Abstract Algebra

About Author: Dr. Andrew Misseldine is originally from Boise, Idaho. He received his B.S. and M.S. in Mathematics from Boise State University in 2008 and 2010, respectively, and his Ph.D. in Mathematics in 2014 from Brigham Young University. He accepted a Lecturer position at Southern Utah University within the Mathematics Department in Fall 2014. Andrew later became an Assistant Professor in 2016 and received tenure and promotion to Associate Professor in 2022. In 2024, Andrew became Chair of the Department of Mathematics at SUU, where he currently serves.

Andrew has served on several committees within and without the Mathematics Department. He served eight years as a mathematics judge for the Sterling Scholar competition and served four years on the Faculty Senate, with two years on the Faculty Senate Executive Committee. Andrew was a Provost Fellow in 2022. He is currently serving on the CLEP College Mathematics Development Committee for ETS. He publishes scholarly papers regularly in the topics of Algebraic Combinatorics and Representation Theory. He has been a proponent for Open Educational Resources for many years, advocating for them in the classroom, publishing articles about their usage in mathematics education, and even self-publishing a few open mathematical texts himself (such as the very one you are reading now). Andrew hosts a Youtube channel (youtube.com/@Misseldine) about undergraduate mathematics, receiving about 25,000 views per month. For "fun," Andrew teaches middle and high school students cryptography to excite them about mathematics, as the standard K12 curriculum, which is designed to prepare students for Calculus one day, turns many talented students off to mathematics.

Andrew resides in Cedar City, UT, with his wife, Heather, and seven childern. In his free time, Andrew loves playing games (board, card, or video) with his friends and family, playing his cello, reading fantasy novels, watching "Bluey" (with or without his kids), and taking his family to Disneyland. Andrew lived in South Korea during 2004-2006, serving as a missionary for his church. Even all these decades later, he still cherishes the language, the culture, and the food of Korea. This gives context to the many artifacts that decorate his office, including iconography from the Legend of Zelda, Pokémon, Lord of the Rings, South Korea, many family photos, and, of course, mathy things.

Dr. Misseldine loves to hear from his readers. If you have any questions, concerns, or feedback, please contact Dr. Misseldine at andrewmisseldine@suu.edu. You may also visit his website at `website` or the aforementioned YouTube channel for more information about Dr. Misseldine's other textbooks, lecture videos, and scholarly papers.

Abstract Algebra

Andrew Misseldine

November 21, 2025

Contents

Chapter 1

Abstract Prealgebra

"What we call the beginning is often the end. And to make an end is to make a beginning. The end is where we start from." – T. S. Eliot

Lecture Videos

Sets | Subsets | Algebra of Sets | Properties of Set Algebra

1.1 Sets

Definition 1.1.1. A **set** is a well-defined collection of distinct objects. The objects of a set are called its **elements**. By **well-defined**, we mean that there is a rule that enables one to determine whether a given object is an element of the set or not. If a set has no elements, it is called the **empty set** and is denoted $\emptyset$ or $\{\}$.

A set is usually specified either by listing all of its elements inside a pair of braces or by stating the property that determines whether or not an object x belongs to the set. We might write

$$X = \{x_1, x_2, \ldots, x_n\}$$

for a set containing elements $x_1, x_2, \ldots, x_n$ or

$$X = \{x \mid x \text{ satisfies } \mathfrak{p}\}$$

if each x in X satisfies a certain property $\mathfrak{p}$.

Because elements in a set are distinct, we do not let repeats to occur in a set. In other words, the sets $\{1, 3, 2, 2\}$ and $\{1, 3, 2\}$ are the same, that is,

$$\{1, 3, 2, 2\} = \{1, 3, 2\}.$$

Also, the order or arrangement of the elements in a set is irrelevant. So,

$$\{1, 2, 3\} = \{1, 3, 2\} = \{3, 2, 1\}.$$

Definition 1.1.2. When considering if an object is an element of a set or not, the symbol $\in$ means "is an element of" and $\notin$ means "is not an element of."

Example 1.1.3. Let $A = \{1, 2, 3, 4\}$. To denote that 2 is an element of the set A, we write $2 \in A$. To denote that 5 is not an element of A, we write $5 \notin A$.

Some famous sets include:

$$\begin{aligned}
\mathbb{N} &= \{n \mid n \text{ is a natural number}\} = \{0, 1, 2, 3, 4, \ldots\};\\
\mathbb{Z} &= \{n \mid n \text{ is an integer}\} = \{\ldots, -2, -1, 0, 1, 2, 3, 4, \ldots\};\\
\mathbb{Q} &= \{r \mid r \text{ is a rational number}\} = \{p/q \mid p, q \in \mathbb{Z} \text{ where } q \neq 0\};\\
\mathbb{R} &= \{x \mid x \text{ is a real number}\};\\
\mathbb{C} &= \{z \mid z \text{ is a complex number}\}.
\end{aligned}$$

One can compare real numbers by determining whether one is bigger than the other or if they are the same. A similar comparison exists for sets.

Definition 1.1.4. Let A and B be two sets. If A and B have exactly the same elements as each other, then $A = B$.

If every element of A is an element of B, then A is a **subset** of B and we denote this as $A \subseteq B$. For example,

$$\{1, 2\} \subseteq \{1, 2, 3\}.$$

Note that it is always true that $\emptyset \subseteq A$ for any set A. The reason this holds is that there exists no counterexamples, that is, there does not exist an element of $\emptyset$ which is not in A.

If $A \subseteq B$ and $A \neq B$, then $A \subset B$ and A is a **proper subset** of B.

The **power set** of A, denoted $\mathcal{P}(A)$ is the set of all subsets of A. Hence, $X \in \mathcal{P}(A)$ if and only if $X \subseteq A$.

Example 1.1.5. Consider the three sets A = the set of all even numbers, $B = \{2, 4, 6\}$, and $C = \{2, 3, 4, 6\}$.

Here, $B \subseteq A$ since every element of B is also an even number, so is an element of A. Of course, $B \neq A$ since $8 \in A$ but $8 \notin B$. So, $B \subset A$.

It is also true that $B \subset C$. On the other hand, $C \not\subseteq A$ since $3 \in C$ but $3 \notin A$.

Definition 1.1.6. If A and B are sets, then the **intersection** of A and B, denoted $A \cap B$, is the set consisting of elements that belong to both A *and* B. The **union** of A and B, denoted $A \cup B$, is the set consisting of elements that belong to A *or* B (or both).

For a list of sets $A_1, A_2, \ldots, A_n$, then

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n \quad \text{and} \quad \bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n.$$

Two sets are called **disjoint** if their intersection is empty.

Example 1.1.7. Let $A = \{1, 3, 5, 8\}$, $B = \{3, 5, 7\}$, and $C = \{2, 4, 6, 8\}$.

(a) $A \cap B = \boxed{\{3, 5\}}$

(b) $A \cup B = \boxed{\{1, 3, 5, 7, 8\}}$

(c) To calculate $B \cap (A \cup C)$, we first calculate $A \cup C$.

$$A \cup C = \{1, 2, 3, 4, 5, 6, 8\}.$$

Then

$$B \cap (A \cup C) = \boxed{\{3, 5\}}.$$

Usually, we will work with subsets of a bigger set, which we call the **universal set** U.

Definition 1.1.8. If A is a set, the **complement** of A, denoted $\overline{A}$, is the set of all elements in the universal set U that are not in A. Other texts sometimes denote this as A', A^c, or $\sim A$.

Example 1.1.9. If the universal set $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $A = \{1, 2, 5, 7, 9\}$, and $B = \{1, 2, 4\}$, then

$$\overline{A} = \boxed{\{3, 4, 6, 8\}} \qquad \text{and} \qquad \overline{B} = \boxed{\{3, 5, 6, 7, 8, 9\}}.$$

Example 1.1.10. Consider the sets $A = \{\text{red, green, blue}\}$ and $B = \{\text{red, orange, yellow, green, blue, purple}\}$ inside the universal set of all colors. Then

$$B \cap \overline{A} = \boxed{\{\text{orange, yellow, purple}\}}.$$

Note that in the previous example, the set $B \cap \overline{A}$ is the set of all elements of B not contained in A. This is called the **set difference** and is often denoted $B \setminus A$ or $B - A$. In this context, the set difference only depends on B and A and not on the universal set U even though A' depends on U.

Still considering the previous example, note that $A \setminus B = A \cap \overline{B} = \emptyset$ since every element of A is contained in B. In fact, $A \subseteq B$. In general, if $A \subseteq B$ then $A \setminus B = \emptyset$.

Proposition 1.1.11. *For any set $A, B, C \subseteq U$,*

(i) (Idempotency) $A \cup A = A$ and $A \cap A = A$;

(ii) (Identity) $A \cup \emptyset = A$ and $A \cap U = A$;

(iii) (Absorption) $A \cup U = U$, $A \cap \emptyset = \emptyset$, and $A \setminus \emptyset = A$;

(iv) (Complements) $A \cup \overline{A} = U$, $A \cap \overline{A} = \emptyset$, and $A \setminus A = \emptyset$;

(v) (Commutativity) $A \cup B = B \cup A$ and $A \cap B = B \cap A$;

(vi) (Associativity) $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$;

(vii) (Distributivity) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;

(viii) (De Morgan's Laws) $\overline{A \cap B} = \overline{A} \cup \overline{B}$ and $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

Proof. Many of these identities are straight forward to prove and many will be proven in the homework. To illustrate the basic template for proving two sets are equal we prove the first distributive law $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. To prove that two sets are equal, we will show that the two sets are subsets of each other.

We begin by showing that $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$. Let $x \in A \cup (B \cap C)$. So, $x \in A$ or $x \in B \cap C$. Suppose first that $x \in A$. Now, $A \subseteq A \cup B, A \cup C$. So, $x \in A \cup B, A \cup C$. Thus, $x \in (A \cup B) \cap (A \cup C)$. Next, suppose that $x \in B \cap C$. So, $x \in B$ and $x \in C$. Now, $B \subseteq A \cup B$ and $C \subseteq A \cup C$. Thus,

$x \in (A \cup B) \cap (A \cup C)$. Either way, we get $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$.

Next, we show that $(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)$. Let $x \in (A \cup B) \cap (A \cup C)$. So, $x \in A \cup B$ and $x \in A \cup C$. If $x \in A$, then $x \in A \cup (B \cap C)$ and we're done. Suppose then that $x \notin A$. That implies that $x \in B$ and $x \in C$. Thus, $x \in B \cap C$. Therefore, $x \in A \cup (B \cap C)$, which implies that $(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)$.

