Abstract Algebra

C. Caruvana

Fall 2025

Contents

1 Brief History and A Motivating Example

The following historical commentary has been gathered from the Wikipedia articles on History of Algebra and Abstract Algebra.

The word *algebra* is derived from the Arabic *al-jabr* which appeared in the title of a treatise written in 830 by Al-Khwarizmi, a Persian mathematician. The treatise itself was about linear and quadratic equations. Note, however, that societies all around the world had independently developed their own studies of solving algebraic equations well before Al-Khwarizmi's treatise.

For the ancient Babylonians and the Greeks, algebraic concepts were largely geometric. In fact, Greek and Vedic Indian mathematicians used geometry to solve certain algebraic equations. Much later, Descartes (1596-1650) introduced modern notation and showed that problems of geometry can be expressed and solved in terms of algebra.

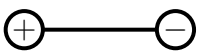
Diophantus and Brahmagupta, independently, moved away from the geometric perspective toward one focused on finding numbers that satisfied given equations.

Sharaf al-Din al-Tusi began the transition from static equation solving to a more functional approach, where functions are seen as dynamic entities representing motion.

It is not until the 19th and 20th centuries that *abstract algebra* is developed.

Abstract algebra, which we understand as the study of "algebraic structure," emerged through the identification of common themes in problems of number theory, geometry, analysis, and algebraic equation solving. In this semester, we will be focusing on a branch of abstract algebra known as *group theory*. Loosely speaking, a group is a collection of objects with a particular kind of operation. We will define them carefully after considering some motivating examples and setting up our mathematical formalism.

Example 1.1. Consider, for example, the possible symmetric positions of a line segment:



There are two possible actions on this structure: "do nothing" or reversing the poles. Note that reversing the poles twice has the same effect as "doing nothing."

Another way to model this example is to use "words." Let P represent the "positive" side and N represent the "negative" side. As presented, the word PN represents the initial orientation. Reversing the poles yields the word NP. Thus, the possible states are thus PN and NP.

We can also abstract away the "actions" here of *doing nothing* and *reversing the poles*. Let 0 represent the action of "doing nothing" and 1 represent the action of "reversing the poles." In the following table, consider the rows as indicating which of the two actions to be applied first, and then the columns indicate which operation is done afterwards.

	0	1
0	0	1
1	1	0

As noted above and displayed in the second row, second column of the table, reversing the poles twice results in the same action as the "doing nothing" action.

2 Brief Commentary on Mathematical Proofs

One of the primary activities in this course is proof-writing. Mathematical proofs are formal arguments that establish the logical necessity of a given statement. Generally speaking, a *statement* is any expression which has a truth value. For example, "2 is even" is a statement (and is true) and "$12 + 5 = 60$" is a statement (which is false). An expression like $x + 5$ is not a statement because there is not a coherent way to assign a truth value to it.

The general structure of discourse here consists of two things: there are the objects of discourse and there are statements referring to the objects.

Of primary importance are conditional statements of the form "if p, then q," where p and q are statements themselves. We will also deal with quantified statements:

- $\forall x\ P(x)$ is the statement "for every x, x satisfies P," and
- $\exists x\ P(x)$ is the statement "there exists an x such that x satisfies P,."

2.1 Direct Proof

A standard mathematical proof proving that the implication "if p, then q" is true starts by assuming p and deducing the logical necessity of the truth of q. We begin with a very basic example, introducing the conventional proof-writing format in the process.

Claim. Suppose x represents a real number. If $3x + 2 = 17$, then $x = 5$.

Proof. Suppose $3x + 2 = 17$. By subtracting 2 from both sides of the equation, we obtain that $3x = 15$. Then, dividing by 3 on both sides of the equation yields that $x = 5$. □

Formal proofs are written in a narrative fashion using the expected natural language, sometimes assisted by some mathematical symbols. Like anything else, proof-writing is a skill that improves with practice.

2.2 Proof by Contrapositive

Another valid proof technique for proving "if p, then q," is to prove what is known as the corresponding *contrapositive* statement: "if it's not true that q, then it's not true that p."

Let's see a basic example.

Claim. Suppose x represents a real number. If $x^3 - 1 < 0$, then $x < 1$.

Proof. We proceed by way of the contrapositive. So, suppose $x \geq 1$. It follows that $x^3 \geq 1$ and, then, that $x^3 - 1 \geq 0$. □

Sanity Check 2.1. If x is a real number and $x \geq 1$, why is it that $x^3 \geq 1$?

Exercise 2.2. An integer n is said to be *even* if there is an integer k such that $n = 2k$. In other words, n is even if 2 divides n with no remainder. Prove that, given an integer n, if n^2 is even, then n is even.

2.3 Proof by Contradiction

Proof by contradiction, also known as *reductio ad absurdum*, is an ancient argument style used often in Platonic dialogues. The basic format is this: First you pose the objection, "suppose what you claim to be true is false." You then attempt to deduct a logical impossibility. The standard introductory proof by contradiction (known certainly to the ancient Greeks) is that of the irrationality of $\sqrt{2}$.

Claim. The number $\sqrt{2}$ is not rational.

Proof. By way of contradiction, suppose that $\sqrt{2}$ is rational. By taking out any common factors, we can thus write $\sqrt{2} = \dfrac{p}{q}$ where p and q are integers with no common factors. Squaring both sides, we obtain that $2 = \dfrac{p^2}{q^2}$ and, thus, that $2q^2 = p^2$. It follows that p^2 is even and, thus, by Exercise 2.2, p is even. That is, $p = 2k$ for some integer k. It follows that

$$\begin{aligned} 2q^2 &= (2k)^2 \\ &= 4k^2. \end{aligned}$$

Dividing by 2 yields $q^2 = 2k^2$. Hence, q^2 is even which, again, by Exercise 2.2, asserts that q is even. We now see that p and q are both even, contradicting the assumption that p and q had no common factors. Therefore, $\sqrt{2}$ is irrational. □

Sanity Check 2.3. Why can we write ratios of integers in so-called reduced terms? (*Hint.* See the *Fundamental Theorem of Arithmetic.*)

3 Sets, Relations, and Functions

The theory of sets provides us with a twofold benefit: they offer us a foundation on which to formalize our mathematics and they off us a relatively simple structure with which to continue practicing the basics of proof-writing.

