$$

for any integers x and y. Considering that $qx + ry$ is an integer, we find that $a \mid (cx + dy)$. □

Proposition 4.2.3 (Divisibility and Absolute Value)**.** *Consider any nonzero integers a and b. Each of the following properties relating divisibility of integers and the absolute value function holds.*

1.) (**Divisibility Decreases Absolute Value**) *If $a \mid b$, then $|a| \leq |b|$.*

2.) (**Divisibility Detects Absolute Value**) *If $a \mid b$ and $b \mid a$, then $|a| = |b|$.*

Proof. We will prove the first statement; the second statement then follows from the first statement by noting that if $a \mid b$ and $b \mid a$, then $|a| \leq |b|$ and $|b| \leq |a|$ so that equality holds. By definition, if $a \mid b$, then there exists an integer q such that $b = aq$. Even more, by assumption that b is nonzero, we must have that $|q| \geq 1$. Consequently, we conclude that $|b| = |aq| = |a||q| \geq |a|$, as desired. □

Remark 4.2.4. Each of the properties of divisibility of integers we have discussed thus far can be phrased in terms of congruence modulo some nonzero integers a and b. We remind the reader that a pair of integers c and d are congruent modulo a nonzero integer n if and only if n divides $d - c$ if and only if $n \mid (d - c)$. Conventionally, if c and d are congruent modulo n, we write $d \equiv c \pmod{n}$.

1.) (**Product Property**) If $c \equiv 0 \pmod{a}$ or $d \equiv 0 \pmod{a}$, then $cd \equiv 0 \pmod{a}$.

2.) (**Transitive Property**) If $b \equiv 0 \pmod{a}$ and $c \equiv 0 \pmod{b}$, then $c \equiv 0 \pmod{a}$.

3.) (**Homogeneity Property**) If $c \equiv 0 \pmod{a}$ and $d \equiv 0 \pmod{b}$, then $cd \equiv 0 \pmod{ab}$.

4.) (**Linearity Property**) If $c \equiv 0 \pmod{a}$ and $d \equiv 0 \pmod{a}$, then $cx + dy \equiv 0 \pmod{a}$.

5.) (**Divisibility Decreases Absolute Value**) If $b \equiv 0 \pmod{a}$, then $|a| \leq |b|$.

6.) (**Divisibility Detects Absolute Value**) If $b \equiv 0 \pmod{a}$ and $a \equiv 0 \pmod{b}$, then $|a| = |b|$.

Even with this very basic notion of divisibility, there are many interesting examples to consider.

Example 4.2.5. Prove that if a, b, c are integers, a and b are nonzero, $a^2 \mid b$, and $b^3 \mid c$, then $a^6 \mid c$.

Proof. By definition, if $a^2 \mid b$, then there exists an integer q such that $b = a^2q$. Likewise, if $b^3 \mid c$, then there exists an integer r such that $c = b^3r$. Considering that $b = a^2q$, we find that $b^3 = (a^2q)^3 = a^6q^3$ so that $c = b^3r = (a^6q^3)r = a^6(q^3r)$. We conclude that $a^6 \mid c$ because q^3r is an integer. □

Example 4.2.6. Prove that for any integers a and b, if $2 \mid ab$, then $2 \mid a$ or $2 \mid b$.

Proof. We will prove the contrapositive. We may assume to this end that 2 does not divide either a or b. Consequently, the integers a and b must be odd, hence there exist integers k and ℓ such that $a = 2k + 1$ and $b = 2\ell + 1$ so that $ab = (2k + 1)(2\ell + 1) = 4k\ell + 2k + 2\ell + 1 = 2(2k\ell + k + \ell) + 1$. Considering that $2k\ell + k + \ell$ is an integer, we conclude that 2 does not divide ab. □

Remark 4.2.7. Even without the notion of divisibility of integers, we could have established the result of the preceding example using the notion of parity of integers: indeed, we claim that if ab is even for some integers a and b, then a or b must be even. Compare with the third line of the proof.

Example 4.2.8. Prove that if n is an integer such that $7 \mid 4n$, then $7 \mid n$.

Proof. By definition, if $7 \mid 4n$, then there exists an integer q such that $4n = 7q$. Considering that $4n$ is even, we must have that q is even; otherwise, if q were odd, then $q = 2k + 1$ for some integer k so that $4n = 7q = 7(2k + 1) = 2(7k + 3) + 1$ is odd — a contradiction. Consequently, there exists an integer k such that $q = 2k$ and $4n = 7q = 14k$. Cancelling one factor of 2 from each side of this identity yields that $2n = 7k$ so that $2 \mid 7k$. By Example 4.2.6, we conclude that $2 \mid k$. Consequently, there exists an integer ℓ such that $k = 2\ell$ and $q = 2k = 4\ell$. Considering that $4n = 7q$, we find that $4n = 7(4\ell)$. Cancelling one factor of 4 from each side of this identity yields $n = 7\ell$ so that $7 \mid n$. □

Remark 4.2.9. Examples 4.2.6 and 4.2.8 can be generalized to demonstrate that for any integers a and b and any prime p, we have that $p \mid ab$ if and only if $p \mid a$ or $p \mid b$ (see Exercise 4.8.6).

Example 4.2.10. Prove that if n if an integer such that $2 \mid (n^2 + 3)$, then $4 \mid (n^2 + 3)$.

Proof. By definition, if $2 \mid (n^2 + 3)$, then there exists an integer k such that $n^2 + 3 = 2k$, hence we have that $n^2 = 2k - 3 = 2(k - 2) + 1$ is odd so that n is odd. Consequently, we have that

$$n^2 + 3 = (2\ell + 1)^2 + 3 = (4\ell^2 + 4\ell + 1) + 3 = 4(\ell^2 + \ell + 1)$$

for some integer ℓ. Considering that $\ell^2 + \ell + 1$ is an integer, we conclude that $4 \mid (n^2 + 3)$. □

Given any nonzero integers a and b, we say that a nonzero integer c is a **common divisor** of a and b if and only if $c \mid a$ and $c \mid b$, i.e., c divides a and c divides b. We distinguish among all common divisors of a and b the unique **greatest common divisor** $d = \gcd(a, b)$ of a and b satisfying that

(a.) $d \mid a$ and $d \mid b$, i.e., d is a common divisor of a and b and

(b.) if c is any common divisor of a and b, then $c \mid d$.

Consequently, $\gcd(a, b)$ is the "largest" common divisor of a and b with respect to divisibility.

Example 4.2.11. Consider the integers $a = 12$ and $b = 30$. By writing down the prime factorizations of a and b, their greatest common divisor can be easily determined. Observe that $12 = 2^2 \cdot 3$ and $30 = 2 \cdot 3 \cdot 5$. Consequently, the greatest common divisor of 12 and 30 is $2 \cdot 3$, i.e., $\gcd(12, 30) = 6$.

Example 4.2.12. Consider the integers $a = 24$ and $b = 16$. By writing down the prime factorizations of a and b, their greatest common divisor can easily be read off. Observe that $24 = 4 \cdot 6 = 2^3 \cdot 3$ and $16 = 4^2 = 2^4$. Consequently, the greatest common divisor of 24 and 16 is 2^3, i.e., $\gcd(24, 16) = 8$.

Generally, for any nonzero integers a and b, we may determine $\gcd(a, b)$ from the prime factorizations of a and b as in Examples 4.2.11 and 4.2.12 (see Exercise 4.8.12). Certainly, it is possible that $\gcd(a, b) = 1$. One immediate instance of this is that both a and b are prime. Generalizing this notion, we say that positive integers a and b are **relatively prime** if and only if $\gcd(a, b) = 1$.

Example 4.2.13. Observe that 2 and 3 are relatively prime because they are distinct primes, hence they have no prime factors in common. Consequently, we have that $\gcd(2, 3) = 1$.

Example 4.2.14. We claim that 30 and 77 are relatively prime. Observe that the prime factorization of 30 is $30 = 2 \cdot 3 \cdot 5$, and the prime factorization of 77 is $77 = 7 \cdot 11$. Because they have no prime factors in common, we conclude that $\gcd(30, 77) = 1$, hence 30 and 77 are relatively prime.

4.3 Division Algorithm

Even as early as grade school, we learn the process of dividing an integer by a nonzero integer. Each time we divide an integer a by a nonzero integer b, we obtain an integer q and a non-negative integer r that is strictly smaller than $|b|$ such that $a = qb + r$. Explicitly, we say that a is the **dividend**; b is the **divisor**; q is the **quotient**; and r is the **remainder** of the division. Our aim throughout this section is to establish that this process is well-founded, i.e., the process of division of an integer a by a nonzero integer b unambiguously results in integers q and r such that $a = qb + r$ and $0 \leq r < |b|$. We will also establish an algorithm that will allow us to efficiently find the integers q and r.

Example 4.3.1. Consider the case that $a = 11$ and $b = 2$. One can easily see that $11 = 5 \cdot 2 + 1$, hence the integers $q = 5$ and $r = 1$ satisfy the requirements that $a = qb + r$ and $0 \leq r < |b|$.

Example 4.3.2. Consider the case that $a = -17$ and $b = 6$. One can easily see that $-17 = -3 \cdot 6 + 1$, hence the integers $q = -3$ and $r = 1$ satisfy the requirements that $a = qb + r$ and $0 \leq r < |b|$.

Example 4.3.3. Consider the case that $a = -8$ and $b = -9$. One can easily see that $-8 = 1(-9) + 1$, hence the integers $q = 1$ and $r = 1$ satisfy the requirements that $a = qb + r$ and $0 \leq r < |b|$.

Each of the previous examples can be completed by noticing that the integer multiples of b are completely determined by b. Consequently, we may consider all integer multiples of b that do not exceed a, i.e., we may consider the collection $R(a, b) = \{a - qb \mid q \text{ is an integer and } a \geq qb\}$. Our idea is to find the largest (in absolute value) integer q such that $a \geq qb$; then, the difference $a - qb$ must be non-negative (by assumption) and strictly smaller than b (otherwise, we could increase q). Using this intuition as our guide, let us return to find $R(a, b)$ in our previous examples.

Example 4.3.4. By definition, we have that $R(11, 2) = \{11 - 2q \mid q \text{ is an integer and } 11 \geq 2q\}$. Observe that $11 \geq 2q$ if and only if $q \leq 11/2$, hence the only valid values of q in $R(11, 2)$ are $q \leq 5$. Consequently, we have that $-2q \geq -10$ so that $11 - 2q \geq 1$. By consecutively decreasing the value of $q \leq 5$, we find that $R(11, 2) = \{1, 3, 5, 7, \dots\}$ consists of all odd positive integers.

Example 4.3.5. We have that $R(-17, 6) = \{-17 - 6q \mid q \text{ is an integer and } -17 \geq 6q\}$. Observe that $-17 \geq 6q$ if and only if $q \leq -17/6$, hence the only valid values of q in $R(-17, 6)$ are $q \leq -3$. Consequently, we conclude that $R(-17, 6) = \{-17 - 6q \mid q \leq -3 \text{ is an integer}\} = \{1, 7, 13, 19, \dots\}$.

Example 4.3.6. We have that $R(-8, -9) = \{-8 + 9q \mid q \text{ is an integer and } -8 \geq -9q\}$. Observe that $-8 \geq -9q$ if and only if $q \geq 8/9$, hence the only valid values of q in $R(-8, -9)$ are $q \geq 1$. Consequently, we conclude that $R(-8, -9) = \{-8 + 9q \mid q \geq 1 \text{ is an integer}\} = \{1, 10, 19, 28, \dots\}$.

Generalizing the collection $R(a, b)$ and using the Well-Ordering Principle yields the following.