Combining the two statements, this shows that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. □

[i] See §1.2 Sets and Equivalence Relations in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, suppose that

$$A = \{n \mid n \in \mathbb{N}, n \text{ is even}\}, \quad B = \{n \mid n \in \mathbb{N}, n \text{ is prime}\}, \quad C = \{n \mid n \in \mathbb{N}, 5 \mid n\}.$$

Describe the given set.

1. $A \cap B$
2. $B \cap C$
3. $A \cup B$
4. $A \cap (B \cup C)$

For Exercises 5–19, for any set $A, B, C \subseteq U$, prove the given set identity.
(Pick 2 from Exercises 5-14; Pick 2 from Exercises 15-19)

5. Proposition 1.1.11 (i)
6. Proposition 1.1.11 (ii)
7. Proposition 1.1.11 (iii)
8. Proposition 1.1.11 (iv)
9. Proposition 1.1.11 (v)
10. Proposition 1.1.11 (vi)
11. Proposition 1.1.11 (vii)
12. Proposition 1.1.11 (viii)
13. $A \subseteq B$ iff $A \cap B = A$
14. $A \cup B = (A \cap B) \cup (A \smallsetminus B) \cup (B \smallsetminus A)$
15. $(A \cap B) \smallsetminus B = \emptyset$
16. $(A \cup B) \smallsetminus B = A \smallsetminus B$
17. $A \smallsetminus (B \cup C) = (A \smallsetminus B) \cap (A \smallsetminus C)$
18. $A \cap (B \smallsetminus C) = (A \cap B) \smallsetminus (A \cap C)$
19. $(A \smallsetminus B) \cup (B \smallsetminus A) = (A \cup B) \smallsetminus (A \cap B)$

"The function of education is to teach one to think intensively and to think critically. Intelligence plus character – that is the goal of true education." – Martin Luther King, Jr.

Lecture Videos

Cartesian Products

Function Composition

Injective and Surjective Functions

Properties of Functions

1.2 Functions

Definition 1.2.1. Given sets A and B, we can define a new set $A \times B$, called the **Cartesian product** of A and B, as a set of ordered pairs. That is,

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

We define the Cartesian product of n sets to be

$$\prod_{i=1}^{n} A_i = A_1 \times \ldots \times A_n = \{(a_1, \ldots, a_n) \mid a_i \in A_i \text{ for } i = 1, \ldots, n\}.$$

If $A = A_1 = A_2 = \ldots = A_n$, we often write A^n for $A \times \ldots \times A$ (where A would be written n times). For example, the set $\mathbb{R}^3$ consists of all of 3-tuples of real numbers.

Example 1.2.2. If $A = \{x, y\}$, $B = \{1, 2, 3\}$, and $C = \emptyset$, then $A \times B$ is the set

$$A \times B = \{(x, 1), (x, 2), (x, 3), (y, 1), (y, 2), (y, 3)\} \quad \text{and} \quad A \times C = \emptyset.$$

Definition 1.2.3. A **relation** R between the sets A and B is a subset of their Cartesian product, that is, $R \subseteq A \times B$. For elements $a \in A$ and $b \in B$, we say that a is **related to** b with respect to R, denoted $a\ R\ b$, if and only if $(a, b) \in R$.

We say that a relation $f \subseteq A \times B$ is a **function** or **mapping** from set A to set B, denoted $f : A \to B$ or $A \xrightarrow{f} B$, if each element $a \in A$ occurs in exactly one ordered pair of the form (a, b) in f, that is, each $a \in A$ is related to exactly one $b \in B$. This unique element b is called the **image** of a with respect to f, and this unique relationship is expressed as $f(a) = b$ or $f : a \mapsto b$. The set A is called the **domain** of f, the set B is called the **codomain** of f, the set

$$f(A) = \{f(a) \mid a \in A\}$$

is called the **image** (or **range**) of f, and, any subset $X \subseteq B$, the set

$$f^{-1}(X) = \{a \in A \mid f(a) \in X\}$$

is called the **pre-image** of X with respect to f.

Let B^A denote the set of all functions of the form $f : A \to B$. This notation is used because if $|A| = a$ and $|B| = b$, then $|B^A| = b^a$, where $|S|$ denotes the cardinality of the set S.

Example 1.2.4. In calculus, functions between real numbers are often defined using formulas such as $f(x) = x^3$ or $g(x) = e^x$. Such function expressions are to be understood as meaning the mapping $f : x \mapsto f(x)$ is given by substitution, for example $f(2) = (2)^3 = 8$ and $g(0) = e^0 = 1$.

In this context, the domain and codomain is usually not stated, but the convention is to set them as the maximal subsets of $\mathbb{R}$ such that $f(x) \in \mathbb{R}$ for each $x \in \operatorname{dom} f$. For example, if $f_1(x) = \dfrac{1}{x}$ and $f_2(x) = \sqrt{x}$, then $f_1(0)$ and $f_2(-1)$ do not evaluate to be real numbers. Therefore, the convention is to restrict the sets so that

$$f_1 : (\infty, 0) \cup (0, \infty) \to (\infty, 0) \cup (0, \infty) \quad \text{and} \quad f_2 : [0, \infty) \to [0, \infty).$$

On the other hand, the functions above have the form

$$f : \mathbb{R} \to \mathbb{R} \qquad \text{and} \qquad g : \mathbb{R} \to (0, \infty).$$

Example 1.2.5. Consider the relation $f : \mathbb{Q} \to \mathbb{Z}$ given by $f(p/q) = p$. We know that $\dfrac{1}{2} = \dfrac{2}{4}$, but is $f\left(\dfrac{1}{2}\right) = 1$ or 2? This relation cannot be a mapping because it is not *well-defined*. A relation is **well-defined** if each element in the domain is assigned to a unique element in the range. This is particularly a problem when "mappings" are defined on sets of equivalence classes[i] according to representatives from the classes. If not carefully defined, such "mappings" might not be functions.

Definition 1.2.6. Given two functions $f : B \to C$ and $g : A \to B$[ii] the **composition** of f and g, denoted $f \circ g : A \to C$, is defined by

$$(f \circ g)(x) = f(g(x)).$$

Example 1.2.7. If $f(x) = x^3$ and $g(x) = e^x$, then $(f \circ g)(x) = (e^x)^3 = e^{3x}$.

Definition 1.2.8. Let $f : A \to B$ be a function. Let $a_1, a_2 \in A$. We say that f is **injective** or **one-to-one** if $f(a_1) = f(a_2)$ implies that $a_1 = a_2$, or in other words, for each $b \in f(A)$ there is a unique $a \in A$ such that $f(a) = b$.

We say that f is **surjective** or **onto** if the image of f equals the codomain of f ($f(A) = B$), or in other words, for all $b \in B$ there exists an $a \in A$ such that $f(a) = b$.

If f is both injective and surjective, then we say that f is **bijective**.

Example 1.2.9. If $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$, and $h : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$, $g(x) = e^x$, and $h(x) = x^2$, then f is bijective, g is injective but not surjective, and h is neither injective nor surjective. On the other hand, the mapping $g_1 : \mathbb{R} \to (0, \infty)$ and $h_1 : \mathbb{R} \to [0, \infty)$ given by $g_1(x) = e^x$ and $h_2(x) = x^2$ are both surjective maps but only g_1 is injective. If we consider $h_2(x) : [0, \infty) \to [0, \infty)$ given by $h_2(x) = x^2$, then h_2 is bijective. For these reasons, the convention for defining functions in calculus is very imprecise for higher mathematics.

Theorem 1.2.10. *Let $f : A \to B$, $g : B \to C$, and $h : C \to D$. Then*

(i) The composition of mappings is associative, that is, $(h \circ g) \circ f = h \circ (g \circ f)$;

(ii) If f and g are both surjective, then $g \circ f$ is surjective;

(iii) If f and g are both injective, then $g \circ f$ is injective;

(iv) If f and g are both bijective, then $g \circ f$ is bijective.

Proof. Let $a \in A$. Then

$$[(h \circ g) \circ f](a) = (h \circ g)(f(a)) = h(g(f(a))) = h((g \circ f)(a)) = [h \circ (g \circ f)](a),$$

which proves (i).

Let $c \in C$. Since g is surjective, there exists some $b \in B$ such that $g(b) = c$. Likewise, since f is surjective, there exists some $a \in A$ such that $f(a) = b$. Thus,

$$(g \circ f)(a) = g(f(a)) = g(b) = c,$$

which proves (ii).

Part (iii) is left as an exercise to the student. Part (iv) follows immediately from (ii) and (iii). □

[i] Equivalence relations will be reviewed in the Section 1.4.

[ii] It is not necessary that the codomain of g and domain of f be equal. Instead function composition can be defined when the domain of f is a subset of the codomain of g.

[iii] See §1.2 Sets and Equivalence Relations in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, suppose that

$$A = \{a, b, c\}, \quad B = \{1, 2, 3\}, \quad C = \{x\}, \quad D = \emptyset.$$

List all of the elements in the given set.

1. $A \times B$
2. $B \times A$
3. $A \times B \times C$
4. $A \times D$

For Exercises 5–8, determine if the given relation $f : \mathbb{Q} \to \mathbb{Q}$ is a function. Explain why or why not.

5. $f\left(\frac{p}{q}\right) = \frac{p+1}{p-2}$
6. $f\left(\frac{p}{q}\right) = \frac{3p}{3q}$
7. $f\left(\frac{p}{q}\right) = \frac{p+q}{q^2}$
8. $f\left(\frac{p}{q}\right) = \frac{3p^2}{7q^2} - \frac{p}{q}$

For Exercises 9–12, determine if the given function is injective or surjective.

9. $f : \mathbb{R} \to \mathbb{R} : f(x) = e^x$
10. $f : \mathbb{Z} \to \mathbb{Z} : f(n) = n^2 + 3$
11. $f : \mathbb{R} \to \mathbb{R} : f(x) = \sin x$
12. $f : \mathbb{R} \to \mathbb{R} : f(x) = x^3 - 3x$

For Exercises 13–14, define a function $f : \mathbb{N} \to \mathbb{N}$ with the given properties. Explain why f has the given properties.

13. Injective but not surjective
14. Surjective but not injective

For Exercises 15–19, let $f : A \to B$ and $g : B \to C$ be functions. Prove the given statement.

15. If f and g are both injective functions, then $g \circ f$ is injective.
16. If $g \circ f$ is surjective, then g is surjective.
17. If $g \circ f$ is injective, then f is injective.
18. If $g \circ f$ is injective and f is surjective, then g is injective.
19. If $g \circ f$ is surjective and g is injective, then f is surjective.

"My happiness grows in direct proportion to my acceptance, and in inverse proportion to my expectations."
– Michael J. Fox

Lecture Videos

Invertible Functions

Invertible Functions are Bijective

Permutations

1.3 Inverse Functions and Permutations

Example 1.3.1. Consider the mapping $f : A \to A$, with $A \neq \emptyset$ defined by the rule $f(a) = a$ for all $a \in A$. This mapping is known as the **identity function**, is denoted id_A, and is always bijective.