3.1 Sets

A *set* is a well-defined collection of primitive objects. In the case of pure set theory, the *only* primitive objects are sets, themselves. In the language of set theory, we use the symbol $\in$ as a relation between an object (it can be a set) x and a set A in the following way: $x \in A$ is the statement that x is an element of the set A. For example, if we say that X is the set of all positive real numbers, then $1 \in X$ and $-5 \notin X$.

Guided by convention, we use $\mathbb{R}$ to refer to the set of all real numbers, $\mathbb{Z}$ to refer to the set of all integers, $\mathbb{Q}$ to refer to the set of all rational numbers, and $\mathbb{N}$ to refer to the set of all natural numbers. Not all mathematicians agree whether $\mathbb{N}$ contains 0 or not. To align with our chosen textbook, we will observe the convention that $\mathbb{N}$ consists of only the positive integers.

One can define (finite) or refer to sets using *set-roster* notation; e.g. $X = \{a, b, c\}$. Here, X is asserted to be the set containing as its elements a, b, and c.

The most common method to define or refer to sets is the *set-builder* notation;

$$X = \{x \in \mathbb{R} : x > 0\}.$$

Here, X is asserted to be the set of positive real numbers. The expression $\{x \in \mathbb{R} : x > 0\}$ can be read as:

> ➢ "The set of real numbers x such that x is positive."

Another set that will be important to us is the set $\mathbb{C}$ consisting of all complex numbers; that is, $\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}$ where i is chosen to be a solution to the equation $x^2 + 1 = 0$.

Note that our expression for $\mathbb{C}$ above is not of the form $\{x \in X : P(x)\}$ where X is a set and P is a *predicate* of x. We won't concern ourselves here with the details of this notation and will simply accept it as a valid incarnation of set-builder notation.

Definition 3.1. Two sets A and B are said to be *equal*, denoted $A = B$, if they consist of exactly the same elements.

Definition 3.2. For two sets A and B, we say that B is a *subset* of A, denoted by $B \subseteq A$, if every element of B is an element of A. If $B \subseteq A$ and $B \neq A$, then we say that B is a *proper subset* of A, which we may denote by $B \subsetneq A$.[1]

Pro Tip. Two sets A and B are equal if and only if $A \subseteq B$ and $B \subseteq A$. Hence, a common strategy for proving that two sets are equal is to prove that they are both subsets of the other.

Definition 3.3. The *empty set*, denoted by either $\emptyset$ or $\{\}$, is the unique set containing no elements.

3.2 Operations with Sets

We can form new sets from old with some basic operations.

Definition 3.4. Given two sets A and B, the *union* of A and B, denoted by $A \cup B$, is the set consisting of all x such that either $x \in A$ or $x \in B$.

Comment. In mathematics, the "or" is always the *inclusive* or. So, a phrase of the form "p or q" is only to be interpreted as meaning "p, or q, or both." We will address one way to deal with the exclusive or below.

Definition 3.5. Given two sets A and B, the *intersection* of A and B, denoted by $A \cap B$, is the set consisting of all x such that $x \in A$ and $x \in B$.

Exercise 3.1. Prove that, for sets A and B,

- $A \cup B = B \cup A$ and
- $A \cap B = B \cap A$.[2]

Exercise 3.2. Prove that, for sets A, B, and C,

- $A \cup (B \cup C) = (A \cup B) \cup C$ and
- $A \cap (B \cap C) = (A \cap B) \cap C$.[3]

Definition 3.6. Given two sets A and B, we define the *set difference* $A \setminus B$ to be

$$\{x \in A : x \notin B\}.$$

When a given context U (referred to as the *universal set*) is understood, we define the *complement* of a set $A \subseteq U$ to be $A' = U \setminus A$.

[1] Our chosen textbook uses the notation $\subset$ in place of $\subseteq$. As such, we must avoid using the visual relationships between $\leq$ and $\subseteq$, and $<$ and $\subset$ here.

[2] That is, the union and intersection operations are *commutative*.

[3] That is, the union and intersection operations are *associative*.

The *symmetric difference* between two sets exemplifies the concept of the exclusive or. The standard convention for the symmetric difference is as follows:

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

We will return to this later.

Definition 3.7. Two sets A and B are said to be *disjoint* if $A \cap B = \emptyset$.

The operations of union and intersection extend well beyond contexts where only two sets are involved.

Definition 3.8. Suppose $\mathcal{A}$ is a set of sets (that is, each $A \in \mathcal{A}$ is a set). We define

$$\bigcup \mathcal{A} = \bigcup_{A \in \mathcal{A}} A = \{x : \exists A \in \mathcal{A}\ (x \in A)\}.$$

Similarly, we define

$$\bigcap \mathcal{A} = \bigcap_{A \in \mathcal{A}} A = \{x : \forall A \in \mathcal{A}\ (x \in A)\}.$$

These will often show up over *countable* sets; e.g. if A_n is a set for each $n \in \mathbb{N}$, then

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n \in \mathbb{N}} A_n = \{x : \exists n \in \mathbb{N}\ (x \in A_n)\}.$$

Proposition 3.9. For any set A,

- $A \cup A = A \cap A = A \cup \emptyset = A$ and
- $A \setminus A = A \cap \emptyset = \emptyset$.