Theorem 4.3.7 (Division Algorithm)**.** *Given any integer a and any nonzero integer b, there exist unique integers q and r such that $a = qb + r$ and $0 \leq r < |b|$.*

Proof. Consider the collection $R(a, b) = \{a - qb \mid q \text{ is an integer and } a \geq qb\}$. By definition, $R(a, b)$ consists of non-negative integers. Observe that if $a \geq 0$, then $R(a, b)$ is nonempty because we may take $q = 0$ to demonstrate that $R(a, b)$ contains a. On the other hand, if $a < 0$, then if $b \geq 1$, then $R(a, b)$ is yet again nonempty because we may take $q = a - 1$ to demonstrate that $R(a, b)$ contains $a - qb$ since $a \geq a - 1 \geq (a - 1)b = qb$. Last, if $a < 0$ and $b \leq -1$, then $R(a, b)$ is once more nonempty because we may take $q = -(a - 1)$ to demonstrate that $R(a, b)$ contains $a - qb$

since $a \geq a-1 \geq -(a-1)b = qb$. Ultimately, this shows that $R(a,b)$ is a nonempty subset of non-negative integers, hence the Well-Ordering Principle implies that there exists a smallest element $r(a,b) = a - qb$ with respect to the total order $\leq$ on the real numbers. Rearranging this identity with $r = r(a,b)$ yields that $a = qb + r$. Considering that $r \geq 0$, it suffices to see that $r < |b|$. On the contrary, suppose that $a - qb = r \geq |b|$. Observe that if $b \geq 1$, then $|b| = b$ yields that $a - qb \geq b$ and $a - (q+1)b \geq 0$. Considering that $a-(q+1)b$ is smaller than the smallest element $r(a,b) = a - qb$ of $R(a,b)$, we obtain a contradiction. Likewise, if $b \leq -1$, then $|b| = -b$ yields that $a - qb \geq -b$ and $a - (q-1)b \geq 0$. Considering that $b \leq -1$, we find that $a - (q-1)b = a - qb + b < a - qb = r(a,b)$. Once again, this contradicts the fact that $r(a,b)$ is the smallest element of $R(a,b)$. Consequently, we conclude that there exist integers q and r such that $a = qb + r$ and $0 \leq r < |b|$.

We must prove next that these integers are unique. We accomplish this by assuming that there exist integers q' and r' such that $a = q'b + r'$ and $0 \leq r' < |b|$. Considering that $a = qb + r$ by the previous paragraph, we conclude that $qb + r = q'b + r'$ so that $b(q - q') = r' - r$. Observe that if $q' = q$, then it is clear that $r' = r$, hence our proof is complete. Consequently, we may assume on the contrary that $q - q'$ is nonzero, hence we must have that $|b| \leq |r' - r|$. Observe that if $r' > r$, then $|r' - r| = r' - r$ implies that $r' \geq |b| + r \geq |b|$ — a contradiction. Likewise, if $r' < r$, then $|r' - r| = r - r'$ implies that $r \geq |b| + r' \geq |b|$ — a contradiction. Either way, we conclude that $r' = r$ so that $b(q - q') = 0$. By hypothesis that b is nonzero, we conclude that $q - q' = 0$ or $q' = q$. □

We have therefore rigorously verified the nontrivial method of division that we have taken for granted since grade school! We remind the reader at this point that if $a = qb + r$ for the unique integers q and r such that $0 \leq r < |b|$, then we refer to the integer a as the **dividend**; the integer b as the **divisor**; the integer q as the **quotient** of a modulo b; and the integer r as the **remainder** of a modulo b. Crucially, the remainder of a modulo b is non-negative and strictly smaller than the absolute value of the divisor. We note that all though the Division Algorithm does not have explicit steps to compute the quotient or remainder of an integer a modulo a nonzero integer b, the proof is constructive in the sense that the unique integers q and $0 \leq r < |b|$ can be deduced from the collection $R(a,b) = \{a - qb \mid q \text{ is an integer and } a \geq qb\}$, as we have done in previous examples.

One of the most fruitful applications of the Division Algorithm is the generalization of the proof by cases technique for divisibility proofs involving any positive integer $n \geq 2$. Explicitly, if we wish to prove that a positive integer $a \geq 2$ divides an integer b, then by the Division Algorithm, we may write $b = qa + r$ for some integers q and r such that $0 \leq r < |a|$. Consequently, it suffices to check each of the $|a|$ cases that $0 \leq r \leq |a| - 1$. We have already tacitly used this kind of proof by cases: in fact, every integer is either even or odd because the remainder an integer modulo 2 is either 0 or 1. Concretely, we illustrate this more general idea for divisibility proofs involving the integer 3.

Example 4.3.8. Prove that if n is an integer, then $3 \mid (2n^2 + 1)$ if and only if $3 \nmid n$.

Proof. We will assume first that $3 \nmid n$. By the Division Algorithm, there are two cases.

1.) Observe that if $n = 3q + 1$ for some integer q, then

$$2n^2 + 1 = 2(3q+1)^2 + 1 = 2(9q^2 + 6q + 1) + 1. = 3(6q^2 + 4q + 1).$$

Considering that $6q^2 + 4q + 1$ is an integer, we conclude that $3 \mid (2n^2 + 1)$.

2.) Observe that if $n = 3q + 2$ for some integer q, then

$$2n^2 + 1 = 2(3q + 2)^2 + 1 = 2(9q^2 + 12q + 4) + 1. = 3(6q^2 + 8q + 3).$$

Considering that $6q^2 + 8q + 3$ is an integer, we conclude that $3 \mid (2n^2 + 1)$.

Conversely, we will prove the contrapositive. We may assume to this end that $3 \mid n$. By definition of divides, there exists an integer q such that $n = 3q$. Consequently, we have that

$$2n^2 + 1 = 2(3q)^2 + 1 = 18q^2 + 1 = 3(6q^2) + 1.$$

Certainly, this is not divisible by 3 because 1 is not divisible by 3, hence $3 \nmid (2n^2 + 1)$. □

Example 4.3.9. Prove that if n is an odd integer such that $3 \nmid n$, then $24 \mid (n^2 - 1)$.

Proof. We will assume that n is an odd integer. By definition of an odd integer, there exists an integer k such that $n = 2k + 1$. By the Division Algorithm, if $3 \nmid n$, then there are two cases.

1.) Observe that if $n = 3q + 1$ for some integer q, then $2k + 1 = 3q + 1$ yields that $2k = 3q$. By Example 4.2.6, we must have that $2 \mid q$ so that $q = 2\ell$ for some integer ℓ, $n = 6\ell + 1$, and

$$n^2 - 1 = (6\ell + 1)^2 - 1 = (36\ell^2 + 12\ell + 1) - 1 = 12(3\ell^2 + \ell).$$

We claim that ℓ is even. On the contrary, if ℓ were odd, then we would have that $\ell = 2m + 1$ for some integer m. Combining this identity with our previous identity that $n = 6\ell + 1$ yields that $n = 6(2m+1)+1 = 12m+2 = 2(6m+1)$ — a contradiction. Consequently, there exists an integer m such that $\ell = 2m$ and $n^2 - 1 = 12[3(2m)^2 + 2m] = 24(6m^2 + m)$.

2.) Observe that if $n = 3q + 2$ for some integer q, then $2k + 1 = 3q + 2$ yields that $2k = 3q + 1$. Consequently, we must have that q is odd; otherwise, if it were the case that $q = 2\ell$ for some integer ℓ, then $2k = 3(2\ell) + 1 = 2(3\ell) + 1$ is odd — a contradiction. We conclude that there exists an integer ℓ such that $q = 2\ell + 1$ and $n = 3q + 2 = 3(2\ell + 1) + 2 = 6\ell + 5$. Observe that

$$n^2 - 1 = (6\ell + 5)^2 - 1 = (36\ell^2 + 60\ell + 25) - 1 = 12(3\ell^2 + 5\ell + 2).$$

We claim that $3\ell^2 + 5\ell + 2$ is even. Certainly, this holds if ℓ is even because the sum of three even integers is even; on the other hand, if $\ell = 2m + 1$ for some integer m, then

$$3\ell^2 + 5\ell + 2 = 3(2m + 1)^2 + 5(2m + 1) + 2 = 3(4m^2 + 4m + 1) + 10m + 7 = 2(6m^2 + 11m + 5).$$

Either way, we conclude that $2 \mid (3\ell^2 + 5\ell + 2)$ so that $24 \mid (n^2 - 1)$, as desired. □

Before we state our next theorem, we remind the reader that if a and b are any nonzero integers and c is any nonzero integer such that $c \mid a$ and $c \mid b$, then we say that c is a common divisor of a and b; the greatest common divisor of a and b is the unique integer $d = \gcd(a, b)$ such that

(a.) $d \mid a$ and $d \mid b$, i.e., d is a common divisor of a and b and

(b.) if c is any common divisor of a and b, then $c \mid d$.

We say that a pair of nonzero integers a and b are relatively prime if and only if $\gcd(a,b) = 1$. Our next theorem states that $\gcd(a,b)$ can be realized as an integer-linear combination of a and b.

Theorem 4.3.10 (Bézout's Identity)**.** *Given any nonzero integers a and b, there exist integers x and y such that $\gcd(a,b) = ax + by$. Even more, $\gcd(a,b)$ divides $av + bw$ for all integers v and w.*

Proof. Consider the set $L(a,b) = \{ax + by \mid x, y \text{ are integers and } ax + by \geq 1\}$ of positive $\mathbb{Z}$-linear combinations of any nonzero integers a and b. Observe that one of $a+b$, $a-b$, $-a+b$, or $-a-b$ lies in $L(a,b)$, hence $L(a,b)$ is nonempty. By the Well-Ordering Principle, there exists a smallest element $d(a,b) = ax + by$ with respect to the total order $\leq$. We will establish that $\gcd(a,b) = d(a,b)$.

By the Division Algorithm, there exist unique integers q_a and r_a such that $a = q_a d(a,b) + r_a$ and $0 \leq r_a < d(a,b)$. By rearranging this identity and using that $d(a,b) = ax + by$, we find that

$$r_a = a - q_a d(a,b) = a - q_a(ax + by) = (1 - q_a x)a - (q_a y)b.$$

Observe that if r_a were nonzero, then it would lie in $L(a,b)$ and satisfy $1 \leq r_a < d(a,b)$, but this is impossible because $d(a,b)$ is the smallest element of $L(a,b)$. Consequently, it must be the case that $r_a = 0$. Likewise, the Division Algorithm with b in place of a yields that $d(a,b)$ divides b. Ultimately, this proves that $d(a,b) \mid a$ and $d(a,b) \mid b$, hence $d(a,b)$ is a common divisor of both a and b.

Consider any other common divisor c of a and b. We must prove that $c \mid d(a,b)$. By assumption, there exist integers q_a and q_b such that $a = q_a c$ and $b = q_b c$, from which it follows that

$$d(a,b) = ax + by = (q_a c)x + (q_b c)y = (q_a x + q_b y)c.$$

By definition, this implies that c divides $d(a,b)$ so that $\gcd(a,b) = d(a,b) = ax + by$, as desired.

Last, by the previous two paragraphs, there exist integers q_a and q_b such that $a = q_a \gcd(a,b)$ and $b = q_b \gcd(a,b)$, hence $\gcd(a,b)$ divides $av + bw$ for any integers v and w by Proposition 4.2.2. □

Corollary 4.3.11 (Uniqueness of GCD)**.** *Greatest common divisors of nonzero integers are unique.*

Proof. By the proof of Bézout's Identity, $\gcd(a,b)$ is unique for any nonzero integers a and b since it is by construction the smallest (w.r.t. the total order $\leq$) positive integer satisfying a property. □

Corollary 4.3.12 (Characterization of Relatively Prime Integers)**.** *Given any nonzero integers a and b, we have that a and b are relatively prime if and only if $ax + by = 1$ for some integers x, y.*

Even though Bézout's Identity guarantees that the existence of integers x and y such that $\gcd(a,b) = ax + by$ for any pair of nonzero integers a and b, neither the statement of this fact nor its proof provides any tools for explicitly finding these integers x and y. We conclude this section by constructing a step-by-step process for producing the integers x and y for which $\gcd(a,b) = ax + by$. Contrary to the Division Algorithm (that is not in fact an algorithm after all), we will obtain a programmable, reproducible algorithm for this procedure that can be readily coded for computing.

Example 4.3.13. Consider the case that $a = 24$ and $b = 16$. We know already that $\gcd(a,b) = 8$, and it is not difficult to see that $8 = 24 \cdot 1 + 16(-1)$; however, this fact can also be seen as follows: by the Division Algorithm, we have that $24 = 1 \cdot 16 + 8$, hence we have that $8 = 24 \cdot 1 + 16(-1)$.