Definition 1.3.2. Let $f : A \to B$ be a function. We say that f is **invertible** if there exists a function $f^{-1} : B \to A$, called the **inverse** of f, such that

$$f^{-1} \circ f = id_A \qquad \text{and} \qquad f \circ f^{-1} = id_B.$$

When an inverse exists, it is unique. We will see that this is a property true for all groups.

Example 1.3.3. The functions $f(x) = x^3$ and $g(x) = e^x$ are invertible, real-valued functions since $f^{-1}(x) = \sqrt[3]{x}$ and $g^{-1}(x) = \ln x$. Note that $f^{-1} : \mathbb{R} \to \mathbb{R}$ and $g^{-1} : (0, \infty) \to \mathbb{R}$.

On the other hand, the function $h_1 : \mathbb{R} \to [0, \infty)$ given by $h_1(x) = x^2$ is NOT invertible. It is tempting to believe that $h_1^{-1}(x) = \sqrt{x}$, but this is where the domain/codomain conventions of calculus can be ambiguous. Note that although $\sqrt{x}^2 = x$ for all $x \geq 0$, we have $\sqrt{x^2} = |x|$, which does not equal x for $x < 0$. On the other hand, the mapping $h_2 : [0, \infty) \to [0, \infty)$ given by $h_2(x) = x^2$ is invertible with $h_2^{-1}(x) = \sqrt{x}$. It is important to recognize that as functions $h_1 \neq h_2$ even though they are defined by the same mathematical formula.

The reason that h_1 is not invertible in the previous example essentially comes down to the fact that h_1 is not bijective.

Theorem 1.3.4. *A mapping is invertible if and only if it is both one-to-one and onto.*

Proof. Let $f : A \to B$. If f is bijective, then for each $b \in B$ there is a unique $a \in A$ such that $f(a) = b$. Define a mapping $g : B \to A$ such that $g(b) = a$ exactly when $f(a) = b$. By the unique statement mentioned before, g is well-defined. Next, note that $f \circ g = id_B$ and $g \circ f = id_A$. Therefore, f is invertible.

If f is invertible, then $f \circ f^{-1} = id_B$, which is a surjective map. By exercise 1.3.22 (b), f is surjective. Likewise, $f^{-1} \circ f = id_A$, which is an injective map. By exercise 1.3.22 (c), f is injective. Therefore, f is bijective. □

Example 1.3.5. Let A be an $m \times n$ matrix. Then this matrix naturally produces a linear transfor-

mation $T : \mathbb{R}^n \to \mathbb{R}^m$ using the rule

$$T(\boldsymbol{x}) = A\boldsymbol{x}.$$

As was learned in Linear Algebra class, an $n \times n$ matrix A is invertible, it has an inverse A^{-1} if and only if A is nonsingular, that is, $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution. Notice that this statement is a special case of the previous theorem. We note that A being nonsingular is the same thing as T being bijective. Note that $A\boldsymbol{x} = \boldsymbol{b}$ having a solution means that $\boldsymbol{b}$ is the image of some vector $\boldsymbol{x}$ via T, that is, $T(\boldsymbol{x}) = \boldsymbol{b}$. This is surjectivity. Likewise, since $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution means there is exactly one vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \boldsymbol{b}$. Thus, there is only one vector $\boldsymbol{x}$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$. This is injectivity.

Definition 1.3.6. For any set X, a bijective mapping $\pi : X \to X$ is called a **permutation** of X. The set of permutations on X is denoted S_X. When the set has the form $X = \{1, 2, \ldots, n\} \subseteq \mathbb{N}$, we denote S_X as S_n.

When working with permutations on a finite set X, it can be convenient to use a two–row tableaux to express the permutation, where the first row is an arrangement of the elements of X and the second row is the images of the first row with respect to π. Thus, the second row will be a rearrangement of the elements of X, which is why these maps are called permutations. For example, the permutation $\pi : X \to X$ where $X = \{1, \ldots, n\}$ might have the form

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

Theorem 1.3.7. *Let X be a finite, nonempty set. Then $|S_X| = |X|!$.*

Proof. Without the loss of generality, we may assume that $X = \{1, \ldots, n\}$, where $n = |X|$. Let $\pi : X \to X$ be a permutation. Then there are n choices for the assignment $\pi(1)$ since 1 can be mapped to any element of X. Likewise, there are $n - 1$ choices for the assignment $\pi(2)$ since 2 can be mapped to any element of X other than $\pi(1)$. This is to guarantee that π is injective. Likewise, there are $n - 3$ choices for $\pi(3)$, since it can be any element of X except $\pi(1)$ and $\pi(2)$. Continuing in this pattern, there will always be $n - k + 1$ choices for $\pi(k)$. Finally, the product of these choices will give us the correct number of permutation which is $n \cdot (n-1) \cdot (n-2) \cdots 1 = n!$. □

Example 1.3.8. Suppose that $X = \{1, 2, 3\}$. Define a map $\pi : X \to X$ by

$$\pi(1) = 2, \quad \pi(2) = 1, \quad \pi(3) = 3$$

This is a bijective map and hence a permutation. In particular,

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ \pi(1) & \pi(2) & \pi(3) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

There are five other elements of S_3:

$$id_X = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

It is always possible to compose two permutations since they have the same domains as codomains, and this composite will itself be a permutation. A three-row tableaux can be used to compute their composite. The right two rows will be the same as the permutation on the right, but the third row will be the image of this second row by the permutation on the left. For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Permutations always have inverses since they are bijective. They can be easily construct by reflecting the above permutation tableaux (and reordering the elements, if necessary). For example, $\pi^{-1} : S \to S$ is given as

$$\pi^{-1} = \begin{pmatrix} \pi(1) & \pi(2) & \pi(3) \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \pi.$$

Interesting enough, this permutation is equal to its own inverse, although this is not always the case.

[i] See §1.2 Sets and Equivalence Relations in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–5, let $f : X \to Y$ be a function with $A_1, A_2 \subseteq X$ and $B_1, B_2 \subseteq Y$. Note we are not assuming that f is invertibilty, but the pre-image $f^{-1}(B_1) = \{x \in X \mid f(x) \in B_1\}$ is well-defined regardless. Prove the given identity.

1. $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$
2. $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$;
   Given an example where equality fails.
3. $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$
4. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$
5. $f^{-1}(Y \setminus B_1) = X \setminus f^{-1}(B_1)$
6. Let $f : A \to B$ and $g : B \to C$ be invertible functions. Prove that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. This implies that $g \circ f$ is likewise invertible.
7. Define a function on the real numbers by $f(x) = \dfrac{x+1}{x-1}$. Compute $\operatorname{dom} f$, $\operatorname{im} f$, f^{-1}, $f \circ f^{-1}$, and $f^{-1} \circ f$.

"The worst form of inequality is to try to make unequal things equal." – Aristotle

Lecture Videos


Equivalence Relations


Examples of Equivalence Relations


Partitions


Modular Arithmetic as an Equivalence Relations

1.4 Equivalence Classes

Definition 1.4.1. An **equivalence relation** $\sim$ on a set X is a relation $\sim \subseteq X \times X$ such that

(i) (Reflexive property) $x \sim x$ for all $x \in X$;

(ii) (Symmetric property) $x \sim y$ implies $y \sim x$;

(iii) (Transitive property) If $x \sim y$ and $y \sim z$, then $x \sim z$.

Let $x \in X$. Then define the set $[x] = \{y \in X \mid x \sim y\}$. This set $[x]$ is called the **equivalence class** of x and each element of this set is called a **representative** of the equivalence class.

Example 1.4.2. Let $\sim$ be a relation on $X = \{(a, b) \mid a, b \in \mathbb{Z}, b \neq 0\}$ given by

$$(a, b) \sim (c, d) \quad \text{if and only if} \quad ad = bc.$$

This forms an equivalence relation on X.

(i) (Reflexive property) Note that $ab = ab$. So, $(a, b) \sim (a, b)$.

(ii) (Symmetry property) If $(a, b) \sim (c, d)$ then $ad = bc$. This implies that $cb = da$. Thus, $(c, d) \sim (a, b)$.

(iii) (Transitive property) Let $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then $ad = bc$ and $cf = de$. Then $adf = bcf = bde$. Since $d \neq 0$, we can cancel it from both sides and get $af = be$. Thus, $(a, b) \sim (e, f)$.

Therefore, $\sim$ is an equivalence relation. Note that $\mathbb{Q} = \{[x] \mid x \in X\}$, the set of equivalence classes of ordered pairs of integers. The ratio in lowest terms is typically chosen for the representative of these equivalence classes.

Example 1.4.3. Let A, B be $n \times n$ matrices. We say that A is **similar** to B, denoted $A \sim B$, if there is a nonsingular $n \times n$ matrix P such that

$$A = PBP^{-1}.$$

Similarity is an equivalence relation on $n \times n$ matrices.

(i) (Reflexive property) Note that $A = I_n A I_n^{-1}$, where I_n is the $n \times n$ identity matrix. So, $A \sim A$.

(ii) (Symmetry property) If $A \sim B$ then $A = PBP^{-1}$. Then $B = P^{-1}AP = P^{-1}A(P^{-1})^{-1})$, where P^{-1} is an invertible matrix. Thus, $B \sim A$.

(iii) (Transitive property) Let $A \sim B$ and $B \sim C$. Then there exists invertible $n \times n$ matrices P and Q such that $A = PBP^{-1}$ and $B = QCQ^{-1}$. Then $A = PBP^{-1} = P(QCQ^{-1})P^{-1} = (PQ)C(PQ)^{-1}$, where PQ is a nonsingular matrix. Thus, $A \sim C$.

Therefore, $\sim$ is an equivalence relation as claimed. Similar matrices share many linear algebraic properties, for example, they have the same determinants, traces, and eigenvalues. The Jordan Canonical Form is often chosen to represent these similarity classes.

Example 1.4.4. Let $f, g : \mathbb{R} \to \mathbb{R}$ be two differentiable functions. Then say that $f \cong g$ if and only if $f' = g'$. This is an equivalence relation on the set of differentiable real-valued functions. From calculus, we know that $f(x) = g(x) + C$ and when integrating the antiderivative with $C = 0$ is usually chosen to represent this class.

Definition 1.4.5. A **partition** $\mathcal{P}$ of a set X is a collection of nonempty subsets X_1, X_2, ... such that $X_i \cap X_j = \emptyset$ for $i \neq j$ and $\bigcup_k X_k = X$.