Proposition 3.10 (Distributive Properties)**.** For sets A, B, and C,

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Theorem 3.11 (DeMorgan's Laws)**.** For any two subsets A and B of a given universal set U,

- $(A \cap B)' = A' \cup B'$ and
- $(A \cup B)' = A' \cap B'$.

Exercise 3.3. Prove Theorem 3.11.

3.3 Cartesian Products, Relations, and Functions

Cartesian products inherit their name from Descartes.

Definition 3.12. Given sets X and Y, the *Cartesian product* of X with Y is

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

For example, the usual coordinate plane is $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Definition 3.13. For sets X and Y, a *relation* between X and Y is a subset of $X \times Y$. When $R \subseteq X \times Y$, we sometimes use the notation xRy to mean $(x, y) \in R$.

Definition 3.14. Given two sets X and Y, a *function* f with *domain* X and *codomain* Y is a relation between X and Y with the following properties:

- For every $x \in X$, there is some $y \in Y$ such that $(x, y) \in f$.[4]
- For every $x \in X$ and $y, z \in Y$, if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$.[5]

We use the notation $f : X \to Y$ to denote that f is a function from X to Y. We will also use the notation $f(x)$ to denote the unique member of Y such that $(x, f(x)) \in f$. The *image* or *range* of a function $f : X \to Y$ is defined to be $f[X] = \{f(x) : x \in X\}$.[6]

Comment. We will say that two functions f and g are *equal*, denoted $f = g$, if they are equal *as sets*.

Example 3.15. The relation

$$R = \{(x, y) \in \mathbb{R}^2 : x^2 = y^4\}$$

is not a function. Consider the fact that both $(1, 1) \in R$ and $(1, -1) \in R$.

You may recall exercises from previous courses with the flavor of "rewrite *blah* as a function of *blech*" which involved algebraic manipulation of expressions. For example, in the equation $x = y^2$, x can be seen as a function of y, but y *cannot* be seen as a function of x. In these kinds examples, the issue of the domain tends to be ignored due to course focus, and they are typically intervals of real numbers, anyway.

We would like to point out here, though, that there are situations in which an equation can express y as a function of x, but cannot be rewritten using familiar algebraic tools in the form $y = f(x)$.

[4] In some contexts, *partial* functions are used and only have to satisfy the second property listed here (that is, they need not have full domain).

[5] This is formally stating that there is a unique element $y \in Y$ which satisfies $(x, y) \in f$. In less formal language, this is a version of the Vertical Line Test.

[6] Note the divergence in the notation here from the chosen textbook's. Either is acceptable, though it is sometimes useful to distinguish between a function being applied to particular elements of its domain and sets consisting of such applications.

Exercise 3.4. Show that the relation

$$R = \{(x, y) \in \mathbb{R}^2 : x = y + y^5\}$$

is a function. (*Hint.* You can use techniques from Calculus to show that $y = x + x^5$ is a bijection. Then, by the discussion below, it must have an inverse function, and that inverse function is exactly R.)

Before we get to function inverses, we will need the notion of *composition*.

Definition 3.16. Suppose $R \subseteq X \times Y$ and $S \subseteq Y \times Z$. The *composition* relation $S \circ R \subseteq X \times Z$ is defined to be

$$S \circ R = \{(x, z) \in X \times Z : \exists y \in Y \ ((x, y) \in R, (y, z) \in S)\}.$$

Exercise 3.5. Suppose $f : X \to Y$ and $g : Y \to Z$. Show that the composition $g \circ f$ is a function $X \to Z$.

Since relations serve as a broader context than functions, we start by defining the *inverse* of a relation, which always exists and is, itself, a relation.

Definition 3.17. Given $R \subseteq X \times Y$, we define the *inverse relation* of R to be

$$R^{-1} = \{(y, x) \in Y \times X : (x, y) \in R\}.$$

Under what conditions is R^{-1} guaranteed to be a function from Y to X? Examining Definition 3.14 in this context should naturally draw one to the following notions.

Definition 3.18. Suppose $f : X \to Y$ is a function. Then

- f is *injective* or *one-to-one* provided that, for $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$, then $x_1 = x_2$.[7]
- f is *surjective* or *onto* if, for every $y \in Y$, there is some $x \in X$ such that $f(x) = y$.
- f is *bijective* if it is both injective and surjective.

Definition 3.19. For any space X, we define the *identity* map $\mathrm{id}_X : X \to X$ by the rule $\mathrm{id}_X(x) = x$. Note that id_X is a bijection.

Though the following does not align yet with Definition 3.17, we present it in this incarnation for the sake of familiarity.

Definition 3.20. Suppose $f : X \to Y$ and $g : Y \to X$. The function g is said to be the *inverse* function of f, denoted by f^{-1}, if $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. In the case that a function $f : X \to Y$ has an inverse function, f is said to be *invertible*.

Theorem 3.21. A function is invertible if and only if it is bijective.

[7] This is a version of the Horizontal Line Test.

In other words, a function is invertible if and only if its inverse relation is, itself, a function.

Example 3.22. Consider $S : \mathbb{R} \to [0, \infty)$ defined by $S(x) = x^2$. Note that S is surjective but not injective. Hence, S is not invertible.

Note, however, that the standard square root function sqrt : $[0, \infty) \to [0, \infty)$, which chooses, for each $y \geq 0$, the *non-negative* solution to $x^2 = y$, is a function with the property that $S \circ \text{sqrt} = \text{id}_{[0,\infty)}$. On the other hand, for any $x \in \mathbb{R}$, $\text{sqrt} \circ S(x) = |x|$, the absolute value of x.