Example 4.3.14. Consider the case that $a = 110$ and $b = 24$. Observe that the unique prime factorizations of 110 and 24 are $110 = 10 \cdot 11 = 2 \cdot 5 \cdot 11$ and $24 = 2^3 \cdot 3$, respectively. By Exercise 4.8.12, it follows that $\gcd(110, 24) = 2$. By successively implementing the Division Algorithm, we may find the integers x and y such that $110x + 24y = 2$, as guaranteed to us by Bézout's Identity. Explicitly, we begin by running the Division Algorithm with $a = 110$ and $b = 24$ to find the unique integers q_1 and $0 \leq r_1 < 24$ such that $110 = 24q_1 + r_1$; then, we repeat the Division Algorithm with 24 and r_1 to produce the unique integers q_2 and $0 \leq r_2 < r_1$ such that $24 = q_2 r_1 + r_2$. Continuing in this manner produces a strictly decreasing sequence $r_1 > r_2 > \cdots > r_n$ of non-negative integers at the nth step. Bearing in mind the Well-Ordering Principle, this sequence must have a least element, hence the process must eventually terminate. Putting this process to the test, we find that

$$\begin{aligned} 110 &= 4 \cdot 24 + 14, \\ 24 &= 1 \cdot 14 + 10, \\ 14 &= 1 \cdot 10 + 4, \text{ and} \\ 10 &= 2 \cdot 4 + 2. \end{aligned}$$

We determine the integers x and y such that $110x + 24y = 2$ by unravelling this process in reverse. Explicitly, our last identity gives that $10 - 2 \cdot 4 = 2$; the identity before that gives that $4 = 14 - 1 \cdot 10$, hence we have that $-2 \cdot 14 + 3 \cdot 10 = 10 - 2 \cdot (14 - 1 \cdot 10) = 2$; the identity before $14 = 1 \cdot 10 + 4$ gives that $10 = 24 - 1 \cdot 14$, hence we have that $3 \cdot 24 - 5 \cdot 14 = -2 \cdot 14 + 3 \cdot (24 - 1 \cdot 14) = 2$; and at last, the identity before $24 = 1 \cdot 14 + 10$ gives that $14 = 110 - 4 \cdot 24$, hence we have that

$$110(-5) + 24(23) = 3 \cdot 24 - 5 \cdot (110 - 4 \cdot 24) = 2.$$

Algorithm 4.3.15 (Euclidean Algorithm)**.** Consider any nonzero integers a and b such that $a \geq b$. We may produce integers x and y such that $\gcd(a, b) = ax + by$ according to the following.

1.) Use the Division Algorithm to find integers q_1 and r_1 such that $a = q_1 b + r_1$ and $0 \leq r_1 < |b|$.

2.) Use the Division Algorithm to find integers q_2 and r_2 such that $b = q_2 r_1 + r_2$ and $0 \leq r_2 < r_1$.

3.) Use the Division Algorithm to find integers q_3 and r_3 such that $r_1 = q_3 r_2 + r_3$ and $0 \leq r_3 < r_2$.

4.) Continue in this manner until the remainder r_{n+1} divides r_n. By the Well-Ordering Principle, this must eventually occur, and moreover, it must occur in a finite number of steps.

5.) Use the fact that $r_{n-1} = q_{n+1} r_n + r_{n+1}$ to express that $r_{n+1} = r_{n-1} - q_{n+1} r_n$.

6.) Use the fact that $r_{n-2} = q_n r_{n-1} + r_n$ to express that $r_n = r_{n-2} - q_n r_{n-1}$; then, use the fact that $r_{n+1} = r_{n-1} - q_{n+1} r_n$ to express that $r_{n+1} = r_{n-1} - q_{n+1}(r_{n-2} - q_n r_{n-1})$ so that

$$r_{n+1} = (q_n q_{n+1} + 1) r_{n-1} - q_{n+1} r_{n-2}.$$

7.) Continue in this manner to produce integers x and y such that $r_{n+1} = ax + by$.

By Bézout's Identity and Proposition 4.2.3, we must have that $\gcd(a, b) \leq r_{n+1}$. Conversely, because r_{n+1} divides r_n by (4.), it must divide r_k for all integers $1 \leq k \leq n$ by steps (5.) through (7.) above. Consequently, by step (2.) above, we conclude that r_{n+1} must divide b, and by step (1.) above, we conclude that r_{n+1} must divide a. Ultimately, this shows that r_{n+1} is a common divisor of a and b, hence we must have that r_{n+1} divides $\gcd(a, b)$; in particular, we have that $r_{n+1} = \gcd(a, b)$.

4.4 Proofs Involving Sets, Set Operations, and Functions

Combined, the calculus of logic of Chapter 2 and the basic proof techniques of Chapter 3 allow us to deduce further properties of sets and set operations. On their own, naïve set theory and formal logic are two rich and interesting areas of mathematics, but their utility in the broader patchwork of pure and applied mathematics and computer science makes them indelible tools in our toolkit.

Before we proceed to any new material, we provide first a reinterpretation of Chapter 1 in the language of Chapter 2. We will assume to this end that X and Y are some (possibly empty) sets.

- We may view the **set membership** $x \in X$ as the following statement.

  $$M(x, X)\colon \text{ We have that } x \text{ is an element of the set } X.$$

  Consequently, the negation $x \notin X$ of the set membership $x \in X$ is the following statement.

  $$\neg M(x, X)\colon \text{ We have that } x \text{ is not an element of the set } X.$$

- We may view the **subset containment** $X \subseteq Y$ as the following statement.

  $$C(X, Y)\colon \text{ For every element } x \in X, \text{ we have that } x \in Y.$$

  Considering that any universally quantified statement can be viewed as a conditional statement, we may view the subset containment $X \subseteq Y$ as the following conditional statement.

  $$C(X, Y)\colon \text{ If } x \in X, \text{ then } x \in Y.$$

  Consequently, the empty set $\varnothing$ is a subset of every set X: indeed, $C(\varnothing, X)$ is vacuously true! Observe that the negation $X \nsubseteq Y$ of the subset containment is an existence statement.

  $$\neg C(X, Y)\colon \text{ There exists an element } x \in X \text{ such that } x \notin Y.$$

- We may view the **proper subset containment** $X \subsetneq Y$ as the following statement.

  $$C^*(X, Y)\colon \text{ We have that } X \text{ is a subset of } Y \text{ and there exists an element } y \in Y \setminus X.$$

  Consequently, the proper subset containment is a conjunctive statement.

- We may view the **set equality** $X = Y$ as the following conjunctive statement.

  $$E(X, Y)\colon \text{ We have that } X \subseteq Y \text{ and } Y \subseteq X.$$

- Elements of either the set X or the Y define the **set union** $X \cup Y$ of X and Y.

  $$X \cup Y = \{w \mid (w \in X) \vee (w \in Y) \text{ is true}\}$$

- Elements of both the set X and the set Y define the **set intersection** $X \cap Y$ of X and Y.

  $$X \cap Y = \{w \mid (w \in X) \wedge (w \in Y) \text{ is true}\}$$

- Elements of the set Y but not the set X define the **relative complement** $Y \setminus X$ of X in Y.

$$Y \setminus X = \{w \mid (w \in Y) \wedge (w \notin X) \text{ is true}\}$$

- We may view the **Cartesian product** $X \times Y$ of the sets X and Y as the collection of all ordered pairs (x, y) for which x is an element of X and y is an element of Y.

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

By using the above dictionary between set theory and logic, we can prove many facts about sets.

Example 4.4.1. Prove that for any sets X, Y, and W such that $X \subseteq W$ and $Y \subseteq W$, we have that

$$X \setminus Y = X \cap (W \setminus Y).$$

Proof. By the above definition of set equality, we must demonstrate that $X \setminus Y \subseteq X \cap (W \setminus Y)$ and $X \cap (W \setminus Y) \subseteq X \setminus Y$. By definition of $X \setminus Y$, if $x \in X \setminus Y$, then $x \in X$ and $x \notin Y$. By assumption that $X \subseteq W$, we find that $x \in W$ and $x \notin Y$ so that $x \in X$ and $x \in W \setminus Y$. We conclude that $x \in X \cap (W \setminus Y)$, from which it follows that $X \setminus Y \subseteq X \cap (W \setminus Y)$. Conversely, if $x \in X \cap (W \setminus Y)$, then $x \in X$ and $x \in W \setminus Y$. By definition of $W \setminus Y$, we have that $x \in W$ and $x \notin Y$. We conclude that $x \in X \setminus Y$ since $x \in X$ and $x \notin Y$, from which it follows that $X \cap (W \setminus Y) \subseteq X \setminus Y$. □

Example 4.4.2. Prove that for any sets X and Y, we have that $X = (X \cap Y) \cup (X \setminus Y)$.

Proof. By the above definition of set equality, we must demonstrate that $X \subseteq (X \cap Y) \cup (X \setminus Y)$ and $(X \cap Y) \cup (X \setminus Y) \subseteq X$. Given any element $x \in X$, either $x \in Y$ or $x \notin Y$ by the Law of the Excluded Middle: if the former holds, then $x \in X \cap Y$; if the latter holds, then $x \in X \setminus Y$. Either way, it follows that $x \in (X \cap Y) \cup (X \setminus Y)$. Conversely, suppose that $x \in (X \cap Y) \cup (X \setminus Y)$. Each of the sets $X \cap Y$ and $X \setminus Y$ is by definition a subset of X, hence we have that $x \in X$. □

Example 4.4.3. Prove that for any sets X and Y, we have that $X \cup Y = X$ if and only if $Y \subseteq X$.

Proof. By the above definition of set equality, we must demonstrate that if $Y \subseteq X$, then $X \cup Y \subseteq X$ and $X \subseteq X \cup Y$. Observe that the latter inclusion is true by definition of the union, hence it suffices to prove that if $Y \subseteq X$, we have that $X \cup Y \subseteq X$. We will assume to this end that $Y \subseteq X$. Observe that if $w \in X \cup Y$, then by definition of the set union, we have that $w \in X$ or $w \in Y$. Either way, by hypothesis that $Y \subseteq X$, it follows that $w \in X$, hence we conclude that $X \cup Y \subseteq X$.

Conversely, we will assume that $X \cup Y = X$. Given any element $y \in Y$, we have that $y \in X \cup Y$ so that $y \in X$ by assumption that $X \cup Y = X$. We conclude that $Y \subseteq X$, as desired. □

We will assume throughout the rest of this section that X, Y, and W are some (possibly empty) sets for which the inclusions $X \subseteq W$ and $Y \subseteq W$ hold. We remind the reader that in this case, we refer to W as our **universe** (or as the **universal set**), and we may view all elements of X and Y as elements of W via the aforementioned inclusions. We obtain the following membership laws.

Theorem 4.4.4 (Law of the Excluded Middle for Sets)**.** *Consider any (possibly empty) sets $X \subseteq W$. Given any element $w \in W$, we must have that either $w \in X$ or $w \notin X$.*

Theorem 4.4.5 (Law of Non-Contradiction for Sets)**.** *Consider any (possibly empty) sets $X \subseteq W$. Given any element $w \in W$, we cannot have that both $w \in X$ and $w \notin X$.*

We omit the proofs of the aforementioned facts because they follow immediately from the Law of the Excluded Middle and the Law of Non-Contradiction for the set membership statement $M(w, X)$. Even more, we have De Morgan's Laws for the relative complements of $X \cup Y$ and $X \cap Y$ in W.

Theorem 4.4.6 (De Morgan's Laws for Sets)**.** *Consider any (possibly empty) sets $X, Y \subseteq W$.*

1.) *We have that $W \setminus (X \cup Y) = (W \setminus X) \cap (W \setminus Y)$.*

2.) *We have that $W \setminus (X \cap Y) = (W \setminus X) \cup (W \setminus Y)$.*

Proof. (1.) We will first establish the inclusion $W \setminus (X \cup Y) \subseteq (W \setminus X) \cap (W \setminus Y)$. Given any element $w \in W \setminus (X \cup Y)$, we have that $w \in W$ and $w \notin X \cup Y$ by definition of the set relative complement. Consequently, we must have that $w \notin X$ and $w \notin Y$. But this implies that $w \in W \setminus X$ and $w \in W \setminus Y$ so that $w \in (W \setminus X) \cap (W \setminus Y)$. Conversely, suppose that $w \in (W \setminus X) \cap (W \setminus Y)$. By definition of the set intersection, we have that $w \in W \setminus X$ and $w \in W \setminus Y$. By definition of the relative complement, we have that $w \in W$ and $w \notin X$ and $w \notin Y$ so that $w \in W$ and $w \notin X \cup Y$.