Theorem 1.4.6. *Given an equivalence relation $\sim$ on a set X, the equivalence classes of X form a partition of X. Conversely, if $\mathcal{P} = \{X_k\}$ is a partition of a set X, then there is an equivalence relation on X with equivalence classes X_k.*

Proof. Suppose that $\sim$ is an equivalence relation on X. Then each class $[x] \neq \emptyset$ since $x \in [x]$, which follows from the reflexive property. Similarly, $\bigcup_x [x] = X$ since every element $x \in X$ is contained in $[x]$. Finally, if $[x] \cap [y] \neq \emptyset$, then let $z \in [x] \cap [y]$. Then $z \sim x$ and $z \sim y$. By the symmetric property, $x \sim z$. Then $x \sim y$ by the transitive property. This shows that $x \in [y]$, which implies that $[x] \subseteq [y]$. Similarly, we have that $[y] \subseteq [x]$, which implies that $[x] = [y]$. Therefore, the equivalence classes of $\sim$ form a partition of X.

Suppose that $\mathcal{P}$ is a partition of X. We define a relation on X by the rule $x \sim y$ if and only if x and y are contained in the same partition class. We will call $[x]$ the class containing x. Thus, $x \sim y$ means $x \in [y]$. Furthermore, $x \sim y$ if and only if $[x] = [y]$. Therefore, $\sim$ is an equivalence relation on X since $=$ is an equivalence relation on the classes of $\mathcal{P}$. □

Corollary 1.4.7. *Two equivalence classes of an equivalence relation are either disjoint or equal.*

Definition 1.4.8. Let r and s be two integers and suppose that n be a positive integer. We say that r is **congruent** to s **modulo** n, denoted $r \equiv s \pmod{n}$, if and only if $r - s = kn$ for some $k \in \mathbb{Z}$.

Example 1.4.9. Compute the following residues.

(a) $10 \equiv 1 \pmod{3}$ since $10 - 1 = 9 = 3 \cdot 3$.

(b) $15 \equiv 0 \pmod 5$ since $15 - 0 = 15 = 3 \cdot 5$.

(c) $15 \not\equiv 0 \pmod 7$ since $15 = 15 - 0$ is not a multiple of 7.

(d) $30 \equiv 48 \pmod 9$ since $30 - 48 = -18 = -2 \cdot 9$.

(e) $23 \equiv 2 \pmod 7$ since $23 - 2 = 21 = 3 \cdot 7$.

Proposition 1.4.10. *Congruence modulo n is an equivalence relation on $\mathbb{Z}$ for any positive integer n.*

Proof. Let $r, s, t \in \mathbb{Z}$.

(i) (Reflexive property) Note that $r - r = 0 = 0n$. So, $r \equiv r \pmod n$.

(ii) (Symmetry property) If $r \equiv s \pmod n$, then $r - s = kn$ for some $k \in \mathbb{Z}$. Then $s - r = (-k)n$, where $-k \in \mathbb{Z}$. Then $s \equiv r \pmod n$.

(iii) (Transitive property) Let $r \equiv s \pmod n$ and $s \equiv t \pmod n$. Then $r - s = kn$ and $s - t = \ell n$ for some $k, \ell \in \mathbb{Z}$. Adding these together, we get $r - t = (r - s) + (s - t) = kn + \ell n = (k + \ell)n$. Since $k + \ell \in \mathbb{Z}$, we conclude that $r \equiv t \pmod n$, which finishes the proof. □

If we consider the equivalence relation established by the integers modulo 3, then

$$\begin{aligned} [0] &= \{\ldots, -3, 0, 3, 6, \ldots\} \\ [1] &= \{\ldots, -2, 1, 4, 7, \ldots\} \\ [2] &= \{\ldots, -1, 2, 5, 8, \ldots\} \end{aligned}$$

Notice that $[0] \cup [1] \cup [2] = \mathbb{Z}$ and also that the sets are disjoint. The sets [0], [1], and [2] form a partition of the integers. In particular, there will always be n congruence classes modulo n, and each class can be represented by a unique natural number x such that $0 \leq x < n$. In fact, this representative x is just the remainder of the integer when divided by n, a concept to be explored in the next chapter. The set of all congruence classes modulo n is denoted $\mathbb{Z}_n$. This is a set containing n elements.

[i] See §1.2 Sets and Equivalence Relations in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Let $X \neq \emptyset$. Provide an example of a relation on X which is symmetric and transitive but not reflexive.

For Exercises 2–8, determine whether the given relation is an equivalence relation or not. Provide proof or counterexample, as necessary. In the case of an equivalence relation, describe the equivalence classes.

2. For all $x, y \in \mathbb{R}$, $x \sim y$ iff $x \geq y$
3. For all $x, y \in \mathbb{R}$, $x \sim y$ iff $|x - y| \leq 4$
4. For all $m, n \in \mathbb{Z}$, $m \sim n$ iff $mn > 0$
5. For all $m, n \in \mathbb{Z}$, $m \equiv n \pmod 6$
6. For all $(a, b), (c, d) \in \mathbb{R}^2$, $(a, b) \sim (c, d)$ iff $a^2 + b^2 \leq c^2 + d^2$
7. For all $(a, b), (c, d) \in \mathbb{R}^2$, $(a, b) \sim (c, d)$ iff $a^2 + b^2 = c^2 + d^2$
8. For all $(a, b), (c, d) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, $(a, b) \sim (c, d)$ iff $\lambda \in \mathbb{R} \setminus \{0\}$ such that $(a, b) = (\lambda c, \lambda d)$[ii]

[ii] This equivalence relation defines the projective line, which is very important in geometry.

"When an idea reaches critical mass there is no stopping the shift its presence will induce."
– Marianne Williamson

Lecture Videos

Mathematical Induction

Examples of Mathematical Induction

Strong Induction

The Well-Ordering Principle of the Natural Numbers

1.5 Induction and the Well–Ordering Principle

One of the most important proof techniques in algebra is the principle of **mathematical induction**, which is a method for proving statements involving the natural numbers.

Theorem 1.5.1 (First Principle of Mathematical Induction)**.** [i] *Let $S(n)$ be a statement about integers for $n \in \mathbb{N}$ and suppose $S(n_0)$ is true for some integer n_0. If for all integers k with $k \geq n_0$, $S(k)$ implies that $S(k+1)$ is true, then $S(n)$ is true for all integers n greater than or equal to n_0.*

An induction proof typically comes in three phases: a base case, an inductive hypothesis, an application of the inductive hypothesis. The base case is a proof that $S(n_0)$ holds. The inductive hypothesis is the assumption that $S(k)$ holds. Finally, the application of the inductive hypothesis is a proof that the inductive hypothesis implies $S(k+1)$. This is a very versatile proof technique.

Example 1.5.2. Prove the identity $1 + 2 + 3 + \ldots + n = \dfrac{n(n+1)}{2}$.

Proof. We will prove the statement using induction.

(Base Case) We will show the identity holds for $n = 1$. The LHS of the equation will be 1 and the RHS will be $\dfrac{1(1+1)}{2} = \dfrac{2}{2} = 1$. Thus, it holds for the base case.

(Inductive Hypothesis) Suppose that $1 + 2 + 3 + \ldots + k = \dfrac{k(k+1)}{2}$.

(Application of IH) Consider now $1 + 2 + 3 + \ldots + k + (k+1)$. By the inductive hypothesis, we have $1 + 2 + 3 + \ldots + k + (k+1) = \dfrac{k(k+1)}{2} + (k+1)$. Thus,

$$\frac{k(k+1)}{2} + (k+1) = \frac{k(k+1)}{2} + \frac{2(k+1)}{2} = \frac{(k+1)(k+2)}{2} = \frac{(k+1)((k+1)+1)}{2},$$

which proves the identity for $n = k+1$. It follows by induction that the identity holds for all $n \geq 1$. □

An induction proof can be visualized as a string of dominoes. As anyone who has seen a sequence of falling dominoes knows, there are two requirements for a good show: first, someone must push over the first domino (the base case) otherwise nothing happens; second, the dominoes must be close enough together (inductive hypothesis and application) so that one falling knocks the next one over. This domino effect quite accurately describes every induction proof.

Example 1.5.3. Prove every integer $10^{n+1} + 3 \cdot 10^n + 5$ is divisible by 9 for all $n \in \mathbb{N}$.

Proof. We will prove this statement by induction. For our base case, let $n = 0$. Note that

$$10^{0+1} + 3 \cdot 10^0 + 5 = 10 + 3 + 5 = 18 = 9 \cdot 2.$$

For our inductive hypothesis, assume that $10^{k+1} + 3 \cdot 10^k + 5 = 9m$. We will then show that $10^{k+2} + 3 \cdot 10^{k+1} + 5$ is divisible by 9.

$$10^{k+2}+3\cdot 10^{k+1}+5 = 10(10^{k+1}+3\cdot 10^k)+5 = 10(10^{k+1}+3\cdot 10^k+5)+5-10(5) = 90m+45 = 9(10m+5),$$

where the second to last equality follows from the inductive hypothesis. Therefore, by induction, $10^{n+1} + 3 \cdot 10^n + 5$ is divisible by 9 for all natural numbers. □

Example 1.5.4. Prove that $2^n > n$ for all natural numbers n.

Proof. We will prove the statement using induction. Note that $2^0 = 1 > 0$ and $2^1 = 2 > 1$, which shows two base cases of $n = 0, 1$.

Assume that $2^k > k$ for some $k > 1$. Then $2^{k+1} = 2 \cdot 2^k > 2 \cdot k = k + k \geq k + 1$, where the strictly less-than follows from the inductive hypothesis. Therefore, the statement follows from induction. □

Theorem 1.5.5 (Second Principle of Mathematical Induction)**.** *Let $S(n)$ be a statement about integers for $n \in \mathbb{N}$ and suppose $S(n_0)$ is true for some integer n_0. If for all integers k with $k \geq n_0$, $S(n_0), S(n_0 + 1), \ldots, S(k)$ implies that $S(k + 1)$ is true, then $S(n)$ is true for all integers n greater than or equal to n_0.*

This second version of induction is often called **strong induction** and on the surface appears to be a weaker proof technique since more is being assumed in the inductive hypothesis. On the other hand, these two versions of induction are logically equivalent, since the domino effect essentially can turn a weak inductive hypothesis into a strong hypothesis. So there really is no reason to not use the inductive hypothesis, "Assume $S(n)$ is true for all $n < k$."

Definition 1.5.6. Let S be a partially ordered set. We say that S is **well-ordered** if every nonempty subset Z contains a least element.