Exercise 3.6. Define $f : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}$ by the rule $f(k, n) = \dfrac{k}{n}$. Show that f is surjective but not injective.

Even when we restrict our attention to invertible functions, the composition operation fails to be commutative.

Example 3.23. Consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x + 2$ and $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^3$. Note that

$$g \circ f(x) = (x + 2)^3$$

and that

$$f \circ g(x) = x^3 + 2.$$

Since $g \circ f(0) = 8$ and $f \circ g(0) = 2$, we see that $g \circ f \neq f \circ g$.

Exercise 3.7. Find a pair of functions $f, g : \mathbb{R} \to \mathbb{R}$ that aren't inverses of each other such that neither is the identity function $\text{id}_{\mathbb{R}}$ and the equation $g \circ f = f \circ g$ is satisfied.

Exercise 3.8. Suppose $f : W \to X$, $g : X \to Y$, and $h : Y \to Z$. Prove that composition is associative; that is, show that $(h \circ g) \circ f = h \circ (g \circ f)$.

Exercise 3.9. Suppose $f : X \to Y$ and $g : Y \to Z$. Prove each of the following.

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

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

(c) If f and g are both bijections, then so is $g \circ f$.

Exercise 3.10. Consider the set $X = \{0, 1\}$ consisting of two elements. List all of the bijections of X. Can you see any similarity here with Example 1.1?

Another important concept related to functions is that of pre-images or fibers, though the standard notation overloads the notation for function inverses.

Definition 3.24. For a function $f : X \to Y$, the *pre-image* or *fiber* of a point $y \in Y$ is defined to be

$$f^{-1}(y) = \{x \in X : f(x) = y\}.$$

So a function is invertible if and only if every fiber consists of exactly one element of the domain.

Sanity Check 3.11. When f is invertible, what is the discrepancy between $f^{-1}(y)$ used as the *inverse function value at* y and $f^{-1}(y)$ used as the *fiber of* f *at* y?

3.4 Equivalence Relations

Definition 3.25. For a set X, a relation $\simeq \subseteq X \times X$ is said to be an *equivalence relation* if the following three properties hold:

- (Reflexivity) For every $x \in X$, $x \simeq x$.
- (Symmetry) For every $x, y \in X$, if $x \simeq y$, then $y \simeq x$.
- (Transitivity) For every $x, y, z \in X$, if both $x \simeq y$ and $y \simeq z$, then $x \simeq z$.

In this case, we will often use the phrasing "$\simeq$ is an equivalence relation *on* X."

The equality relation itself is an equivalence relation.

Exercise 3.12. Show that, if $R \subseteq X \times X$ is an equivalence relation, then

- $R^{-1} = R$ and
- $R \circ R = R$.

Exercise 3.13. Suppose $f : X \to Y$ and define $\simeq$ on X by the following rule: $x \simeq y$ provided that $f(x) = f(y)$. Show that $\simeq$ is an equivalence relation.

Exercise 3.14. Let $X = \mathbb{Z} \times \mathbb{N}$ and define $(k, m) \simeq (\ell, n)$ by

$$kn = \ell m.$$

Show that $\simeq$ is an equivalence relation.[8]

Definition 3.26. Suppose $\simeq$ is an equivalence relation on a set X. For $x \in X$, we define the $\simeq$-*equivalence class* (or simply called the *equivalence class* when there is no possibility for confusion) to be

$$[x]_{\simeq} = \{y \in X : x \simeq y\}.$$

When the equivalence relation $\simeq$ is understood, we may suppress the subscript and refer to the equivalence class of x as $[x]$.

Comment. Using the notion of equivalence classes, we can formally define the rational numbers to be the set of equivalence classes of the $\simeq$ relation defined in Exercise 3.14.

Exercise 3.15. Suppose $\simeq$ if an equivalence relation on a set X. Show that, for any $x, y \in X$, either $[x] = [y]$ or $[x] \cap [y] = \emptyset$.

Example 3.27. Let $n \geq 2$, $n \in \mathbb{Z}$, and define, for $p, q \in \mathbb{Z}$, $p \equiv q \pmod{n}$ if there exists $k \in \mathbb{Z}$ with $p - q = nk$. Then $\cdot \equiv \cdot \pmod{n}$ is an equivalence relation.

[8] In fact, it is the standard notion of equivalence for rational numbers.

Proof. For $p \in \mathbb{Z}$, note that $p - p = 0 = n \cdot 0$. Since $0 \in \mathbb{Z}$, $p \equiv p \pmod{n}$.

For symmetry, suppose $p \equiv q \pmod{n}$. By definition, that means there is some $k \in \mathbb{Z}$ for which $p - q = nk$. Note that $q - p = -nk = n \cdot (-k)$. Since $-k \in \mathbb{Z}$, we see that $q \equiv p \pmod{n}$.

Finally, for transitivity, suppose $p \equiv q \pmod{n}$ and $q \equiv r \pmod{n}$. Then let $k, \ell \in \mathbb{Z}$ be such that $p - q = nk$ and $q - r = n\ell$. Observe that

$$\begin{aligned} p - r &= p - q + q - r \\ &= nk + n\ell \\ &= n(k + \ell). \end{aligned}$$

Since $k + \ell \in \mathbb{Z}$, we see that $p \equiv r \pmod{n}$, finishing the proof. □

For each $n \geq 2$, $n \in \mathbb{Z}$, we define $\mathbb{Z}_n$ to be the set of equivalence classes for the equivalence relation $\cdot \equiv \cdot \pmod{n}$.

Exercise 3.16. Consider $\mathbb{Z}_2$.

(a) Show that,

- for any even $n \in \mathbb{Z}$, $[n] = [0]$ and
- for any odd $n \in \mathbb{Z}$, $[n] = [1]$.

Conclude that $\mathbb{Z}_2$ consists of exactly two elements.