(2.) We will first establish the inclusion $W \setminus (X \cap Y) \subseteq (W \setminus X) \cup (W \setminus Y)$. Given any element $w \in W \setminus (X \cap Y)$, we have that $w \in W$ and $w \notin X \cap Y$ by definition of the relative complement. Consequently, we must have that either $w \notin X$ or $w \notin Y$. But this implies that $w \in W \setminus X$ or $w \in W \setminus Y$ so that $w \in (W \setminus X) \cup (W \setminus Y)$. Conversely, suppose that $w \in (W \setminus X) \cup (W \setminus Y)$. By definition of the union, we have that $w \in W \setminus X$ or $w \in W \setminus Y$. Consequently, we have that $w \in W$ and either $w \notin X$ or $w \notin Y$. Either way, it follows that $w \notin X \cap Y$ so that $w \in W \setminus (X \cap Y)$. □

Often, we will simultaneously deal with $n \geq 2$ (possibly empty) sets $X_1, X_2, \ldots, X_n$ such that $X_i \subseteq W$ for each integer $1 \leq i \leq n$; in this case, it is easiest to adopt the notation of Section 1.4. We will say that each set X_i is **indexed** by an integer subscript $1 \leq i \leq n$. Consider the set union

$$\bigcup_{i=1}^{n} X_i = X_1 \cup X_2 \cup \cdots \cup X_n = \{w \mid w \in X_i \text{ for some integer } 1 \leq i \leq n\}.$$

Observe that $w \in \cup_{i=1}^{n} X_i$ if and only if the existence statement "$\exists i \in \{1, 2, \ldots, n\}$, $w \in X_i$" is true. Consider the set intersection

$$\bigcap_{i=1}^{n} X_i = X_1 \cap X_2 \cap \cdots \cap X_n = \{w \mid w \in X_i \text{ for every integer } 1 \leq i \leq n\}.$$

Observe that $w \in \cap_{i=1}^{n} X_i$ if and only if the universal statement "$\forall i \in \{1, 2, \ldots, n\}$, $w \in X_i$" is true. Generally, De Morgan's Laws for Sets hold for finite unions and intersections of sets as follows.

Proposition 4.4.7 (Generalized De Morgan's Laws)**.** *Consider any sets $X_1, X_2, \ldots, X_n \subseteq W$.*

1.) *We have that $W \setminus (X_1 \cup X_2 \cup \cdots \cup X_n) = (W \setminus X_1) \cap (W \setminus X_2) \cap \cdots \cap (W \setminus X_n)$.*

2.) *We have that $W \setminus (X_1 \cap X_2 \cap \cdots \cap X_n) = (W \setminus X_1) \cup (W \setminus X_2) \cup \cdots \cup (W \setminus X_n)$.*

Likewise, we have the following distributive laws for finite unions and intersections of sets.

Proposition 4.4.8 (Distributive Laws for Sets)**.** *Consider any sets $X_1, X_2, \ldots, X_n$ and Y.*

1.) *We have that $Y \cap (X_1 \cup X_2 \cup \cdots \cup X_n) = (Y \cap X_1) \cup (Y \cap X_2) \cup \cdots \cup (Y \cap X_n)$.*

2.) *We have that $Y \cup (X_1 \cap X_2 \cap \cdots \cap X_n) = (Y \cup X_1) \cap (Y \cup X_2) \cap \cdots \cap (Y \cup X_n)$.*

Proof. (1.) By definition of set equality, we must establish both of the set containments

$$Y \cap (X_1 \cup X_2 \cup \cdots \cup X_n) \subseteq (Y \cap X_1) \cup (Y \cap X_2) \cup \cdots \cup (Y \cap X_n) \text{ and}$$
$$Y \cap (X_1 \cup X_2 \cup \cdots \cup X_n) \supseteq (Y \cap X_1) \cup (Y \cap X_2) \cup \cdots \cup (Y \cap X_n).$$

Consider any element $x \in Y \cap (X_1 \cup X_2 \cup \cdots \cup X_n)$. By definition of set intersection, we have that $x \in Y$ and $x \in X_1 \cup X_2 \cup \cdots \cup X_n$. Likewise, by definition of set union, we have that $x \in X_i$ for some integer $1 \leq i \leq n$. Consequently, it follows that $x \in Y \cap X_i$ for some integer $1 \leq i \leq n$ so that $x \in (Y \cap X_1) \cup (Y \cap X_2) \cup \cdots \cup (Y \cap X_n)$, and the subset containment $\subseteq$ is established. Conversely, suppose that $x \in (Y \cap X_1) \cup (Y \cap X_2) \cup \cdots \cup (Y \cap X_n)$. By definition of set union, we have that $x \in Y \cap X_i$ for some integer $1 \leq i \leq n$. Consequently, it follows that $x \in Y$ and $x \in X_i$ for some integer $1 \leq i \leq n$. But this implies that $x \in X$ and $x \in X_1 \cup X_2 \cup \cdots \cup X_n$, hence $\supseteq$ holds.

(2.) We reserve the proof of the second distributive law for sets as Exercise 4.8.16. □

By appealing to our dictionary between logic and set theory, we may also prove many important properties of functions. We remind the reader that a function $f : X \to Y$ is simply a subset of the Cartesian product $X \times Y$ with the additional property that for each element $x \in X$, there exists one and only one element $f(x) = y \in Y$ such that $(x, f(x)) \in f$. Each function $f : X \to Y$ gives rise to a set $\text{range}(f) = \{f(x) \mid x \in X\}$ of all **images** of elements of X under f. Conversely, for each subset $W \subseteq Y$, we may consider the set $f^{-1}(W) = \{x \in X \mid f(x) \in W\}$ of **inverse images** of elements of W under f. We refer to the function $f : X \to Y$ as **injective** provided that $f(x) = f(y)$ implies that $x = y$ for all elements $f(x) \in \text{range}(f)$. Likewise, we refer to the function $f : X \to Y$ as **surjective** provided that $Y = \text{range}(f)$, i.e., for every element $y \in Y$, there exists an element $x \in X$ such that $y = f(x)$. We say that a function is **bijective** if it is injective and surjective.

Proposition 4.4.9. *Consider any function $f : X \to Y$ between any two sets X and Y.*

(a.) *If f is injective, then $f^{-1}(f(V)) = V$ for any set $V \subseteq X$.*

(b.) *If f is surjective, then $f(f^{-1}(W)) = W$ for any set $W \subseteq Y$.*

Proof. 1.) By Exercise 4.8.18, it suffices to prove that $f^{-1}(f(V)) \subseteq V$. Let x be an arbitrary element of $f^{-1}(f(V))$. By definition of the inverse image $f^{-1}(f(V))$ of $f(V)$, this means that $f(x) \in f(V)$. By definition of the image $f(V)$ of V, we have that $f(x) = f(v)$ for some element $v \in V$. Last, by assumption that f is injective and $V \subseteq X$, we conclude that $x = v$, hence x is an element of V.

(2.) By Exercise 4.8.18, it suffices to prove that $W \subseteq f(f^{-1}(W))$. Let w be any element of W. By assumption that f is surjective and $W \subseteq Y$, there exists an element $x \in X$ such that $w = f(x)$. By definition of the inverse image $f^{-1}(W)$, it follows that $x \in f^{-1}(W)$. By definition of the image $f(f^{-1}(W))$, we conclude that $w = f(x)$ for some element $x \in f^{-1}(W)$ so that $w \in f(f^{-1}(W))$. □

Conversely, if $f^{-1}(f(V)) = V$ holds for any set $V \subseteq X$, then $f : X \to Y$ must be injective; likewise, if $f(f^{-1}(W)) = W$ for any set $W \subseteq Y$, then f must be surjective (see Exercise 4.8.19).

4.5 Counting Principles

Combinatorics broadly encompasses the mathematics of counting and discrete structures. Even though the problem of counting objects of a finite set might seem ostensibly simple on first thought, many problems in combinatorics are quite subtle and require deep mathematics and elegant proofs. Combinatorics has numerous applications in both the theoretical sciences (e.g., commutative algebra, probability, and statistical physics) and applied sciences (e.g., computer science and evolutionary biology). Our aim in this course is not to concern ourselves with the many existing sophisticated counting problems; rather, we provide in this section three fundamental combinatorial principles.

Going forward, we will refer to a **task** as any process that can be completed in a finite amount of time. Concretely, a task may be as simple as putting on a pair of socks or placing a ball into a bin or as complex as the Towers of Hanoi. Often, we will consider tasks that can be completed in a fixed and finite number of ways. By way of example, if our task is to put on a pair of socks, then the number of ways to accomplish this task is completely determined by the number of pairs of socks we possess. Crucially, when completing a task in the context of this course, we do not consider the physical manner in which the task is carried out; we are only concerned with our number of options.

Theorem 4.5.1 (Addition Rule of Counting)**.** *Consider any procedure that requires the completion of $k \geq 2$ tasks $t_1, t_2, \ldots, t_k$ in a manner such that no pair of tasks can be performed simultaneously. Provided that the task t_i can be performed in n_i ways for each integer $1 \leq i \leq k$, the total number of ways in which the procedure can be carried out is $n_1 + n_2 + \cdots + n_k$.*

Proof. We proceed by induction on the number of tasks $k \geq 2$. Observe that for two tasks t_1 and t_2, we may complete the first task t_1 in n_1 ways; then, we may subsequently complete the second task t_2 in n_2 ways. Considering that we cannot perform both tasks simultaneously by assumption, there are a total of $n_1 + n_2$ ways to complete tasks t_1 and t_2 in either order. We will assume by induction that there are $n_1 + n_2 + \cdots + n_k$ ways in which a procedure involving $k \geq 2$ tasks $t_1, t_2, \ldots, t_k$ can be carried out in a manner such that no pair of tasks can be performed simultaneous. Consider the case that we wish to perform the tasks $t_1, t_2, \ldots, t_{k+1}$ in a manner such that no pair of tasks can be performed simultaneously. By first performing the tasks $t_1, t_2, \ldots, t_k$ and subsequently performing the task t_{k+1}, we obtain a procedure that requires the completion of two tasks T_1 and T_2: explicitly, the task T_1 involves completing the tasks $t_1, t_2, \ldots, t_k$ in a manner such that no pair of tasks can be performed simultaneously, and the task T_2 is simply the task t_{k+1}. By the above case of two tasks, if the task T_1 can be performed in N_1 ways and the task T_2 can be performed in N_2 ways, then the number of ways to complete this procedure is $N_1 + N_2$. By our inductive hypothesis, we have that $N_1 = n_1 + n_2 + \cdots + n_k$, and it is clear that $N_2 = n_{k+1}$ since the task T_2 is nothing but the task t_{k+1}. Consequently, the total number of ways in which the procedure involving the tasks $t_1, t_2, \ldots, t_{k+1}$ can be carried out is $n_1 + n_2 + \cdots + n_k + n_{k+1}$. We conclude the desired result by induction. □

Example 4.5.2. Evergreen University's mathematics program requires undergraduates to obtain an academic minor. Angelina is a sophomore in the mathematics program who needs three credits for a computer science minor, and she is eligible for three 3-credit computer science courses in the Fall 2024 semester. Even more, Angelina could also enroll in microeconomics, German, or physics, but she cannot take any of these class concurrently because they are offered at the same time. By the Addition Rule of Counting, there are $3 + 1 + 1 + 1 = 6$ options for Angelina's schedule.

One of the most important consequences of the Addition Rule of Counting is the following.