Theorem 1.5.7 (The Well-Ordering Principle)**.** *The set of natural numbers $\mathbb{N}$ is well-ordered.*

The natural numbers are the canonical example of a well-ordered set. In fact, the well-ordering principle is equivalent to the induction principle. This can be shown by arguing that a statement is true by induction if and only if it is true by a smallest counterexample argument, a proof by contradiction which grabs a smallest counterexample guaranteed by the well-ordering principle and from which a smaller counterexample is constructed. This again means that any set which is well-ordered is $\mathbb{N}$-like. Note that the set $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are ordered set but are not well-ordered since minimal elements need not exist in their subsets.

Example 1.5.8. Let n be a positive integer. Prove that sum of the first n odd natural numbers is n^2, that is,

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

Proof. We will prove the statement by smallest counterexample. Suppose to the contrary that the above equation is false. Let S be the set of all counterexamples to this equation. As a subset of natural numbers, the well-ordering principle implies that there is a smallest counterexample, call it k, that is,

$$1 + 3 + 5 + \ldots + (2k - 1) \neq k^2.$$

We know that $k \neq 1$ since $1 = 1^2$. So $k - 1$ is a natural smaller than k, which means the equation hold for $k - 1$, that is,

$$1 + 3 + 5 + \ldots + [2(k - 1) - 1] = (k - 1)^2.$$

But if we add $2k - 1$ to both sides of the equation we get

$$1 + 3 + 5 + \ldots + 2k - 1 = (k - 1)^2 + 2k - 1 = k^2 - 2k + 1 + 2k - 1 = k^2,$$

which is a contradiction. Therefore, the statement holds for all positive integers. □

[i] We should mention that Induction is not a theorem to be proven. Instead, it is an axiom of the natural numbers, meaning that it is part of the definition of what $\mathbb{N}$ means. We will never need to prove axioms. Instead, in algebra, we show that specific models satisfy axioms and thus take on all the associated theory derived from those axioms. For example, any set that has an induction principle is essentially $\mathbb{N}$-like.

[ii] See §2.1 Mathematical Induction in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–10, prove the given statement.
(Pick 3 from Exercises 1-6; Pick 2 from Exercises 7-10)

1. For $n \in \mathbb{N}$, $\displaystyle\sum_{k=0}^{n} k^2 = 1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$.
2. For $n \in \mathbb{N}$, $\displaystyle\sum_{k=0}^{n} k^3 = 1^3 + 2^3 + \ldots + n^3 = \frac{n^2(n+1)^2}{4}$.
3. For $n \in \mathbb{N}$, $\displaystyle\sum_{k=0}^{n} (3k+1) = 1 + 4 + 7 + \ldots + (3n+1) = \frac{(n+1)(3n+2)}{2}$.
4. For $n \geq 4$, $n! > 2^n$.
5. For $n \in \mathbb{N}$, $3 \mid (10^{n+1} + 10^n + 1)$.
6. For $n \geq 1$, $99 \mid (4 \cdot 10^{2n} + 9 \cdot 10^{2n-1} + 5)$.
7. For $n \in \mathbb{N}$, $\displaystyle\sum_{k=0}^{n} 2^k = 1 + 2 + 2^2 + \ldots + 2^n = 2^{n+1} - 1$.
8. For $n \geq 1$, $\displaystyle\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{1}{2} + \frac{1}{6} + \ldots + \frac{1}{n(n+1)} = \frac{n}{n+1}$.
9. For $n \in \mathbb{N}$ and $x \geq 0$, $(1+x)^n - 1 \geq nx$.
10. For a finite set X, $|\mathcal{P}(X)| = 2^{|X|}$.

"Divide each difficulty into as many parts as is feasible and necessary to resolve it." – Rene Descartes

Lecture Videos

The Division Algorithm

The Euclidean Algorithm

Euclid's Lemma

The Well-Ordering Principle of the Natural Numbers

1.6 Divisibility of Integers

Theorem 1.6.1 (The Division Algorithm)**.** *If $a, b \in \mathbb{Z}$ and if $b > 0$, then there are unique integers q and r such that*

$$a = qb + r$$

where $0 \leq r < b$.

Proof. Let $S = \{a - bk \mid k \in \mathbb{Z}, a - kb \geq 0\}$, which is a subset of the natural numbers. If $a \geq 0$, then $a - b \cdot 0 \in S$. If $a < 0$, then $a - b(2a) = a(1 - 2b) \geq 0$. So, $a - b(2a) \in S$. In either case, S is not empty. By the well-ordering principle, there is a minimal element of S, call it r. Let q be an integer which obtains r, that is, $r = a - bq$.

We claim that $0 \leq r < b$. By definition of S, it must be that $r \geq 0$. Consider the number $a - b(q + 1)$, which is smaller than r. Since r is minimal in S, we can conclude that $a - b(q+1) < 0$ which implies that $a - bq - b = r - b < 0$ or $r < b$. Therefore, there do exist integers q and r such that $a = bq + r$ and $0 \leq r < b$.

To show that these numbers are unique suppose that $a = bq + r$ and $a = bq' + r'$ with $0 \leq r, r' < b$. Without the loss of generality, assume that $r' \geq r$. Removing a from the system of equations, we get $bq + r = bq' + r'$, which implies that $b(q - q') = r' - r \geq 0$. Thus, $b \mid r' - r \leq r' < b$. But the only multiple less than b and greater than or equal to 0 is 0 itself. So, $r' - r = 0$, or $r' = r$. Consequently, $q = q'$ also. Therefore, the numbers q and r are unique. □

While we were able to prove the division algorithm from the well-ordering principle, unfortunately, this argument is a non-constructive proof, that is, although we know q and r exist we do not have any idea what these values are. Fortunately, the long division algorithm from grade school exists (hence the name of the theorem) to help us compute these integers.

Definition 1.6.2. Let a and b be integers. If $b = ak$ for some integer k, we write $a \mid b$. An integer d is called a **common divisor** of a and b if $d \mid a$ and $d \mid b$.

The **greatest common divisor** of integers a and b, denoted $\gcd(a, b)$, is a positive integer d such that d is a common divisor of a and b and if d' is any other common divisor of a and b, then $d' \mid d$. We say that two integers a and b are **relatively prime** if $\gcd(a, b) = 1$.

Theorem 1.6.3 (Greatest Common Divisor Linear Combination)**.** *Let a and b be nonzero integers. Then there exist integers r and s such that*

$$\gcd(a, b) = ar + bs.$$

Furthermore, the greatest common divisor of a and b is unique.

Again the greatest common divisor linear combination theorem can be proven using the well-ordering principle on the set $\{am + bn \mid m, n \in \mathbb{Z}, am + bn > 0\}$, but this proof is again non-constructive. Typically, the Euclidean algorithm is used to construct these values, as is demonstrated in the next example.

Example 1.6.4 (Euclidean Algorithm)**.** Write the greatest common divisor of $a = 945$ and $b = 2415$ as a linear combination of a and b.

The Euclidean algorithm begins by finding $\gcd(a, b)$, which is accomplished by repeated use of the division algorithm:

$$\begin{aligned} 2415 &= 945 \cdot 2 + 525 \\ 945 &= 525 \cdot 1 + 420 \\ 525 &= 420 \cdot 1 + 105 \\ 420 &= 105 \cdot 4 + 0 \end{aligned}$$

Therefore, $\gcd(2415, 945) = 105$.

Next we can use the equations above to build 105 as a linear combination of a and b, working backward. We start at the second to last line.

$$\begin{aligned} 105 &= 525 - 420 \\ &= 525 - (945 - 525) = 2 \cdot 525 - 945 \\ &= 2 \cdot (2415 - 2 \cdot 945) - 945 = 2 \cdot 2415 - 5 \cdot 945 \end{aligned}$$

Therefore, $\boxed{105 = 2a - 5b}$.

Corollary 1.6.5. *Let a and b be two integers that are relatively prime. Then there exist integers r and s such that $ar + bs = 1$.*

Definition 1.6.6. Let p be an integer such that $p > 1$. We say that p is a **prime number** if the only positive numbers that divide p are 1 and p itself. An integer $n > 1$ that is not prime is said to be **composite**.

Lemma 1.6.7 (Euclid)**.** *Let a and b be integers and p be a prime number. If $p \mid ab$, then either $p \mid a$ or $p \mid b$.*

Proof. Suppose that $p \nmid a$. Then $\gcd(p, a) = 1$, which implies there are integers r, s such that $ar + ps = 1$. Multiplying both sides by b, we get

$$b = (ab)r + p(bs).$$

Since p divides the RHS of the equation, we conclude that $p \mid b$. Therefore, $p \mid a$ or $p \mid b$. □

Euclid's lemma is one of the most powerful properties of prime numbers. The following two theorems are consequences of Euclid's lemma.

Theorem 1.6.8 (Euclid)**.** *There exist infinitely many prime numbers.*

Proof. Suppose that to the contrary there are only finitely many primes, $p_1, p_2, \ldots, p_r$. Let $q = p_1 p_2 \ldots p_r + 1$. If q is composite, then it must be divisible by some prime number by Euclid's lemma, say p_i. But this implies that $p_i \mid q - p_1 p_2 \ldots p_r = 1$, a contradiction. Thus, q must be a prime number which is larger than p_i for all i, another contradiction. Therefore, there are infinitely many prime numbers. □

Theorem 1.6.9 (Fundamental Theorem of Arithmetic)**.** *Let n be an integer such that $n > 1$. Then*

$$n = p_1 p_2 \ldots p_r,$$

where $p_1, \ldots, p_r$ are primes (not necessarily distinct). Furthermore, this factorization is unique; that is, if

$$n = q_1 q_2 \ldots q_s,$$

then $r = s$ and the q_i's are just the p_i's rearranged.

[i]See §2.2 The Division Algorithm in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–6, for each pair of integers, a and b, given, compute $\gcd(a, b)$ and find integers r and s such that $\gcd(a, b) = ra + sb$.

1. 14 and 39
2. 234 and 165
3. 1739 and 9923
4. 471 and 562
5. 23,771 abd 19,945
6. -4357 and 3754

For Exercises 7–14, prove the given statement.
(Pick 4 from Exercises 7-14)

7. Let a and b be nonzero integers. If there exist integers r and s such that $ar + bs = 1$, then a and b are relatively prime.
8. Every perfect square is of the form $4k$ or $4k + 1$ for some nonnegative integer k.
9. Let $n \in \mathbb{N}$. Every integer is congruent mod n to precisely one of the integers $0, 1, 2, \ldots, n-1$. Conclude that if r is an integer, then there is exactly one $s \in \mathbb{Z}$ such that $0 \leq s < n$ and $[r] = [s]$. Hence, the integers are indeed partitioned by congruence mod n.

> **Definition 1.6.10.** Let a and b be integers. An integer m is called a **common multiple** of a and b if $a \mid m$ and $b \mid m$.
>
> The **least common multiple** of integers a and b, denoted $\operatorname{lcm}(a, b)$, is a positive integer m such that m is a common multiple of a and b and if m' is any other common multiple of a and b, then $m \mid m'$.