(b) Consider addition in $\mathbb{Z}_2$ to be defined as $[x] + [y] = [x + y]$. Can you see any similarity here with Example 1.1?

Let $\mathbb{Z}^{\mathbb{Z}}$ denote the set of all functions $\mathbb{Z} \to \mathbb{Z}$.[9] For each $k \in \mathbb{Z}$, consider the function $g_k : \mathbb{Z} \to \mathbb{Z}$ defined by $g_k(n) = n + k$. Note that each g_k is a bijection and that $g_0 = \mathrm{id}_{\mathbb{Z}}$. We can also see each g_k as a shifting of $\mathbb{Z}$. For example, if we envision $\mathbb{Z}$ on a horizontal line ordered in increasing order, g_3 has the effect of shifting all points over by three units to the right, g_0 has the effect of doing nothing, and g_{-5} has the effect of shifting all points over by four units to the left.

Given the symbols above, we can define $\varphi : \mathbb{Z} \to \mathbb{Z}$ by $\varphi(k) = g_k$.[10] We will return to this idea later.

4 Introducing Groups

4.1 Getting to Know Groups

Definition 4.1. Let G be a set and $\diamond : G \times G \to G$ be a function, also referred to as a *binary operation*. The set G with the binary operation $\diamond$, $(G, \diamond)$, is said to be a *group* if the following properties hold:

[9] There is a connection here with Cartesian products, but we won't get into those details now.

[10] Formally speaking, g is already a function from $\mathbb{Z}$ to $\mathbb{Z}^{\mathbb{Z}}$.

- (Associativity) For any $a, b, c \in G$, $a \diamond (b \diamond c) = (a \diamond b) \diamond c$.
- (Identity) There exists $e \in G$ such that, for every $g \in G$, $g \diamond e = e \diamond g = g$. Any such $e \in G$ is referred to as the[11] *identity*.
- (Inverses) For any $g \in G$, there is some $h \in G$ such that $g \diamond h = h \diamond g = e$, where e is the identity. As will be shown in Proposition 4.2, inverse elements are unique so we will use g^{-1} to denote the *inverse* of g.

Sanity Check 4.1. Suppose $(G, \diamond)$ is a group. Why is the identity $e \in G$ unique? (*Hint.* Suppose $e_1, e_2 \in G$ satisfy the identity condition for the group. Show that $e_1 = e_2$.)

Proposition 4.2. In any group, the inverse of a given element of the group is unique.

Proof. Let $(G, \diamond)$ be a group with identity e. Suppose $g \in G$ and $h_1, h_2 \in G$ are such that $g \diamond h_1 = h_1 \diamond g = e$ and $g \diamond h_2 = h_2 \diamond g = e$. Observe that

$$\begin{aligned} h_1 &= h_1 \diamond e \\ &= h_1 \diamond (g \diamond h_2) \\ &= (h_1 \diamond g) \diamond h_2 \\ &= e \diamond h_2 \\ &= h_2. \end{aligned}$$

Therefore, $h_1 = h_2$. □

A common practice in algebra is to use juxtaposition for the "multiplicative" operation of the group. In particular, we may say that $(G, \cdot)$ is a group and let $gh = g \cdot h$ for $g, h \in G$.

This convention also inspired "exponent" notation. For a group $(G, \cdot)$, we will consider $g^0 = e$, the identity, and $g^1 = g$. Then, for $n \in \mathbb{N}$, assuming g^n has been defined, we set $g^{n+1} = g^n g$. We also define $g^{-n} = (g^n)^{-1}$.

Exercise 4.2. Let $(G, \cdot)$ be a group. Show that, for $g \in G$, $(g^{-1})^{-1} = g$.

Exercise 4.3. Let $(G, \cdot)$ be a group and $g, h \in G$. Show that $(gh)^{-1} = h^{-1} g^{-1}$.

Exercise 4.4 (Left- and Right-cancellation Laws)**.** Let $(G, \cdot)$ be a group and $a, g, h \in G$. Show that

- $ag = ah \implies g = h$ and
- $ga = ha \implies g = h$.

Exercise 4.5. Let $(G, \cdot)$ be a group and $g, h \in G$. For $n, m \in \mathbb{Z}$, show that

- $g^n g^m = g^{n+m}$ and
- $(g^n)^m = g^{nm}$.

[11] As will be shown in Sanity Check 4.1, there is only one element that satisfies the identity criterion.

Exercise 4.6. Let $(G, \diamond)$ be a group and let $p \in G$. Define $\varphi : G \to G$ by $\varphi(g) = g \diamond p$. Show that φ is a bijection and determine the inverse function of φ.

Example 4.3. The integers $\mathbb{Z}$ with their usual binary operation of addition $+$ forms a group.

Definition 4.4. If a group $(G, \diamond)$ satisfies the property that $g \diamond h = h \diamond g$ for all $g, h \in G$, then G is said to be *abelian* or *commutative*.

Note that the integer group $(\mathbb{Z}, +)$ is abelian.

Exercise 4.7. Suppose $(G, \cdot)$ is an abelian group. Show that, for any $g, h \in G$ and $n \in \mathbb{Z}$, $(gh)^n = g^n h^n$.

Example 4.5. Let $\mathrm{Sym}(\mathbb{R})$ consist of all bijections $\mathbb{R} \to \mathbb{R}$. With the operation of function composition $\circ$, $(\mathrm{Sym}(\mathbb{R}), \circ)$ is a group with identity $\mathrm{id}_{\mathbb{R}}$. By Example 3.23, this group is not commutative.

Example 4.6. Let $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$. Then $(\mathbb{R}^+, \cdot)$, where $\cdot$ is the standard multiplication operation, is an abelian group with identity 1.