Proposition 4.5.3 (Cardinality of a Finite Set Under a Partition)**.** *Consider any finite set S with cardinality $|S|$. Given any partition $S = S_1 \cup S_2 \cup \cdots \cup S_k$ of S, we have that*

$$|S| = |S_1| + |S_2| + \cdots + |S_k|.$$

Proof. We proceed by induction on the integer $k \geq 2$. Consider any partition $S = S_1 \cup S_2$ of a finite set S. By definition of a partition, we must have that $S_1 \cap S_2 = \varnothing$, i.e., S_1 and S_2 have no elements in common. We conclude by the Addition Rule of Counting that $|S| = |S_1| + |S_2|$. Explicitly, the procedure of labelling the distinct elements of S can be carried out by first labelling the distinct elements of the set S_1 and subsequently labelling the distinct elements of S_2. Considering that S_1 and S_2 have no elements in common, the number of ways to label the distinct elements of S is $|S_1| + |S_2|$ since there are as many distinct elements of S_1 and S_2, respectively. We may assume by induction that the statement holds for some integer $k \geq 2$. Consider any partition of S into $k+1$ subsets $S = S_1 \cup S_2 \cup \cdots \cup S_{k+1}$. Observe that $T = S_1 \cup S_2 \cup \cdots \cup S_k$ is a finite set partitioned by $S_1 \cup S_2 \cup \cdots \cup S_k$ since these sets are disjoint by assumption and their union is T. Even more, we have that $T \cap S_{k+1} = (S_1 \cup S_2 \cup \cdots \cup S_k) \cap S_{k+1} = (S_1 \cap S_{k+1}) \cup (S_2 \cap S_{k+1}) \cup \cdots \cup (S_k \cap S_{k+1}) = \varnothing$ because $S_i \cap S_{k+1} = \varnothing$ for each integer $1 \leq i \leq k$ by assumption that $S_1, S_2, \ldots, S_{k+1}$ partition S. We conclude that $S = T \cup S_{k+1}$ is a partition of S. By the base case of our induction, we have that $|S| = |T| + |S_{k+1}|$. By our inductive hypothesis, we have that $|T| = |S_1| + |S_2| + \cdots + |S_k|$. □

Like with the Addition Rule of Counting, there is a rule for carrying out tasks simultaneously.

Theorem 4.5.4 (Multiplication Rule of Counting)**.** *Consider any procedure that requires the completion of $k \geq 2$ tasks $t_1, t_2, \ldots, t_k$. Provided that the task t_i can be performed in n_i ways for each integer $1 \leq i \leq k$, the total number of ways in which the procedure can be carried out is $n_1 n_2 \cdots n_k$.*

Proof. We proceed by induction on the number of tasks $k \geq 2$. Observe that for two tasks t_1 and t_2, we may complete the first task t_1 in n_1 ways; then, for each of these n_1 ways of completing task t_1, there are n_2 ways to complete task t_2. By the Addition Rule of Counting, there are a total of $n_1 n_2 = n_2 + n_2 + \cdots + n_2$ ways in which the procedure can be carried out. We will assume by induction that there are $n_1 n_2 \cdots n_k$ ways in which a procedure involving $k \geq 2$ tasks $t_1, t_2, \ldots, t_k$ can be carried out. Consider the case that we now wish to perform the tasks $t_1, t_2, \ldots, t_{k+1}$. Like in the proof of the Addition Rule of Counting, we may first perform the tasks $t_1, t_2, \ldots, t_k$ and subsequently perform the task t_{k+1}. Each of these can be viewed as a task: the number of ways to perform the first task is $n_1 n_2 \cdots n_k$ by our inductive hypothesis, and the number of ways to perform the second task in n_{k+1} by assumption. Consequently, there are $(n_1 n_2 \cdots n_k) n_{k+1}$ ways in which these two tasks can be carried out, hence the desired formula is established by induction. □

Example 4.5.5. On Casual Fridays, Angelina wears a T-shirt, pants, and tennis shoes; however, she cannot decide which of her four T-shirts, three pairs of pants, and five pairs of tennis shoes to wear. By the Multiplication Rule of Counting, Angelina has $4 \cdot 3 \cdot 5 = 60$ possible outfits.

Example 4.5.6. Lunch at the Evergreen University cafeteria consists of a vegetarian option or a vegan option. Vegetarian students may choose from six entrees, three desserts, and three drinks, hence by the Multiplication Rule of Counting, there are $6 \cdot 3 \cdot 3 = 54$ possible vegetarian meals.

Even more, vegan students may choose from four entrees, two desserts, and three drinks, hence by the Multiplication Rule of Counting, there are $4 \cdot 2 \cdot 3 = 24$ possible vegan meals. Considering that students can choose either the vegan or vegetarian meal option but not both, by the Addition Rule of Counting, Evergreen University students have $24 + 54 = 78$ meal options in the cafeteria.

Example 4.5.7. Every non-negative integer can be represented in **binary** (or **base-2**) by a **string** $a_n a_{n-1} \cdots a_1 a_0$ of **length** $n+1$ such that for each integer $0 \leq i \leq n$, we have that $a_i \in \{0, 1\}$. Compute the total number of binary strings of length $n+1$ for any non-negative integer n; then, determine how many binary strings of length $n \geq 2$ begin with the digit 1 and end with the digit 0.

Solution. Observe that for each integer $0 \leq i \leq n$, there are exactly two possibilities for the ith digit a_i. Consequently, if we consider the procedure consisting of the $n+1$ tasks t_i of assigning the digit a_i, then each task t_i can be performed in 2 ways; therefore, by the Multiplication Rule of Counting, there are 2^{n+1} binary strings of length $n+1$ for any non-negative integer n. On the other hand, if the first digit is 1 and the last digit is 0, then we must only determine the digits $a_1, \ldots, a_{n-1}$. Like before, there are 2^{n-1} ways to assign each of the $n-1$ digits $a_1, \ldots, a_{n-1}$ with either 0 or 1. ◇

By Proposition 1.14.1, a pair of finite sets X and Y admit a bijection $f : X \to Y$ if and only if X and Y possess the same number of elements. Given any nonempty set X of finite cardinality, the collection of bijective functions $f : X \to X$ forms 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 4.5.8. *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 $X = \{x_1, x_2, \ldots, x_n\}$, then there are n distinct choices for the value of $f(x_1)$; there are $n-1$ distinct choices for $f(x_2)$ other than $f(x_1)$; there are $n-2$ distinct choices for $f(x_3)$ other than $f(x_1)$ and $f(x_2)$; and in general, there are $n-i+1$ choices for $f(x_i)$ distinct from $f(x_1), f(x_2), \ldots, f(x_{i-1})$ for each integer $1 \leq i \leq n$. We conclude that there are $n!$ bijections between X and Y since there are $n! = n(n-1)(n-2)\cdots 2 \cdot 1$ possibilities. □

Before we proceed with our third and final counting principle, we recall that for any real number x, the **ceiling** of x is the least integer that is greater than or equal to x. Explicitly, we have that

$$\lceil x \rceil = \min\{n \in \mathbb{Z} \mid n \geq x\}.$$

By way of example, observe that $\lceil 2.5 \rceil = 3$ and $\lceil -1.01 \rceil = -1$. One simple rule of thumb for the ceiling function is that if the decimal approximation of x is known, then the ceiling $\lceil x \rceil$ of x can be obtained by rounding up to the nearest integer if $x \geq 0$ or by chopping off the decimal if $x \leq -1$. Crucially, we note that if a is an integer such that $a < \lceil x \rceil$, then we must have that $a < x$; otherwise, if it were the case that $a \geq x$, then we would have that $a \geq \lceil x \rceil$ — a contradiction.

Theorem 4.5.9 (Pigeonhole Principle)**.** *Consider any nonempty finite set S with $n \geq 1$ elements. Given any partition of S into k subsets $S_1, S_2, \ldots, S_k$, there exists an integer $1 \leq i \leq k$ such that S_i contains at least $\lceil n/k \rceil$ elements. Even more, if $|S_i| \geq n_i$ for each integer $1 \leq i \leq k$, then each set $T \subseteq S$ with $|T| \geq n_1 + n_2 + \cdots + n_k - k + 1$ contains n_i elements of S_i for some integer $1 \leq i \leq k$.*

Proof. On the contrary, suppose that for every integer $1 \leq i \leq k$, we have that $|S_i| < \lceil n/k \rceil$. Consequently, we must have that $|S_i| < n/k$ for all integers $1 \leq i \leq k$. Considering that $S_1, S_2, \ldots, S_k$ form a partition of S into k subsets, by Proposition 4.5.3, it follows that

$$n = |S| = |S_1| + |S_2| + \cdots + |S_k| < \underbrace{(n/k) + (n/k) + \cdots + (n/k)}_{k \text{ summands}} = n.$$

But this is a contradiction. We conclude therefore that $|S_i| \geq \lceil n/k \rceil$ for some integer $1 \leq i \leq k$. Even more, suppose that $|S_i| \geq n_i$ for each integer $1 \leq i \leq k$. Consider any set $T \subseteq S$ such that $|T| \geq n_1 + n_2 + \cdots + n_k - k + 1$. On the contrary, assume that T does not contain n_i elements of S_i for any integer $1 \leq i \leq k$. Observe that $T = T \cap S = T \cap (S_1 \cup S_2 \cup \cdots \cup S_k)$, hence by the Distributive Law for Set Intersection, it follows that $T = (T \cap S_1) \cup (T \cap S_2) \cup \cdots \cup (T \cap S_k)$ is a partition of T. By assumption, we have that $|T \cap S_i| \leq n_i - 1$ for each integer $1 \leq i \leq k$ so that

$$n_1 + n_2 + \cdots + n_k - k + 1 \leq |T| \leq (n_1 - 1) + (n_2 - 1) + \cdots + (n_k - 1) = n_1 + n_2 + \cdots + n_k - k.$$

But this is a contradiction, hence T must contain n_i elements of S_i for some integer $1 \leq i \leq k$. □

We remark that the Pigeonhole Principle is so named because it can be described as a condition on the number of ways to place n pigeons into k pigeon coops: indeed, if we view the elements of the set S as pigeons and the subsets $S_1, S_2, \ldots, S_k$ that partition S as the pigeon coops, then the Pigeonhole Principle says that if we place all n pigeons in some k coops, one pigeon coop will contain at least $\lceil n/k \rceil$ pigeons. We note that intuitively, this makes sense since the best case scenario is that all pigeons are placed in the same coop (in which case, one pigeon coop contains all n pigeons), and the worst case scenario is that we attempt to place each of the n pigeons in a distinct pigeon coop and inevitably wind up with some coop that contains the purported number of pigeons. By this interpretation, the Pigeonhole Principle asserts that if each pigeon is housed in some coop, then we can find a pigeon coop that is "sufficiently full." Consequently, in order to employ the Pigeonhole Principle, it suffices to determine the "pigeons" and the "coops." Let us try out some examples.

Example 4.5.10. Prove that for any 11 integers, at least two integers end in the same digit.

Solution. Before we prove this fact, we must first determine the underlying "pigeons" and "coops." We claim that two integers end in the same number, hence it seems natural to consider the "pigeons" as the 11 integers in question and the "coops" as the last digits of these integers. Every digit of an integer is a number between 0 and 9, so there are only 10 coops for our 11 pigeons. By the Pigeonhole Principle, at least one of these coops must contain at least $\lceil 11/10 \rceil = \lceil 1.1 \rceil = 2$ pigeons! ◇

Proof. Observe that every integer a ends in an integer that is uniquely determined by the remainder of a modulo 10: indeed, by the Division Algorithm, there exist unique integers q and r such that $a = 10q + r$ and $0 \leq r \leq 9$, hence a ends in the digit r. Consequently, we may partition the integers as $\mathbb{Z} = S_0 \cup S_1 \cup S_2 \cup \cdots \cup S_9$ such that $S_i = \{10q + i \mid q \in \mathbb{Z}\}$ for each integer $0 \leq i \leq 9$. Even more, observe that for any subset $T \subseteq \mathbb{Z}$ such that $|T| \geq 11$, we have that

$$T = T \cap \mathbb{Z} = T \cap (S_0 \cup S_1 \cup S_2 \cup \cdots \cup S_9) = (T \cap S_0) \cup (T \cap S_1) \cup (T \cap S_2) \cup \cdots \cup (T \cap S_9)$$

is a partition of T into 10 subsets. Consequently, by the Pigeonhole Principle, there exists an integer $0 \leq i \leq 9$ such that $|T \cap S_i| \geq \lceil 11/10 \rceil = 2$, hence at least two integers in T end in the digit i. □

Example 4.5.11. Consider a random number generator that returns any integer with equal probability. Conjecture the least number of integers that must be randomly generated in order to ensure that at least ten integers have the same remainder modulo 6; then, prove the conjecture.