10. For each pair of integers a and b, there exists a unique least common multiple.
11. For all $a, b \in \mathbb{Z}$, $|ab| = \gcd(a, b) \operatorname{lcm}(a, b)$.
12. For all $a, b \in \mathbb{Z}$, $\operatorname{lcm}(a, b) = ab$ if and only if $\gcd(a, b) = 1$.
13. For all $a, b, c \in \mathbb{Z}$, $\gcd(a, c) = \gcd(b, c) = 1$ if and only if $\gcd(ab, c) = 1$.
14. For all $a, b, c \in \mathbb{Z}$, if $\gcd(a, b) = 1$ and $a \mid bc$, then $a \mid c$.

1.7 Supplemenatry Exercises

(Go to Solutions)

1. Find an example of two nonempty sets A and B for which $A \times B = B \times A$ is true.
2. For positive real numbers $a_1, \ldots, a_n$, prove that $\sqrt[n]{a_1 a_2 \cdots a_n} \leq \dfrac{1}{n} \displaystyle\sum_{k=1}^{n} a_k$.
3. For nth differentiable functions f, g, prove the **Leibniz rule**, namely,
$$(fg)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x),$$
where $f^{(k)}$ denotes the kth derivative of f.
4. Prove that the First and Second Principles of Mathematical Induction are logically equivalent.
5. Prove that the First Principle of Mathematical Induction and the Well–Ordering Principle of logically equivalent.
6. Let a and b be integers such that $\gcd(a, b) = 1$. Let r and s be integers such that $ar + bs = 1$. Prove that
$$\gcd(a, s) = \gcd(r, b) = \gcd(r, s) = 1.$$
7. Let $a, b \in \mathbb{N}$ be relatively prime. If ab is a perfect square, prove that a and b must both be perfect squares.
8. Suppose that a, b, r, s are pairwise relatively prime and that
$$\begin{cases} a^2 + b^2 = r^2 \\ a^2 - b^2 = s^2. \end{cases}$$
Prove that a, r, s are odd and b is even.
9. Let $p \geq 2$. Prove that if $2^p - 1$ is prime, then p must also be prime.
10. Prove that there are an infinite number of primes of the form $6n + 5$.
11. Prove that there are an infinite number of primes of the form $4n - 1$.
12. Prove that $\sqrt{2}$ is irrational.

> **Definition.** The **Fibonacci sequence** is the sequence of natural numbers defined recursively as
> $$f_0 = 0, \quad f_1 = 1, \quad f_n = f_{n-1} + f_{n-2}.$$
> For example, the initial terms of the Fibonacci sequence are
> $$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

13. For $n \in \mathbb{N}$, prove $f_n < 2^n$.
14. For $n \geq 2$, prove $f_{n+1} f_{n-1} = f_n^2 + (-1)^n$.
15. For $n \in \mathbb{N}$, prove $f_n = \dfrac{(1 + \sqrt{5})^n - (1 - \sqrt{5})^n}{2^n \sqrt{5}}$.
16. For $n \in \mathbb{N}$, prove that $\gcd(f_n, f_{n+1}) = 1$.

Chapter 2

Group Theory

"Individual commitment to a group effort - that is what makes a team work, a company work, a society work, a civilization work." – Vince Lombardi

Lecture Videos

Binary Operations | Closure Under an Operations | Groups | Solving Equations in Groups

2.1 Groups

In this chapter we introduce the most fundamental and most important structure in abstract algebra, the group.

Definition 2.1.1. A **binary operation** $\circ$ on a set G is a function $\circ : G \times G \to G$. For an element $(a, b) \in G \times G$, the image of (a, b) under $\circ$ is denoted $a \circ b$ (or just juxtaposition ab when the operation is clear from context), that is, $(a, b) \mapsto a \circ b$. We will use the notation $(G, \circ)$ to denote that G is a set and $\circ$ a binary operation on G.

Example 2.1.2. Addition and multiplication are binary operations on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. Addition and multiplication also are binary operations on $\mathbb{Z}_n$, the set of congruence classes modulo n.

Vector addition is a binary operation on $\mathbb{R}^n$. On the other hand, scalar multiplication of vectors is not a binary operation because it is a product of a scalar and a vector producing a vector, $\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. To be a binary operation, the two factors and the product must all be elements of the same set. Likewise, the dot product of two vectors is not a binary operation since it is a function of the form $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, that is, the product is not a vector but a scalar. Conversely, the cross product on $\mathbb{R}^3$ is a binary operation $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ since the product of two vectors is a vector.

Matrix multiplication is a binary operation on the set of $n \times n$ matrices with real scalars, $M_n(\mathbb{R})$, but not on the set of all matrices because the product of two matrices might be undefined if the dimensions are incompatible. This is just a special case of function composition. Recall that B^A is the set of all functions of the form $f : A \to B$. Then while function composition does not form a binary operation for all functions since many composites are undefined, it does form a binary operation on X^X, that is, on functions of the form $f : X \to X$.

Definition 2.1.3. When a binary operation $\circ$ is defined on a set X, we say that a subset $Y \subseteq X$ is **closed** under $\circ$ if the restriction of $\circ$ to Y forms a binary operation on Y, that is, if $a, b \in Y$ then $a \circ b \in Y$.

Example 2.1.4. Subtraction is a binary operation for $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. Note that subtraction is NOT a binary operation for $\mathbb{N}$ since the difference of two natural numbers need not be a natural number, e.g. $3 - 7 = -4 \notin \mathbb{N}$. In other words, $\mathbb{N}$ is not closed under subtraction.

Division is not a binary operation for $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, nor $\mathbb{C}$ since division by zero is undefined. Let $\mathbb{Z}^*$, $\mathbb{Q}^*$, $\mathbb{R}^*$, and $\mathbb{C}^*$ denote the subset of nonzero numbers of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, respectively. Then division is a binary operation on $\mathbb{Q}^*$, $\mathbb{R}^*$, and $\mathbb{C}^*$, but not a binary operation for $\mathbb{Z}^*$ since the quotient of

two integers need not be a nonzero integer, e.g. $1 \div 2 = \frac{1}{2} \notin \mathbb{Z}^*$. Thus, $\mathbb{Z}^*$ is not closed under division.

The set of permutations S_X is closed under composition inside of X^X. In this case, function composition is typically called permutation multiplication.

Definition 2.1.5. We say that $(G, \circ)$ is a **group** if the following three axioms are satisfied by the binary operation:

(iv) (**associativity**) For all $g, h, k \in G$, it holds that

$$g \circ (h \circ k) = (g \circ h) \circ k.$$

(v) (**identity**) There exists an element $e \in G$ such that for all $g \in G$ we have

$$g \circ e = e \circ g = g.$$

(vi) (**inverses**) For all $g \in G$ there is an element $g^{-1} \in G$ such that

$$g \circ g^{-1} = g^{-1} \circ g = e.$$

When the binary operation $\circ$ is clear from context, we will say that G is a group instead of $(G, \circ)$.

Furthermore, we say G is an **Abelian group** if G is a group which satisfies an additional axiom:

(vii) (**commutativity**) For all $g, h \in G$, it holds that

$$g \circ h = h \circ g.$$

For Abelian groups, the operation is often denoted as $+$, the identity as 0, and the inverse of a as $-a$.

Example 2.1.6. The structures $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, and $(\mathbb{C}, +)$ are all Abelian groups, where the identity element is 0 and the inverse of x is just $-x$. The structure $(\mathbb{N}, +)$ is not a group because not all elements have an additive inverse, e.g. $-1 \notin \mathbb{N}$. The set $\mathbb{Z}^+$ of positive integers with addition is also not a group since it has no identity element.

Similarly, $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$, and $(\mathbb{C}^*, \cdot)$ are all Abelian groups, where the identity element is 1 and the inverse of x is just $1/x$, but $(\mathbb{N}^*, \cdot)$ and $(\mathbb{Z}^*, \cdot)$ are not because not all elements have inverses.

The structures $(\mathbb{Z}, -)$, $(\mathbb{Q}, -)$, $(\mathbb{R}, -)$, and $(\mathbb{C}, -)$ are not groups. Although each set has an identity[i] and all elements have inverses, the operation is not associative. Note that $3 - (2 - 1) = 3 - 1 = 2 \neq 0 = 1 - 1 = (3 - 2) - 1$. Of course, the operation of subtraction is noncommutative.

The structure $(X^X, \circ)$ where $\circ$ is just function composition and $(S_X, \circ)$ with permutation multiplication (function composition) are both non-Abelian groups since their binary operation is noncommutative. The group S_X is called the **symmetric group** on X.

Example 2.1.7. The set of congruence classes modulo n, $\mathbb{Z}_n = \{[0], [1], [2], \ldots, [n-1]\}$ is an Abelian group under addition. The identity element of $\mathbb{Z}_n$ is the congruence class of all multiples

of n, namely $[0]$. For a congruence class $[k]$, the inverse is the class $[-k]$. The fact that addition is associative and commutative is an immediate consequence of associativity and commutativity on $(\mathbb{Z}, +)$ and the division algorithm.

When working with $\mathbb{Z}_n$, it is common to identify a class $[k]$ with its unique representative between 0 and n, that is, $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$ and $k + \ell$ is the unique representative of $[k + \ell]$ between 0 and n. With this notation, the identity of $\mathbb{Z}_n$ is 0 and the inverse of k is $n - k$.

Let $\mathbb{Z}_n^* = \{k \in \mathbb{Z}_n \mid \gcd(k, n) = 1\}$ (the book uses the notation $U(n)$ to denote the set of **units** in $\mathbb{Z}_n$). Then $(\mathbb{Z}_n^*, \cdot)$ is also an Abelian group where the identity is 1. The Euclidean algorithm is used to compute inverses in this group because it computes a linear combination $ak + bn = 1$ which implies that $ak = 1 - bn$, that is, $ak \equiv 1 \pmod{n}$. Thus, $a = k^{-1}$.

The importance of the group structure is that groups are exactly the setting where we can solve equations.

Example 2.1.8. Solve the equation $2x + 1 \equiv 5 \pmod{7}$ for x.

To begin we apply the additive inverse of 1 to both sides of the equation:

$$\begin{aligned} (2x+1)+(-1) &\equiv 5+(-1) \pmod{7} \\ 2x+(1+(-1)) &\equiv 4 \pmod{7} \\ 2x+0 &\equiv 4 \pmod{7} \\ 2x &\equiv 4 \pmod{7} \end{aligned}$$

Notice that to "move" 1 to the other side of the equation we used inverses, associativity, and identity.