Note that Examples 4.5 and 4.6 can justify the overloading of the $\cdot^{-1}$ operator seen in, for example, Calculus courses. In some cases, it is used to refer to the functional inverse, and in other cases, it is used to refer to the multiplicative inverse.

Exercise 4.8. Show that $(\mathbb{R}, \cdot)$, where $\cdot$ is multiplication, is not a group.

Exercise 4.9. Is $(\mathbb{N}, \cdot)$ a group? Why or why not?

Example 4.7. Expanding on Exercise 3.16, let $n \geq 2$, $n \in \mathbb{Z}$, and define $\oplus : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ by the rule

$$[x] \oplus [y] = [x + y].$$

Then $(\mathbb{Z}_n, \oplus)$ is an abelian group with identity $[0]$. As is common, we can identify each $[x]$ with the unique $0 \leq y < n$ for which $x \equiv y \pmod{n}$. In Figure 1, we provide the particular *Cayley table* for each of the groups $\mathbb{Z}_2$, $\mathbb{Z}_3$, and $\mathbb{Z}_5$, with the identifications specified above.

$\oplus$	0	1
0	0	1
1	1	0

$\oplus$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\oplus$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Figure 1: Cayley tables for $(\mathbb{Z}_2, \oplus)$, $(\mathbb{Z}_3, \oplus)$, and $(\mathbb{Z}_5, \oplus)$, respectively.

Example 4.8. For $n \geq 2$, $n \in \mathbb{Z}$, define $* : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ by the rule

$$[x] * [y] = [x \cdot y].$$

In Figure 2, we consider two Cayley tables for this operation in the contexts $n = 5, 6$.

Ignoring the zero rows and columns, we obtain the tables in Figure 3, highlighting zero entries in red. Note that $(\mathbb{Z}_5 \setminus \{0\}, *)$ forms a group but $(\mathbb{Z}_6 \setminus \{0\}, *)$ does not.

In fact, if we let $\mathbb{Z}_n^+ = \mathbb{Z}_n \setminus \{0\}$, then $(\mathbb{Z}_n^+, *)$ is a group if and only if n is prime. Recall that an integer $n \geq 2$ is said to be *composite* if there are integers $2 \leq a, b < n$ such that $ab = n$. If the integer $n \geq 2$ is not composite, it is *prime*.

Recall the standard notation of divisibility: if $n \in \mathbb{Z}$ and $d \in \mathbb{N}$, we write $d \mid n$ if there is an integer k such that $n = dk$.

A fact that we will use here without proof is

Theorem 4.9 (The Fundamental Theorem of Arithmetic)**.** Every integer greater than 1 has a unique prime factorization.

Using The Fundamental Theorem of Arithmentic, one can prove

Lemma 4.10 ("Euclid's Lemma")**.** If a prime p divides a product ab, then either $p \mid a$ or $p \mid b$.

Note that, by mathematical induction, Euclid's Lemma can be extended to the following assertion:

If a prime p divides a product $a_1 a_2 \cdots a_n$, then $p \mid a_j$ for some $j \in \{1, 2, \ldots, n\}$.

The following lemma can be seen as a rephrasing of Euclid's Lemma.

Lemma 4.11. If p is a prime number, then, for integers $0 \leq a, b < p$, if $ab \equiv 0 \pmod{p}$, then either $a = 0$ or $b = 0$.

$*$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

$*$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Figure 2: Cayley tables for $(\mathbb{Z}_5, *)$ and $(\mathbb{Z}_6, *)$, respectively.

$*$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$*$	1	2	3	4	5
1	1	2	3	4	5
2	2	4	0	2	4
3	3	0	3	0	3
4	4	2	0	4	2
5	5	4	3	2	1

Figure 3: Cayley tables for $(\mathbb{Z}_5 \setminus \{0\}, *)$ and $(\mathbb{Z}_6 \setminus \{0\}, *)$, respectively.

Theorem 4.12. Let $p \geq 2$, $p \in \mathbb{Z}$, and consider $\mathbb{Z}_n^+$ with the operation $*$ defined in Example 4.8. Then $(\mathbb{Z}_p^+, *)$ is a group if and only if p is prime.

Proof. First note that, if p is composite, then we can write $p = ab$ for some $2 \leq a, b < p$. Note then that $ab \equiv 0 \pmod{p}$ and so $\mathbb{Z}_p^+$ is not closed under the operation $*$. Consequently, $(\mathbb{Z}_p^+, *)$ is not a group.

Now suppose p is prime. By Lemma 4.11, $\mathbb{Z}_p^+$ is closed under $*$. We also know that 1 satisfies the identity criterion and that multiplication is associative. So the only thing to show is that every element of $\mathbb{Z}_p^+$ has an inverse. So consider $2 \leq n < p$ and the set $\{n^k : k \in \mathbb{N}\}$. Note that $\{n^k : k \in \mathbb{N}\}$ is an infinite subset of natural numbers. However, since $\mathbb{Z}_p^+$ is finite, $\{n^k \pmod{p} : k \in \mathbb{N}\}$ is also finite. Consequently, there must be some $j, k \in \mathbb{N}$, $j < k$, such that $n^j \equiv n^k \pmod{p}$. It follows that

$$p \mid n^j - n^k = n^j(1 - n^{k-j}).$$

Since $p \nmid n$, it must be the case that $p \mid 1 - n^{k-j}$, which is equivalent to $n^{k-j} \equiv 1 \pmod{p}$. Since $2 \leq n < p$, we must have that $k - j > 1$. Hence, $n \cdot n^{k-j-1} \equiv 1 \pmod{p}$. Therefore, n has a multiplicative inverse. □

We are now equipped to prove

Theorem 4.13 (Wilson's Theorem)**.** An integer $p > 1$ is prime if and only if

$$(p-1)! \equiv -1 \pmod{p}$$

where $x!$ is the factorial of x. Equivalently, the integer $p > 1$ is prime if and only if

$$\frac{(p-1)! + 1}{p}$$

is an integer.