Solution. Our "pigeons" in this case are the n integers whose remainders modulo 6 we seek. We require that ten of these n integers have the same remainder modulo 6, so our "coops" in this case must be the distinct remainders modulo 6, of which there are six. By the Pigeonhole Principle, it suffices to determine the least positive integer n such that $\lceil n/6 \rceil \geq 10$. Observe that if $n = 54$, then this inequality does not hold since $54/6 = 9$; however, if $n = 55$, then this inequality holds. $\diamond$

Proof. By the Division Algorithm, every integer can be written as $6q + r$ for some unique integers q and r such that $0 \leq r \leq 5$. Consequently, we obtain a partition $\mathbb{Z} = S_0 \cup S_1 \cup S_2 \cup \cdots \cup S_6$ such that $S_i = \{6q + i \mid q \in \mathbb{Z}\}$. By the Pigeonhole Principle, it suffices to find the least integer n such that $\lceil n/6 \rceil \geq 10$. Considering that $54 = 6 \cdot 9$, we must have that $n \geq 55$. We claim that $n = 55$ is the least number of integers that must be randomly generated in order to ensure that at least ten integers have the same remainder modulo 6. Once again, by the Pigeonhole Principle, every subset $T \subseteq \mathbb{Z}$ of cardinality $|T| = 55$ satisfies that $|T \cap S_i| \geq \lceil 55/6 \rceil = 10$ for some integer $0 \leq i \leq 5$, hence if we randomly generate 55 integers, then at least ten have the same remainder modulo 6. $\square$

Example 4.5.12. Prove that for any integer $n \geq 1$, if $n + 1$ integers are randomly selected from the set $S = \{1, 2, \ldots, 2n\}$, then some pair of integers selected is relatively prime to one another.

Solution. Unlike the previous examples, it is easiest in this case to first define a criterion with which to determine our "pigeons" and "coops." Concretely, according to the claim, we require that some pair of integers among the $n+1$ integers between 1 and $2n$ is relatively prime. Bearing in mind that this is a "weak" condition (since it only restricts the prime factorization of the integers), we may first note that every pair of consecutive integers is relatively prime, so we may view our "pigeons" as the elements of S and our "coops" as the sets $\{i, i + 1\}$ for each odd integer $1 \leq i \leq 2n - 1$. $\diamond$

Proof. Consider the partition $S = S_1 \cup S_3 \cup \cdots \cup S_{2n-1}$ such that $S_i = \{i, i+1\}$ for each odd integer $1 \leq i \leq 2n - 1$. Observe that there are n odd integers $1 \leq i \leq 2n - 1$, hence we have a partition of S into n subsets. Given any subset $T \subseteq S$ such that $|T| \geq n + 1$, we have that

$$T = T \cap S = T \cap (S_1 \cup S_3 \cup \cdots \cup S_{2n-1}) = (T \cap S_1) \cup (T \cap S_3) \cup \cdots \cup (T \cap S_{2n-1})$$

is a partition of T into n subsets. Consequently, by the Pigeonhole Principle, there exists an odd integer $1 \leq i \leq 2n - 1$ such that $|T \cap S_i| \geq \lceil (n + 1)/n \rceil = 2$, hence T contains two consecutive integers in S. Considering that consecutive integers are relatively prime, our proof is complete. $\square$

Example 4.5.13. Prove that among any five integers that are randomly selected from the finite set $S = \{1, 2, 3, \ldots, 10\}$, at least one of these integers is composite; then, determine the least number of integers randomly selected from S to guarantee that at least one integer selected is prime.

Proof. Consider the partition $S = S_1 \cup S_2$ such that $S_1 = \{1, 4, 6, 8, 9, 10\}$ consists of all composite integers of S and $S_2 = \{2, 3, 5, 7\}$ consists of all primes of S. Given any subset $T \subseteq S$, observe that $T = T \cap S = T \cap (S_1 \cup S_2) = (T \cap S_1) \cup (T \cap S_2)$ is a partition of T so that $|T| = |T \cap S_1| + |T \cap S_2|$. We note that if $|T| \geq 5$, then $|T \cap S_1| = |T| - |T \cap S_2| \geq 1$, hence T contains a composite integer. Conversely, if $|T| \leq 6$, then $|T \cap S_2| = |T| - |T \cap S_1| \leq 0$, hence $|T \cap S_2| \geq 1$ only if $|T| \geq 7$. $\square$

4.6 Permutations and Combinations

Last, we discuss the combinatorics of ordered and unordered arrangements of distinct objects chosen without repetition. Given any integer $n \geq 0$, we define the positive integer **n-factorial**

$$n! = \prod_{i=1}^{n} = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$$

as the product of all positive integers less than or equal to n. Conventionally, we set $0! = 1$. Even more, we note that for every integer $1 \leq k \leq n$, we have that $k!$ is a factor of $n!$ since it holds that

$$n! = n(n-1)(n-2)\cdots(k+1)k(k-1)(k-2)\cdots 3 \cdot 2 \cdot 1 = n(n-1)(n-2)\cdots(k+1)k!.$$

Explicitly, for any integer $0 \leq k \leq n$, we have that $n!$ is divisible by $(n-k)!$ since it holds that

$$n! = n(n-1)(n-2)\cdots(n-k+1)(n-k)!.$$

By dividing both sides of this identity by the integer $(n-k)!$, we find that

$$\frac{n!}{(n-k)!} = n(n-1)(n-2)\cdots(n-k+1).$$

Example 4.6.1. By definition of the factorial, we have that $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$.

Example 4.6.2. Observe that $7! = 7 \cdot 6 \cdot 5!$ is divisible by $5!$ and $\dfrac{7!}{5!} = 7 \cdot 6 = 42$.

Given any collection of n distinct objects, we refer to any ordered arrangement of these n distinct objects as a **permutation** of the objects. Even more, for any integer $0 \leq k \leq n$, we refer to any ordered arrangement of any k distinct objects among the n distinct objects as a **k-permutation** of the objects. We use the symbol $P(n,k)$ or ${}_nP_k$ for the count of k-permutations of n distinct objects.

Proposition 4.6.3 (Ordered Arrangements Without Repetition)**.** *Given any integers $0 \leq k \leq n$, the number of ordered arrangements of k objects taken without repetition from n distinct objects is*

$$P(n,k) = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

Proof. Consider the procedure of arranging k distinct objects from a collection of n distinct objects. Begin with k bins; then, successively place one object in each bin so that no bin contains more than one object. Considering that each of the k bins is initially vacant, we may place the first object in any of the n bins; then, because we cannot place two objects in one bin, we may place the second object in any of the remaining $n-1$ bins. Generally, for each integer $1 \leq i \leq k$, there are $n-i+1$ ways to place the ith object into a bin. Consequently, by the Multiplication Rule of Counting,

$$P(n,k) = \prod_{i=1}^{k}(n-i+1) = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}. \qquad \square$$

Example 4.6.4. Compute the number of seating arrangements of five students out of ten students.

Solution. Each of the ten students is distinct, hence there are $P(10,5)$ ways to seat five of them. $\diamond$

Example 4.6.5. Compute the number of six-letters words in English with no repeated letters.

Solution. Each of the 26 letters of the English alphabet is distinct, and we seek to order six of them without repetition. Consequently, there are $P(26, 6)$ six-letter words with no repeated letters. ◇

Example 4.6.6. Compute the number of seven-digit phone numbers with distinct digits.

Solution. Each of the ten digits $0, 1, 2, \ldots, 9$ is distinct, and we seek to order seven of them without repetition. Consequently, there are $P(10, 7)$ seven-digit phone numbers with distinct digits. ◇

Compared to a permutation in which order matters, a collection of n distinct objects in which order does not matter is a **combination**. Concretely, any unordered collection of n distinct objects is a combination. Consequently, a k-**combination** is any unordered selection of k distinct objects from n distinct objects. We write $C(n, k)$ or ${}_nC_k$ to count of k-combinations of n distinct objects.

Proposition 4.6.7 (Unordered Selections Without Repetition)**.** *Given any integers $0 \leq k \leq n$, the number of unordered selections of k objects taken without repetition from n distinct objects is*

$$C(n, k) = \frac{n!}{k!(n-k)!}.$$

Proof. Consider the procedure of arranging k distinct objects from a collection of n distinct objects. Begin by selecting k distinct objects from n distinct objects. By definition, there are $C(n, k)$ ways to perform task. Even more, by Proposition 4.6.3, there are then $k!$ ways to arrange these k distinct objects. Consequently, it follows that $P(n, k) = k!C(n, k)$ by the Multiplication Rule of Counting. By appealing once again to Proposition 4.6.3, we conclude the desired result that

$$C(n, k) = \frac{P(n, k)}{k!} = \frac{n!}{k!(n-k)!}.$$ □

Example 4.6.8. Compute the number of five-member committees chosen from eight people.

Solution. Considering that order does not matter, there are $C(8, 5)$ such committees. ◇

Example 4.6.9. Compute the number of six-bit binary strings with one digit of 1.

Solution. Considering that order does not matter, there are $C(6, 1)$ such six-bit binary strings. ◇

Example 4.6.10. Compute the number of six-bit binary strings with at most four digits of 1.

Solution. Observe that there are exactly five kinds of six-bit binary strings with at most four digits of 1: each one is determined by its number n of digits of 1 for each integer $0 \leq n \leq 4$. Each six-bit binary string is of the form $a_1a_2a_3a_4a_5a_6$ for some integers $0 \leq a_i \leq 1$, hence it suffices to select n digits of the six-bit string without repetition for each integer $0 \leq n \leq 4$. Considering that there are $C(6, n)$ ways to select n distinct objects without repetition among 6 distinct objects, there are $C(6, n)$ six-bit binary strings with n digits of 1. Consequently, the Addition Rule of Counting yields $C(6, 0) + C(6, 1) + C(6, 2) + C(6, 3) + C(6, 4)$ six-bit binary strings with at most four digits of 1. ◇

Conventionally, the number of k-combinations of n distinct objects is denoted by $C(n,k) = \binom{n}{k}$ and referred to as the **binomial coefficient** n **choose** k. Considering our previous examples, this terminology is justified: the number of k-combinations of n distinct objects is precisely the number of ways to select (or choose) k distinct objects from n distinct objects without repetition. Crucially, the binomial coefficients $C(n,k)$ determine the familiar **arithmetic triangle** as follows.

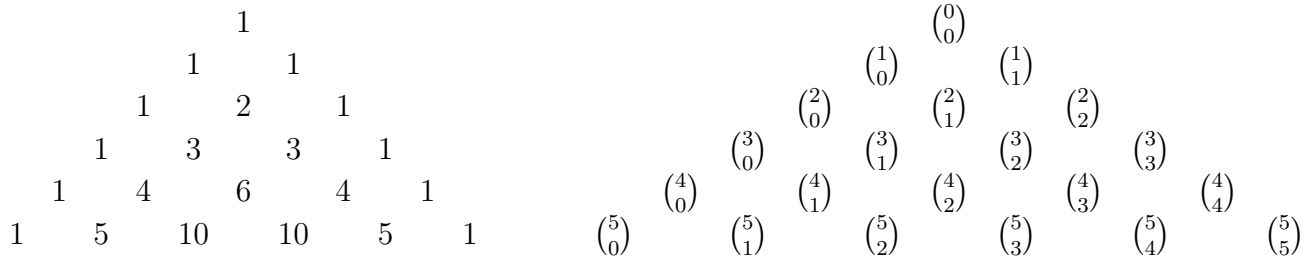


Figure 4.1: the Arithmetic Triangle

Quite remarkably, the binomial coefficients enjoy a delightful symmetry and many useful properties that have made them a classical object of study in combinatorics and commutative algebra.

Proposition 4.6.11 (Properties of Binomial Coefficients)**.** *Consider any integers* $0 \leq k \leq n$.