The Euclidean algorithm (or guess-and-check) can be used to show that $(4)2 + (-1)7 = 1$. Thus, $2^{-1} \equiv 4 \pmod{7}$.

$$\begin{aligned} (4)(2x) &\equiv 4(4) \pmod{7} \\ (4 \cdot 2)x &\equiv 16 \pmod{7} \\ 8x &\equiv 16 \pmod{7} \\ x &\equiv \boxed{2} \pmod{7} \end{aligned}$$

[i] Actually, these algebraic structures only have a **right identity**, that is, an element e such that $g \circ e = g, \forall g \in G$. Similarly, we can define a **left identity** as an element e such that $e \circ g, \forall g \in G$. A left- or right-identity is called a **one-sided identity**. The identity defined in Definition 2.1.5 could more precisely be called a **two-sided identity**. It can be proven that with an associative operation, a one-sided identity is necessarily a two-sided identity is unique. Analogous definitions and statement can be said about one- and two-sided inverses.

[ii] See §3.1 Integer Equivalence Classes and Symmetries and §3.2 Definitions and Examples in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–6, solve the modular congruence for x.

1. $3x \equiv 2 \pmod 7$
2. $5x + 1 \equiv 13 \pmod{23}$
3. $5x + 1 \equiv 13 \pmod{26}$
4. $9x \equiv 3 \pmod 5$
5. $5x \equiv 1 \pmod 6$
6. $3x \equiv 1 \pmod 6$

For Exercises 7–11, prove the given set and operation form a group. Determine, with proof, whether the group is abelian or not.
(Pick 2 from Exercises 7-11)

7. Let $G = \mathbb{R} \setminus \{-1\}$ equipped with operation $x * y = x + y + xy$.
8. Let $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ equipped with standard multiplication.
9. Let $G = \mathbb{R}^* \times \mathbb{Z} =$ equipped with operation $\circ$ defined as
$$(x, a) \circ (y, b) = (xy, a + b).$$
10. Let $\mathbb{Z}_2^n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in \mathbb{Z}_2\}$[iii] equipped with vector addition
$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$
11. Let $H = \left\{ \begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \middle| \, x, y, z \in \mathbb{R} \right\}$[iv] equipped with matrix multiplication
$$\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x' & y' \\ 0 & 1 & z' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x + x' & y + xz' + y' \\ 0 & 1 & z + z' \\ 0 & 0 & 1 \end{pmatrix}.$$

For Exercises 12–20, prove the given properties of modular arithmetic.
(Pick 2 from Exercises 12-20)

12. Addition and multiplication modulo n are well defined operations.
13. For all $a \in \mathbb{Z}_n$, $0 + a \equiv a + 0 \equiv a \pmod n$.
14. For all $a \in \mathbb{Z}_n$, $1a \equiv a1 \equiv a \pmod n$.
15. For all $a \in \mathbb{Z}_n$, there exists $b \in \mathbb{Z}_n$ such that $a + b \equiv b + a \equiv 0 \pmod n$.
16. For all $a, b \in \mathbb{Z}_n$, $a + b \equiv b + a \pmod n$.
17. For all $a, b \in \mathbb{Z}_n$, $ab \equiv ba \pmod n$.
18. For all $a, b, c \in \mathbb{Z}_n$, $a + (b + c) \equiv (a + b) + c \pmod n$.
19. For all $a, b, c \in \mathbb{Z}_n$, $a(bc) \equiv (ab)c \pmod n$.
20. For all $a, b, c \in \mathbb{Z}_n$, $a(b + c) \equiv ab + ac \pmod n$.

[iii] This group is important in algebraic coding theory, which we study later in Chapter 6.
[iv] This group is known as the **Heisenberg group**, an important group in quantum physics.

"There are basically two types of people. People who accomplish things, and people who claim to have accomplished things. The first group is less crowded." – Mark Twain

Lecture Videos

Cayley Tables =

The Quaternion Group

The Dihedral Group

2.2 Cayley Tables

Definition 2.2.1. A **Cayley table** of a group G is a table where rows and columns are marked by elements of a group and the interior of the table in row g and column h is the product gh.

Example 2.2.2. The Cayley table for $(\mathbb{Z}_5, +)$ is given to the right.

$+$	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

The identity in a group can be identified as the unique element with a row or column which matches with the group indicators perfectly. Every other row or column will be a rearrangement of the group elements.

Also, the inverse of g can be found on the Cayley table by finding the column which contains the identity in the row of g. That column corresponds to g^{-1}. For example, in row 2, the identity, 0, is in row 3. Thus, $-2 \equiv 3 \pmod 5$.

Example 2.2.3. The Cayley table for $(\mathbb{Z}_8^*, \cdot)$ is given to the right.

$\cdot$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

Commutativity can be identified from the Cayley table, because the table of an Abelian group will be symmetric across the main diagonal. This is apparent in this and the last example.

In any row or column of a Cayley table, every element of the group appears once and only once. Such a table is called a **latin square**, like in a game of Sudoku.

Example 2.2.4. The set $M_n(\mathbb{R})$ of all $n \times n$ matrices with real scalars with respect to matrix multiplicaiton is not a group, because not every matrix has an inverse. On the other hand, let $\mathrm{GL}_n(\mathbb{R})$ be the set of nonsingular, real, $n \times n$ matrices. This does form a group, known as the **general linear group**. The identity of this group is I_n, the $n \times n$ identity matrix.

Example 2.2.5. There are a variety of ways of building groups out of subsets of the general linear group. For example, let

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$J = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad K = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

be matrices in $\mathrm{GL}_2(\mathbb{C})$. Then the set $Q_8 = \{1, -1, I, -I, J, -J, K, -K\}$ forms a group under matrix multiplication, known as the **quaternion group**. Its Cayley table is displayed.

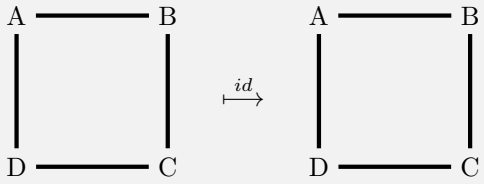
	1	-1	I	$-I$	J	$-J$	K	$-K$
1	1	-1	I	$-I$	J	$-J$	K	$-K$
-1	-1	1	$-I$	I	$-J$	J	$-K$	K
I	I	$-I$	-1	1	K	$-K$	$-J$	J
$-I$	$-I$	I	1	-1	$-K$	K	J	$-J$
J	J	$-J$	$-K$	K	-1	1	I	$-I$
$-J$	$-J$	J	K	$-K$	1	-1	$-I$	I
K	K	$-K$	J	$-J$	$-I$	I	-1	1
$-K$	$-K$	K	$-J$	J	I	$-I$	1	-1

Note that Q_8 is a non-Abelian group since $IJ = K \neq -K = JI$.

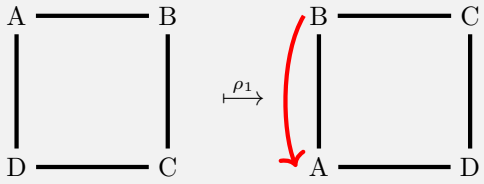
Definition 2.2.6. A **symmetry** of a geometric figure is a rearrangement of the figure preserving the arrangement of its sides and vertices as well as its distances and angles. The set of all symmetries of the regular n-gon is denoted D_n, called the **dihedral group**.

Note that a symmetry of a polygon is just a special type of permutation of the vertices. As such, multiplication of symmetries is function composition and is associative. The identity function is a trivial symmetry, and the inverse of a symmetry is again a symmetry. Symmetry groups are typically noncommutative.

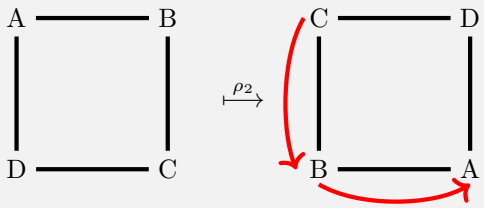
Example 2.2.7. In this example, we will explore D_4, the eight symmetries of a square.



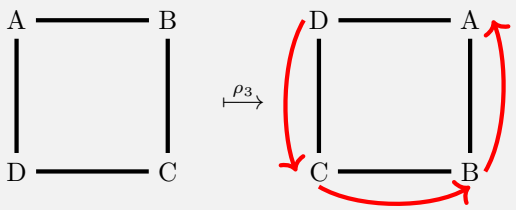
$$id = \begin{pmatrix} A & B & C & D \\ A & B & C & D \end{pmatrix}$$



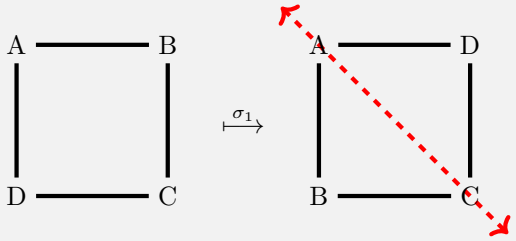
$$\rho_1 = \begin{pmatrix} A & B & C & D \\ D & A & B & C \end{pmatrix}$$



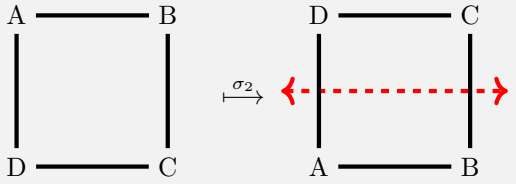
$$\rho_2 = \begin{pmatrix} A & B & C & D \\ C & D & A & B \end{pmatrix}$$



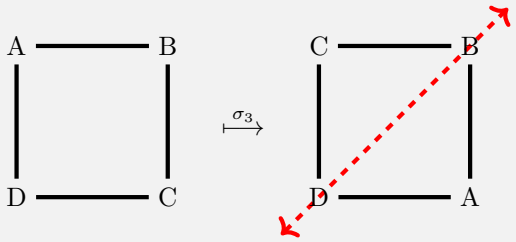
$$\rho_3 = \begin{pmatrix} A & B & C & D \\ B & C & D & A \end{pmatrix}$$