Proof. [12] First, suppose p is composite. We proceed here by cases.

First, suppose $p > 4$. Then we write $p = ab$ where $2 \leq a, b < p$. Since $p > 4$, we can assume without loss of generality that $a \geq 3$. Note now that $2 \leq a - 1$ and $1 \leq b - 1$, and, thus,

$$2 \leq (a-1)(b-1) = ab - a - b + 1 \implies a + b \leq ab - 1 = p - 1.$$

[12] This proof is adapted from comments appearing in math.se/307 and math.se/164852.

It follows that

$$(p-1)! = 1 \cdot 2 \cdots a(a+1) \cdots (a+b) \cdots (ab-1).$$

In particular, there is an integer m for which

$$(p-1)! = a(a+1) \cdots (a+b)m.$$

Note that

$$b \mid (a+1)(a+2) \cdots (a+b)$$

since the right-hand expression is the product of b consecutive integers. Hence, there is an integer k for which

$$(a+1)(a+2) \cdots (a+b) = bk.$$

We can thus rewrite our equation above as

$$(p-1)! = abkm = pkm.$$

It follows that $(p-1)! \equiv 0 \pmod{p}$.

In the case that $p = 4$, note that $3! = 6 \equiv 2 \pmod{4}$. Since $2 \not\equiv -1 \pmod{4}$, we have finished this direction of the proof.

Finally, suppose p is prime. In this portion of the proof, we wish to show that the product $(p-1)!$ consists of two types of elemental products, one of which produces a factor of -1 and the other producing a factor of 1. So consider

$$P = \{(n, m) \in \mathbb{Z}_p^+ \times \mathbb{Z}_p^+ : n \leq m \wedge nm \equiv 1 \pmod{p}\}.$$

Since, by Theorem 4.12, $(\mathbb{Z}_p^+, *)$ is a group, we have that, for every $x \in \mathbb{Z}_p^+$ there is some $y \in \mathbb{Z}_p^+$ such that either $(x, y) \in P$ or $(y, x) \in P$.

We now consider

$$H = \{n \in \mathbb{Z}_p^+ : \exists m \in \mathbb{Z}_p^+ \ (n, m) \in P\}.$$

By the uniqueness of inverses, we know that $n \mapsto n^{-1}$, $H \to \mathbb{Z}_p^+$, is an injective function. Also, by the property mentioned above, note that, for any $x \in \mathbb{Z}_p^+ \setminus H$, there is some $y \in H$ such that $x = y^{-1}$.

We now split H into two disjoint subsets:

$$H_{id} = \{n \in H : n = n^{-1}\} \text{ and } H_\star = \{n \in H : n \neq n^{-1}\}.$$

We start by showing that $H_{id} = \{[1], [-1]\}$. So let $n \in H_{id}$ and note that $n = n^{-1}$ is equivalent to $n^2 \equiv 1 \pmod{p}$. It follows that

$$p \mid n^2 - 1 = (n+1)(n-1) \implies p \mid n+1 \vee p \mid n-1.$$

Hence, if $n^2 \equiv 1 \pmod{p}$, then either $n \equiv 1 \pmod{p}$ or $n \equiv -1 \pmod{p}$. This establishes that $H_{id} \subseteq \{[1], [-1]\}$. For equality, simply note that $p-1 \equiv -1 \pmod{p}$ and that $(-1)^2 = 1$.

Finally, enumerate $H_\star = \{a_1, \ldots, a_k\}$ and note that

$$(p-1)! = (p-1)a_1 a_1^{-1} a_2 a_2^{-1} \cdots a_k a_k^{-1} \equiv -1 \pmod{p}.$$

This finishes the proof. □

Example 4.14. Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$. Then define

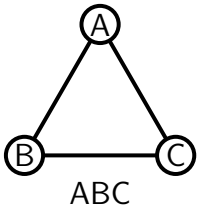
$$\mathrm{GL}(2,\mathbb{F}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc \neq 0 \right\}$$

and

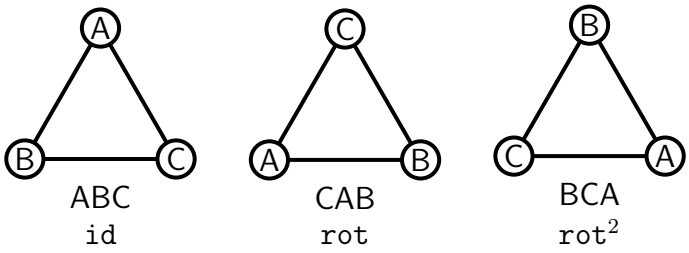
$$\mathrm{SL}(2,\mathbb{F}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc = 1 \right\}.$$

With the operation of matrix multiplication, both $\mathrm{GL}(2,\mathbb{F})$ and $\mathrm{SL}(2,\mathbb{F})$ are nonabelian groups. $\mathrm{GL}(2,\mathbb{F})$ is referred to as the *general linear group* and $\mathrm{SL}(2,\mathbb{F})$ is referred to as the *special linear group*.