1.) *We have that* $C(n,k) = C(n,n-k)$. *Consequently, every row of the Arithmetic Triangle is a palindrome (i.e., it appears the same read forward as it does read backward).*

2.) *We have that* $C(n+1,k) = C(n,k-1) + C(n,k)$. *Consequently, the* k*th entry of the* $(n+1)$*th row of the Arithmetic Triangle can be determined as the sum of the* $(k-1)$*th and* k*th entry of the* n*th row of the Arithmetic Triangle. Geometrically, we can determine the* k*th entry of the* $(n+1)$*th row of the Arithmetic Triangle as the sum of the entries above-left and above-right.*

3.) *We have that* $C(n,k)$ *is the exponent of* $x^k y^{n-k}$ *in the polynomial expansion of the binomial* $(x+y)^n$. *Consequently, the terminology of "binomial coefficient" is justified.*

Proof. We leave the proof of the proposition as Exercise 4.8.22 for the reader. □

We have thus far only considered k-permutations without repetition; however, with some care, it is possible to remove this restriction. Given any collection of n objects such that n_1 objects are of one kind, n_2 objects are of a second kind, and so on up to n_k objects of a kth kind, the number of ways to arrange these $n = n_1 + n_2 + \cdots + n_k$ objects is denoted by the **multinomial coefficient**

$$P(n;\, n_1, n_2, \ldots, n_k) = \binom{n}{n_1, n_2, \ldots, n_k}.$$

Proposition 4.6.12 (Ordered Arrangements with Repetition)**.** *Given any non-negative integers* $n_1, n_2, \ldots, n_k$ *such that* $n = n_1 + n_2 + \cdots + n_k$, *the number of ordered arrangements of* n_1 *objects of one kind,* n_2 *objects of a second kind, and so on up to* n_k *objects of a* k*th kind is*

$$P(n;\, n_1, n_2, \ldots, n_k) = \binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}.$$

Proof. Considering each of the n objects in this collection as distinct, by Proposition 4.6.3, there are $n!$ ordered arrangements of these n objects under this convention. Conversely, each arrangement of these n objects is uniquely determined by the location of the objects of one kind in relation to the objects of another kind because objects of the same kind are not distinct. Consequently, it follows that $P(n;\, n_1, n_2, \ldots, n_k)$ can be determined by first choosing n_1 bins in which to place the objects of the first kind; then, choosing n_2 bins in which to place objects of the second kind; and so on for each integer $1 \leq i \leq k$. Crucially, this count involves unordered arrangements of objects of the ith kind for each integer $1 \leq i \leq k$; however, by labelling each of the n_i objects of the ith kind with a distinct label for each integer $1 \leq i \leq k$, we find that there are $n_i!$ ordered arrangements of the n_i objects of the ith kind for each integer $1 \leq i \leq k$. Consequently, by the Multiplication Rule of Counting, it follows that $n! = n_1!n_2! \cdots n_k!P(n;\, n_1, n_2, \ldots, n_k)$, and the desired formula holds. □

Example 4.6.13. Compute the number of 11-letter strings formed from the word MISSISSIPPI.

Solution. Considering that there is one copy of M, four copies of I, four copies of S, and two copies of P in MISSISSIPPI, it follows that the number of 11-letter strings formed from MISSISSIPPI is

$$P(11;\, 1, 4, 4, 2) = \binom{11}{1, 4, 4, 2} = \frac{11!}{1!4!4!2!} = 34650. \qquad \diamond$$

Example 4.6.14. Compute the number of 8-bit binary strings with an even number of digits of 1.

Solution. Considering that there are two distinct digits 0 and 1 in any 8-bit binary string, it follows that the number of 8-bit binary strings with an even number of digits of 1 is uniquely determined as a multinomial coefficient by the number of digits of 1; thus, the number of such 8-bit strings is

$$P(8;\, 0, 8) + P(8;\, 2, 6) + P(8;\, 4, 4) + P(8;\, 6, 2) + P(8;\, 8, 0) = 128. \qquad \diamond$$

Example 4.6.15. Compute the number of ways in which a fair coin can be flipped a total of eight times such that at least as many heads appear as tails appear after all eight flips.

Solution. Considering that a sequence of eight flips of a fair coin can be viewed as an eight-digit string using only the letters H (heads) and T (tails), it follows that the number of ways in which a fair coin can be flipped a total of eight times is uniquely determined as a multinomial coefficient by the number of times heads appears, i.e., the number of appearances of the letter H in the eight-digit sequence corresponding to the sequence of coin flips. Consequently, if we flip a fair coin eight times and we witness at least as many heads appear as tails appear, then we must record heads k times for some integer $4 \leq k \leq 8$. We conclude that the number of ways in which this occurs is

$$P(8;\, 8, 0) + P(8;\, 7, 1) + P(8;\, 6, 2) + P(8;\, 5, 3) + P(8;\, 4, 4) = 163. \qquad \diamond$$

Last, we fix our attention on combinations of objects selected with repetition as follows.

Proposition 4.6.16 (Unordered Selections with Repetition)**.** *Given any positive integers n and k, the number of unordered selections of k objects with repetition from a set of n distinct kinds of objects is equal to the number of unordered selections of $n-1$ distinct objects from $n+k-1$ objects*

$$\binom{n+k-1}{n-1}.$$

Proof. Given any k objects from a set of n distinct kinds of objects, we may first group the selected objects according to their kind and subsequently separate each group of like objects with a line. Effectively, we have placed the k objects into $n+k-1$ distinct bins, as depicted below.

$$\underbrace{*\,*\,\cdots\,*}_{\text{objects of the first kind}} \;\big|\; \underbrace{*\,*\,\cdots\,*}_{\text{objects of the second kind}} \;\big|\cdots\big|\; \underbrace{*\,*\,\cdots\,*}_{\text{objects of the } n\text{th kind}}$$

Concretely, we have one bin for each of the k objects in our selection based on its kind. Even more, some bins could be empty if no objects of that kind are selected. Between each group of objects of a different kind, we require a bin for the line that divides one kind of object from the other, hence there are a total of $n-1$ bins corresponding to these dividing lines. Consequently, in order to count the number of unordered selections of k objects from a set of n distinct kinds of objects, it suffices to count the number of ways to place the $n-1$ dividing lines into the $n+k-1$ distinct bins. Considering that these lines are not distinct, their placement corresponds to unordered selections of $n-1$ distinct objects taken without repetition from a collection of $n+k-1$ distinct objects. □

Remark 4.6.17. Bearing in mind the proof technique employed to prove the formula for Unordered Selections with Repetition, we will henceforth refer to any question involving unordered selections of objects with repetition as a **stars and bars** problem. (Compare with the illustration above.)

Example 4.6.18. Compute the number of baker's dozens of donuts formed from 20 kinds of donuts.

Solution. Each baker's dozen of donuts is uniquely determined by choosing 13 donuts from the 20 kinds of donuts with repetition. Consequently, the total number of baker's dozens of donuts is

$$\binom{20+13-1}{20-1} = \binom{32}{19} = \frac{32!}{19!13!} = 347373600.$$

Considering the astronomical number of possibilities, it is a wonder anyone could make a decision! ◇

Example 4.6.19. Compute the number of two-scoop bowls formed from ten flavors of ice cream.

Solution. Each bowl of ice cream is uniquely determined by choosing two flavors from the ten flavors with repetition. Consequently, the total number of two-scoop bowls of ice cream is

$$\binom{10+2-1}{10-1} = \binom{11}{9} = \frac{11!}{9!2!} = 55.$$

Counting in a different way, we could first determine the number of two-scoop bowls of ice cream with two distinct flavors and add this to the number of two-scoop bowls of ice cream with one flavor. Considering that order does not matter and repetition is not allowed, the number of two-scoop bowls of ice cream with distinct flavors is $\binom{10}{2}$. On the other hand, the number of two-scoop bowls of ice cream with the same flavor is 10. One can readily verify that $\binom{10}{2} + 10 = 45 + 10 = 55$. ◇

Example 4.6.20. Compute the number of non-negative integer solutions of the following equation.

$$x_1 + x_2 + x_3 + x_4 + x_5 = 10$$

Solution. By definition of a non-negative integer solution, we seek five integers x_1, x_2, x_3, x_4, and x_5 such that $x_i \geq 0$ for each integer $1 \leq i \leq 5$ and $x_1+x_2+x_3+x_4+x_5 = 10$. We may view this as a stars and bars problem since it is equivalent to placing 10 copies of 1 into the five distinct bins x_1, x_2, x_3, x_4, and x_5 with repetition. We conclude that there are $\binom{10+5-1}{5-1} = \binom{14}{4} = 1001$ solutions. ◇

4.7 Chapter 4 Overview

Collectively, **Principle of Mathematical Induction** comprises three equivalent statements: the Principle of Ordinary Induction, the Principle of Complete Induction, and the Well-Ordering Principle. We will assume that n_0 is an integer. Consider any open sentence $P(n)$ defined for all integers $n \geq n_0$. Concretely, the Principle of Ordinary Induction asserts that the statement $P(n)$ is true for all integers $n \geq n_0$ provided that the following pair of statements is true.

(a.) We have that $P(n_0)$ is a true statement.

(b.) If $P(n)$ is a true statement for some integer $n \geq n_0$, then $P(n+1)$ is a true statement.

Likewise, the Principle of Complete Induction asserts that the statement $P(n)$ is true for all integers $n \geq n_0$ provided that the following pair of statements is true.

(c.) We have that $P(n_0)$ is a true statement.

(d.) If $P(k)$ is a true statement for each integer $n_0 \leq k \leq n$, then $P(n+1)$ is a true statement.

One crucial benefit of using complete induction as opposed to ordinary induction is that the stronger hypotheses of complete induction provide more information with which to conveniently write proofs that might otherwise prove difficult with ordinary induction (see Exercise 4.8.3). Even more, the Principle of Mathematical Induction appears in the guise of the Well-Ordering Principle of the non-negative integers — a powerful tool that 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 a nonempty set, then there exists an element $s_0 \in S$ such that $s_0 \leq s$ for all elements $s \in S$.

Using the Well-Ordering Principle, we may rigorously establish that for any integer a and nonzero integer b, there exist unique integers q and r such that $a = qb + r$ and $0 \leq r \leq |b| - 1$; this fact is known as the Division Algorithm. We refer to the integer a as the **dividend**; b is the **divisor**; q is the **quotient**; and r is the **remainder** of a modulo b. Conventionally, if we obtain a remainder of zero when we divide an integer a by a nonzero integer b, then we say that b **divides** a; in this case, there exists a unique integer q such that $a = qb$, and we use the notation $b \mid a$. If a and b are any integers, then a nonzero integer c is called a **common divisor** of a and b if it holds that $c \mid a$ and $c \mid b$; the **greatest common divisor** of a and b is the unique integer $d = \gcd(a, b)$ such that

(a.) $d \mid a$ and $d \mid b$, i.e., d is a common divisor of a and b and

(b.) if c is any common divisor of a and b, then $c \mid d$.

We say that the nonzero integers a and b are **relatively prime** if and only if $\gcd(a, b) = 1$. Bézout's Identity asserts that $\gcd(a, b) = ax + by$ for some integers x and y. We may employ the Euclidean Algorithm to determine the integers x and y that are guaranteed by Bézout's Identity.

Using logical quantifiers allows us to conveniently state many properties of sets, e.g., the Law of the Excluded Middle for Sets, Law of Non-Contradiction for Sets, and De Morgan's Laws for Sets.

Counting principles like the Addition Rule of Counting, the Multiplication Rule of Counting, and the Pigeonhole Principle provide useful techniques to solve many problems in combinatorics. Combined with these are the more subtle formulas for permutations and combinations, e.g., the rules for handling Ordered Arrangements with Repetition and Unordered Selections Without Repetition as well as Ordered Arrangements with Repetition and Unordered Selections with Repetition.

4.8 Chapter 4 Exercises

Exercise 4.8.1. Given any integer $n \geq 0$, prove that $\binom{2n}{n} > 2^n$ using induction.

Exercise 4.8.2. Consider any finite set X with power set $P(X)$.

(a.) Prove that $|P(X)| = 2^{|X|}$ using induction.

(b.) Consider the collection 2^X of all functions $f : X \to X$. Construct an explicit bijection between $P(X)$ and 2^X. Conclude from part (a.) and Proposition 1.14.1 that $|2^X| = 2^{|X|}$.

Exercise 4.8.3. Consider the sequence of **Fibonacci numbers** F_n defined recursively for all non-negative integers by $F_0 = 0$, $F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$. We refer to F_n as the nth Fibonacci number. Quite astoundingly, the Fibonacci numbers appear abundantly in nature.

(a.) Prove that $F_n < 2^n$ for each integer $n \geq 0$.

(b.) Prove that $F_{n+1}F_{n-1} = F_n^2 + (-1)^n$ for each integer $n \geq 2$.