Example 4.15. Consider the symmetries of an equilateral triangle with labeled vertices and code each state of the triangle with a word where the first letter corresponds to the top vertex, the second letter corresponds to the bottom left vertex, and the third letter corresponds to the bottom right vertex:

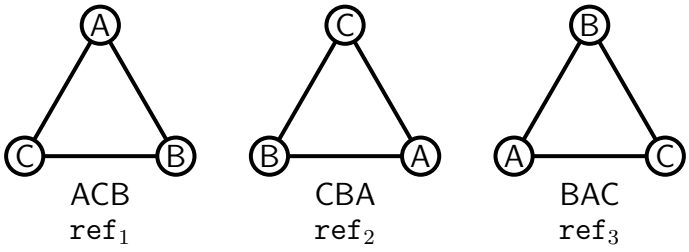


Let `id` represent the null action and let `rot` denote a counterclockwise rotation. We use juxtaposition to indicate actions done in succession, so the for orbit of `rot` looks like this:



Note here[13] that $\texttt{rot}^{-1} = \texttt{rot}^2$ and that $\texttt{rot}^3 = \texttt{id}$.

We also identify three possible reflections:



[13] One clockwise rotation undoes one counter-clockwise rotation.

It is clear that $\mathtt{ref}_j^2 = \mathtt{id}$ for $j = 1, 2, 3$.

Now, for any two action $\mathtt{a}$ and $\mathtt{b}$ listed above, let $\mathtt{a} \cdot \mathtt{b}$ represent the action obtained by first applying $\mathtt{a}$ and then applying $\mathtt{b}$. It can then be verified that

$$\mathtt{rot} \cdot \mathtt{ref}_1 = \mathtt{ref}_2 \qquad \mathtt{rot} \cdot \mathtt{ref}_2 = \mathtt{ref}_3 \qquad \mathtt{rot} \cdot \mathtt{ref}_3 = \mathtt{ref}_1$$

$$\mathtt{rot}^2 \cdot \mathtt{ref}_1 = \mathtt{ref}_3 \qquad \mathtt{rot}^2 \cdot \mathtt{ref}_2 = \mathtt{ref}_1 \qquad \mathtt{rot}^2 \cdot \mathtt{ref}_3 = \mathtt{ref}_2$$

It can be verified directly as above, or by applying Exercise 4.3 to the equations above, that

$$\mathtt{ref}_1 \cdot \mathtt{rot} = \mathtt{ref}_3 \qquad \mathtt{ref}_2 \cdot \mathtt{rot} = \mathtt{ref}_1 \qquad \mathtt{ref}_3 \cdot \mathtt{rot} = \mathtt{ref}_2$$

$$\mathtt{ref}_1 \cdot \mathtt{rot}^2 = \mathtt{ref}_2 \qquad \mathtt{ref}_2 \cdot \mathtt{rot}^2 = \mathtt{ref}_3 \qquad \mathtt{ref}_3 \cdot \mathtt{rot}^2 = \mathtt{ref}_1$$

Let[14]

$$S_3 = \{\mathtt{id}, \mathtt{rot}, \mathtt{rot}^{-1}, \mathtt{ref}_1, \mathtt{ref}_2, \mathtt{ref}_3\}.$$

THen $(S_3, \cdot)$ is a nonabelian group and its corresponding Cayley table is:

$\cdot$	$\mathtt{id}$	$\mathtt{rot}$	$\mathtt{rot}^{-1}$	$\mathtt{ref}_1$	$\mathtt{ref}_2$	$\mathtt{ref}_3$
$\mathtt{id}$	$\mathtt{id}$	$\mathtt{rot}$	$\mathtt{rot}^{-1}$	$\mathtt{ref}_1$	$\mathtt{ref}_2$	$\mathtt{ref}_3$
$\mathtt{rot}$	$\mathtt{rot}$	$\mathtt{rot}^{-1}$	$\mathtt{id}$	$\mathtt{ref}_2$	$\mathtt{ref}_3$	$\mathtt{ref}_1$
$\mathtt{rot}^{-1}$	$\mathtt{rot}^{-1}$	$\mathtt{id}$	$\mathtt{rot}$	$\mathtt{ref}_3$	$\mathtt{ref}_1$	$\mathtt{ref}_2$
$\mathtt{ref}_1$	$\mathtt{ref}_1$	$\mathtt{ref}_3$	$\mathtt{ref}_2$	$\mathtt{id}$	$\mathtt{rot}^{-1}$	$\mathtt{rot}$
$\mathtt{ref}_2$	$\mathtt{ref}_2$	$\mathtt{ref}_1$	$\mathtt{ref}_3$	$\mathtt{rot}$	$\mathtt{id}$	$\mathtt{rot}^{-1}$
$\mathtt{ref}_3$	$\mathtt{ref}_3$	$\mathtt{ref}_2$	$\mathtt{ref}_1$	$\mathtt{rot}^{-1}$	$\mathtt{rot}$	$\mathtt{id}$

Comment. Note that every element of S_3 in Example 4.15 codes a bijection of a given set of three elements. In fact, $(S_3, \cdot)$ is exactly the group of bijections on the set $\{A, B, C\}$ under function composition.

Another thing to note here is that a copy of $(\mathbb{Z}_3, \oplus)$ lives inside of S_3; indeed,

$$(\{\mathtt{id}, \mathtt{rot}, \mathtt{rot}^{-1}\}, \cdot)$$

can be verified to be a copy of $(\mathbb{Z}_3, \oplus)$.

4.2 Subgroups

Definition 4.16. For a group $(G, \cdot)$, we say that $H \subseteq G$ is a *subgroup* of G if $(H, \cdot)$ is a group. The singleton set consisting of the identity element is a subgroup of any given group, and is referred to as the *trivial subgroup*. If H is a proper subset of G and is a subgroup of G, we say that H is a *proper subgroup* of G.

Example 4.17. The integer group $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{R}, +)$.

Exercise 4.10. Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Show that, as defined in Example 4.14, $\mathrm{SL}(2, \mathbb{F})$ is a subgroup of $\mathrm{GL}(2, \mathbb{F})$. (*Hint.* There is an important property about determinants that can be helpful here.)

[14] More on the choice of notation later.