(c.) Prove that $\gcd(F_n, F_{n+1}) = 1$ for all integers $n \geq 0$.

Exercise 4.8.4. Complete the following two steps to prove that the Principle of Ordinary Induction and the Principle of Complete Induction are materially equivalent to one another.

1.) Given any statement $P(n)$ defined for a non-negative integer n, let $Q(n)$ be the statement that $P(k)$ holds for all integers $1 \leq k \leq n$. Use the Principle of Ordinary Induction to prove that the statement $Q(n)$ is true for all integers $n \geq 0$, hence $P(n)$ is true for all integers $n \geq 0$. Unravelling this shows that ordinary induction implies complete induction.

(**Hint:** Observe that $Q(0)$ is vacuously true, hence we may assume that $Q(n)$ is true. By definition, this means that $P(k)$ is true for all integers $1 \leq k \leq n$. What about $P(n+1)$?)

2.) Given any statement $P(n)$ defined for a non-negative integer n, let $Q(n)$ be the statement that $P(k)$ holds for some integer $1 \leq k \leq n$. Use the Principle of Complete Induction to prove that the statement $Q(n)$ is true for all integers $n \geq 0$, hence $P(n)$ is true for all integers $n \geq 0$. Unravelling this shows that complete induction implies strong induction.

(**Hint:** Observe that $Q(0)$ is vacuously true, hence we may assume that $Q(k)$ is true for all integers $1 \leq k \leq n$; in particular, $P(1)$ is true. What does this say about $Q(n+1)$?)

Exercise 4.8.5. Complete the following three steps to prove that the Well-Ordering Principle and the Principle of Ordinary Induction are materially equivalent to one another.

1.) Prove that 0 is the smallest non-negative integer with respect to $\leq$.

2.) Prove that if $S \subseteq \mathbb{Z}_{\geq 0}$ satisfies $0 \in S$ and $n + 1 \in S$ whenever $n \in S$, then $\mathbb{Z}_{\geq 0} \subseteq S$.

3.) Conclude that the Well-Ordering Principle implies the Principle of Ordinary Induction; then, use Exercise 4.8.4 in tandem with the proof of the Well-Ordering Principle to conclude conversely that the Principle of Ordinary Induction implies the Well-Ordering Principle.

Exercise 4.8.6. Recall that a positive integer p is **prime** if and only if the only integers that divide p are $\pm p$ and 1. Prove that if a and b are any integers such that $p \mid ab$, then $p \mid a$ or $p \mid b$.

(**Hint:** We may assume that $p \nmid a$ and show that $p \mid b$; now, use Bézout's Identity.)

Exercise 4.8.7. Given any integers a, b, and c, prove that if $\gcd(a, b) = 1$ and $a \mid bc$, then $a \mid c$.

Exercise 4.8.8. Prove that there are no positive integers a, b, and c such that $a^2 + b^2 = 3c^2$.

Exercise 4.8.9 (Fundamental Theorem of Arithmetic)**.** Given any positive integer a, prove that

(a.) there exist (not necessarily distinct) prime numbers $p_1, \ldots, p_k$ such that $a = p_1 \cdots p_k$ and

(b.) the primes $p_1, \ldots, p_k$ are unique in the sense that if $a = q_1 \ldots q_\ell$, then we must have that $\ell = k$ and $\{p_1, \ldots, p_k\} = \{q_1, \ldots, q_k\}$ (i.e., $q_1, \ldots, q_k$ are simply a rearrangement of $p_1, \ldots, p_k$).

(**Hint:** Consider the collection N of positive integers that **do not** possess such a prime factorization. Use the Well-Ordering Principle to show that if N is nonempty, then there exists a smallest element n with respect to $\leq$. Consider the factors of the positive integer n to conclude that N is empty, hence the existence is established. On the matter of uniqueness, proceed by induction on k.)

Exercise 4.8.10. Prove that every nonzero integer a can be written as $a = \pm p_1^{e_1} \ldots p_n^{e_n}$ for some distinct prime numbers $p_1, \ldots, p_n$ and unique non-negative integers $e_1, \ldots, e_n$ such that

Given any nonzero integers a and b, the **least common multiple** $\operatorname{lcm}(a, b)$ of a and b can be defined in a manner analogous to the greatest common divisor of a and b. Explicitly, we say that an integer m is a **multiple** of a if and only if $a \mid m$. Consequently, m is a **common multiple** of a and b if and only if $a \mid m$ and $b \mid m$; a least common multiple of a and b is $\ell = \operatorname{lcm}(a, b)$ such that

(1.) $a \mid \ell$ and $b \mid \ell$, i.e., ℓ is a common multiple of a and b and

(2.) if ℓ' is any common multiple of a and b, then $\ell \mid \ell'$.

Exercise 4.8.11. Prove that the least common multiple $\operatorname{lcm}(a, b)$ is unique up to sign.

By the Fundamental Theorem of Arithmetic, for any positive integers a and b, there exist prime numbers $p_1, \ldots, p_k$ and unique non-negative integers $e_1, \ldots, e_k, f_1, \ldots, f_k$ such that $a = p_1^{e_1} \cdots p_k^{e_k}$ and $b = p_1^{f_1} \cdots p_k^{f_k}$. Consider these prime factorizations of a and b for the next three exercises.

Exercise 4.8.12. Prove that $\gcd(a, b) = p_1^{\min\{e_1, f_1\}} \cdots p_k^{\min\{e_k, f_k\}}$.

Exercise 4.8.13. Prove that $\operatorname{lcm}(a, b) = p_1^{\max\{e_1, f_1\}} \cdots p_k^{\max\{e_k, f_k\}}$.

Exercise 4.8.14. Conclude from Exercises 4.8.12 and 4.8.13 that $ab = \gcd(a, b) \operatorname{lcm}(a, b)$.

Exercise 4.8.15. Consider any sets W, X, and Y such that $X \subseteq W$ and $Y \subseteq W$.

(a.) Prove that for any subset $Z \subseteq W$ such that $Z \supseteq X$ and $Z \supseteq Y$, it follows that $Z \supseteq X \cup Y$. Conclude that $U = X \cup Y$ is the "smallest" subset of W containing both X and Y.

(b.) Prove that for any subset $Z \subseteq W$ such that $Z \subseteq X$ and $Z \subseteq Y$, it follows that $Z \subseteq X \cap Y$. Conclude that $I = X \cap Y$ is the "largest" subset of W contained in both X and Y.

Consider the relative complement $X' = W \setminus X$ of X in W. We may sometimes refer to X' simply as the **complement** of X if we are dealing only with subsets of W, i.e., if W is our universe.

(c.) Prove that $Y \setminus X = Y \cap X'$. Use part (b.) above to conclude that $C = Y \cap X'$ is the "largest" subset of W that is contained in Y and disjoint from X.

Exercise 4.8.16. Prove the second of the Distributive Laws for Sets.

Exercise 4.8.17. Consider any function $f : X \to Y$ from a set X to a set Y.

(a.) Prove that $f(U \cup V) = f(U) \cup f(V)$ for any sets $U, V \subseteq X$.

(b.) Prove that $f(\cup_{i \in I} V_i) = \cup_{i \in I} f(V_i)$ for any index set I and any sets $V_i \subseteq X$.

(c.) Prove that $f(U \cap V) = f(U) \cap f(V)$ for any sets $U, V \subseteq X$.

(d.) Prove that $f(\cap_{i \in I} V_i) = \cap_{i \in I} f(V_i)$ for any index set I and any sets $V_i \subseteq X$.

(e.) Prove that $f^{-1}(V \cup W) = f^{-1}(V) \cup f^{-1}(W)$ for any sets $V, W \subseteq Y$.

(f.) Prove that $f^{-1}(\cup_{i \in I} W_i) = \cup_{i \in I} f^{-1}(W_i)$ for any index set I and any sets $W_i \subseteq Y$.

(g.) Prove that $f^{-1}(V \cap W) = f^{-1}(V) \cap f^{-1}(W)$ for any sets $V, W \subseteq Y$.

(h.) Prove that $f^{-1}(\cap_{i \in I} W_i) = \cap_{i \in I} f^{-1}(W_i)$ for any index set I and any sets $W_i \subseteq Y$.

Exercise 4.8.18. Consider any function $f : X \to Y$ from a set X to a set Y.

(a.) Prove that $V \subseteq f^{-1}(f(V))$ for any set $V \subseteq X$.

(b.) Exhibit sets $V \subseteq X$ and Y and a function $f : X \to Y$ such that $f^{-1}(f(V)) \not\subseteq V$.
(**Hint:** By Proposition 4.4.9, $f : X \to Y$ cannot be injective.)

(c.) Prove that $f(f^{-1}(W)) \subseteq W$ for any set $W \subseteq Y$.

(d.) Exhibit sets X and $W \subseteq Y$ and a function $f : X \to Y$ such that $W \not\subseteq f(f^{-1}(W))$.
(**Hint:** By Proposition 4.4.9, $f : X \to Y$ cannot be surjective.)

Exercise 4.8.19. Consider any function $f : X \to Y$ from a set X to a set Y.

(a.) Prove that if $f^{-1}(f(V)) = V$ for any set $V \subseteq X$, then f is injective.
(**Hint:** If $f(x_1) = f(x_2)$, then consider the set $V = \{x_1\}$.)

(b.) Prove that if $f(f^{-1}(W)) = W$ for any set $W \subseteq Y$, then f is surjective.
(**Hint:** Consider the set $W = Y$; then, use the definition of $f(f^{-1}(W))$.)

Exercise 4.8.20. Given any set X, consider the **diagonal function** $\delta_X : X \to X \times X$ defined by $\delta_X(x) = (x, x)$ and the **diagonal relation** $\Delta_X = \{(x, x) \mid x \in X\}$ on X. Prove that $\Delta_X = \delta_X(X)$.

Exercise 4.8.21. Given any prime number p, consider the collection $\mathbb{Z}_p$ of equivalence classes of the integers modulo p. Prove that $[a]$ admits a multiplicative inverse if and only if $p \nmid a$.

Exercise 4.8.22. Prove the three Properties of Binomial Coefficients using the definition of $\binom{n}{k}$.

Exercise 4.8.23. Given any integer $n \geq 0$, prove that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

(**Hint:** Give a combinatorial description of what 2^n counts; then, show that the sum on the left-hand side counts the same thing. We will henceforth refer to this as a **combinatorial proof**.)

Exercise 4.8.24. Given any integers $0 \leq n \leq k$, prove that

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Conclude that $\binom{2n}{n} \geq 2^n$ for all integers $n \geq 0$.

(**Hint:** Construct a combinatorial proof by illustrating that these quantities count the same thing.)

Exercise 4.8.25. Compute the number of n-digit integers without a digit of 2, 4, 6, or 8.

Exercise 4.8.26. Compute the number of three-digit integers constructed from the positive integers 1, 2, 3, 5, and 7; then, determine the number of such three-digit integers that are divisible by 5.

Exercise 4.8.27. Compute the number of three-letter strings constructed from the English alphabet without repetition using both uppercase letters and lowercase letters.

Exercise 4.8.28. Compute the number of four-letter English bracelets without repetition.

Exercise 4.8.29. Compute the number of three-letter sets constructed from the English alphabet.

Exercise 4.8.30. Compute the number of three-character sets constructed from the enriched English alphabet, i.e., the English alphabet with the special characters !, @, #, and $.

Exercise 4.8.31. Compute the number of 7-letter strings formed from the word FALAFEL.

Exercise 4.8.32. Compute the number of 5-letter strings formed from the word PHILIPPINES.

Example 4.8.33. Compute the number of sextuples of non-negative integers that sum to 7.

Exercise 4.8.34. Compute the number of non-negative integer solutions of the following equation.

$$\sum_{i=1}^{13} x_i = 113.$$

References

[CPZ18] G. Chartrand, A.D. Polimeni, and P. Zhang. *Mathematical Proofs: a Transition to Advanced Mathematics*. 4th ed. Pearson Education, Inc., 2018.

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

[Gri99] R.P. Grimaldi. *Discrete and Combinatorial Mathematics: an Applied Introduction*. 4th ed. Addison Wesley Longman, Inc., 1999.

[Ham13] R. Hammack. *Book of Proof*. 2.2. Richard Hammack, 2013.