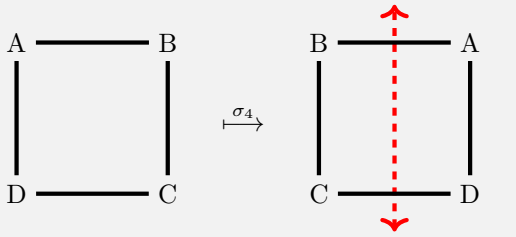
$$\sigma_1 = \begin{pmatrix} A & B & C & D \\ A & D & C & B \end{pmatrix}$$



$$\sigma_2 = \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}$$



$$\sigma_3 = \begin{pmatrix} A & B & C & D \\ C & B & A & D \end{pmatrix}$$



$$\sigma_4 = \begin{pmatrix} A & B & C & D \\ B & A & D & C \end{pmatrix}$$

Multiplication of symmetries is the same permutation multiplication that we introduced earlier. For example,

$$\rho_1\sigma_2 = \begin{pmatrix} A & B & C & D \\ D & A & B & C \end{pmatrix} \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix} = \begin{pmatrix} A & B & C & D \\ D & C & B & A \\ C & B & A & D \end{pmatrix} = \sigma_3$$

and

$$\sigma_4\sigma_1 = \begin{pmatrix} A & B & C & D \\ B & A & D & C \end{pmatrix} \begin{pmatrix} A & B & C & D \\ A & D & C & B \end{pmatrix} = \begin{pmatrix} A & B & C & D \\ A & D & C & B \\ B & C & D & A \end{pmatrix} = \rho_3$$

In this group, id is the identity element, and $\rho_1^{-1} = \rho_3$ while every other element is its own inverse. This group is non-Abelian. The Cayley table for D_4 is provided below.

	id	ρ_1	ρ_2	ρ_3	σ_1	σ_2	σ_3	σ_4
id	id	ρ_1	ρ_2	ρ_3	σ_1	σ_2	σ_3	σ_4
ρ_1	ρ_1	ρ_2	ρ_3	id	σ_2	σ_3	σ_4	σ_1
ρ_2	ρ_2	ρ_3	id	ρ_1	σ_3	σ_4	σ_1	σ_2
ρ_3	ρ_3	id	ρ_1	ρ_2	σ_4	σ_1	σ_2	σ_3
σ_1	σ_1	σ_4	σ_3	σ_2	id	ρ_3	ρ_2	ρ_1
σ_2	σ_2	σ_1	σ_4	σ_3	ρ_1	id	ρ_3	ρ_2
σ_3	σ_3	σ_2	σ_1	σ_4	ρ_2	ρ_1	id	ρ_3
σ_4	σ_4	σ_3	σ_2	σ_1	ρ_3	ρ_2	ρ_1	id

[i]See §3.1 Integer Equivalence Classes and Symmetries and §3.2 Definitions and Examples in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, determine whether the given Cayley table defined over $G = \{a, b, c, d\}$ forms a group.

1.

$\circ$	a	b	c	d
a	a	c	d	a
b	b	b	c	d
c	c	d	a	b
d	d	a	b	c

2.

$\circ$	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

3.

$\circ$	a	b	c	d
a	a	b	c	d
b	b	c	d	a
c	c	d	a	b
d	d	a	b	c

4.

$\circ$	a	b	c	d
a	a	b	c	d
b	b	a	c	d
c	c	b	a	d
d	d	d	b	c

For Exercises 5–8, determine the Cayley table for the given group.

5. The symmetry group of a rectangle (which is not a square).

6. The symmetry group of a rhombus (which is not a square).

7. $(\mathbb{Z}_4, +)$

8. $(\mathbb{Z}_{12}^*, \cdot)$

(Pick 1 from Exercises 9-10)

9. Prove or disprove that every group containing six elements is abelian.

10. Give an example of three different groups with eight elements. Why are the groups different?

"Never doubt that a small group of thoughtful, committed citizens can change the world; indeed, it's the only thing that ever has." – Margaret Mead

Lecture Videos

Order of a Group

Uniqueness of Identities and Inverses

The Cancellation Laws

Exponent Laws

2.3 Properties of Groups

Definition 2.3.1. The **order** of a group $(G, \circ)$ is the cardinality of the set $|G|$.

Example 2.3.2. The group $\mathbb{Z}_n$ under addition has order $|\mathbb{Z}_n| = n$ and the group S_n has order $|S_n| = n!$. These are both examples of **finite groups**. The groups $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ under addition are all examples of groups with infinite order, which we call **infinite groups**.

We will now prove some very important properties about all groups. When working with generic groups multiplicative structure is used almost always. For example, we will denote $g \circ h$ in G simplify as gh and the inverse of g as g^{-1}. Also, we will use exponential notation to represent iterated products: $g^n = \underbrace{g \cdot g \cdots g}_{n \text{ times}}$ and $g^{-n} = \underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{n \text{ times}}$. Although we will primarily use e for the identity of G, we might sometimes use $1 \in G$ as the identity element. In fact, we are justified in saying "the" identity elements because it is unique.

Proposition 2.3.3. *The identity element in a group G is unique; that is, there exists only one element $e \in G$ such that $eg = ge = g$ for all $g \in G$.*

Proof. Let $e', e'' \in G$ such that $ge' = e'g = g$ and $ge'' = e''g = g$ for all $g \in G$. Then

$$e' = e'e'' = e''.$$ □

Likewise, we can say "the" inverse of $g \in G$ since inverses are also unique.

Proposition 2.3.4. *If g is any element in a group G, then the inverse of g is unique.*

Proof. Let $g', g'' \in G$ such that $gg' = g'g = e$ and $gg'' = g''g = e$. Then

$$g' = g'e = g'(gg'') = (g'g)g'' = eg'' = g''.$$ □

Proposition 2.3.5. *Let G be a group. For any $g, h \in G$, then $(gh)^{-1} = h^{-1}g^{-1}$.*

Proof. Since inverses are unique in groups, as shown above, it suffices to show that $h^{-1}g^{-1}$ acts like an inverse to $(gh)^{-1}$. If $h^{-1}g^{-1}$ does act like an inverse to $(gh)^{-1}$ then uniqueness of inverses demand that

$(gh)^{-1} = h^{-1}g^{-1}$. Note that

$$(gh)(h^{-1}g^{-1}) = g(hh^{-1})g^{-1} = g(e)g^{-1} = gg^{-1} = e.$$

Similarly, $(h^{-1}g^{-1})(gh) = e$. Therefore, $h^{-1}g^{-1}$ is the inverse of gh. □

Proposition 2.3.6. *Let G be a group. For any $g \in G$, $(g^{-1})^{-1} = g$.*

Proof. Since inverses are unique in groups, as shown above, it suffices to show that g acts like an inverse to g^{-1}, like the previous proof. Note that

$$gg^{-1} = e = g^{-1}g.$$

Therefore, g is the inverse of g^{-1}. □

Proposition 2.3.7 (Cancellation Laws)**.** *If G is a group and $a, b, c \in G$, then $ab = ac$ implies $b = c$ and $ba = ca$ implies $b = c$.*

Proof. We will prove the case that $ab = ac$ implies that $b = c$. Multiplying both sides of the equation on the left by a^{-1} then gives

$$a^{-1}(ab) = a^{-1}(ac) \quad \Rightarrow \quad (a^{-1}a)b = (a^{-1}a)c \quad \Rightarrow \quad eb = ec \quad \Rightarrow \quad b = c.$$

The case of $ba = ca$ is handled similarly. □

This proposition tells us that the **right** and **left cancellation laws** are true in all groups. In fact, the above proof used all three axioms of group theory: associativity, identity, and inverses, to prove the cancellation laws. In some regard, any algebraic object with cancellation must be group-like.

Proposition 2.3.8. *Let G be a group and $a, b \in G$. Then the equations $ax = b$ and $xa = b$ have unique solutions in G.*

Proof. We will prove the case that $ax = b$ has a unique solution in G. Note that $x = a^{-1}b$ is a solution the equation since

$$a(a^{-1}b) = (aa^{-1})b = eb = b.$$

This shows existence of a solution. For uniqueness, if x and y are both solutions to the above equation, then $ax = b = ay$. Canceling a on the left gives $x = y$. Thus, $ax = b$ has a unique solution in G. The case of $xa = b$ is handled similarly. □

Again, we should emphasize that in the above proof we used all three axioms of group theory to solve equations. Groups are exactly the setting where we solve equations in the manner we first learned in algebra class.

Proposition 2.3.9 (Exponential Laws)**.** *In a group G, the usual exponent laws hold; that is, for all $g, h \in G$ and $m, n \in \mathbb{Z}$,*

(i) $g^m g^n = g^{m+n}$

(ii) $(g^m)^n = g^{mn}$

(iii) $(gh)^n = (h^{-1}g^{-1})^{-n}$. *Furthermore, if G is Abelian, then* $(gh)^n = g^n h^n$.

The proof of the above exponent laws follows from the associativity of the group and an (double) induction argument. It should be emphasized that, in general, $(gh)^n \neq g^n h^n$ in groups and commutativity in some form is needed for equality.

In fact, when a group is Abelian, additive notation is often used instead of multiplicative notation. For example, we will denote $g \circ h$ in G simplify as $g+h$ and the inverse of g as $-g$. Also, we will use multiplicative notation to represent iterated products: $ng = \underbrace{g + g + \ldots + g}_{n \text{ times}}$ and $-ng = \underbrace{-g + -g + \ldots + -g}_{n \text{ times}}$. Although we will primarily use e for the identity of G, we might sometimes use $0 \in G$ as the identity element. The three "exponent laws" look like distributive laws for additive groups: for $g, h \in G$ and $m, n \in \mathbb{Z}$, we have

(i) $mg + ng = (m + n)g$

(ii) $m(ng) = (mn)g$

(iii) $m(g + h) = mg + mh$.

[i] See §3.2 Definitions and Examples in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–9, prove the given statement about a group G.
(Pick 3 from Exercises 1-5; Pick 2 from Exercises 6-9)

1. For any $g_1, g_2, \ldots, g_n \in G$, $(g_1 g_2 \cdots g_n)^{-1} = g_n^{-1} \cdots g_2^{-1} g_1^{-1}$.
2. For any $a, b, c \in G$, if $ba = ca$, then $b = c$ (the remaining part of Proposition 2.3.7).
3. For any $a, b \in G$, the equation $xa = b$ has a unique solution in G (the remaining part of Proposition 2.3.8).
4. Proposition 2.3.9
5. If $a^2 = e$ for all $a \in G$, then G is abelian.
6. If $(ab)^2 = a^2 b^2$ for all $a, b \in G$, then G is abelian.
7. If $ab = a^{-1} b^{-1}$ for all $a, b \in G$, then G is abelian.
8. For $a, b \in G$, if $a^4 b = ba$ and $a^3 = e$, then $ab = ba$.
9. If $|G|$ is even, then there exists some nonidentity $g \in G$ such that $g^2 = e$.

"Someone who hates one group will end up hating everyone – and, ultimately, hating himself or herself."
– Elie Wiesel

Lecture Videos

Subgroups | Not Every Subset is a Subgroup | A Condition for Checking Subgroups | Examples of Subgroups | Another Condition for Checking Subgroups

2.4 Subgroups

Definition 2.4.1. Let $(G, \circ)$ be a group. Let $H \subseteq G$. We say that $(H, \circ)$ is a **subgroup** of G, denoted $H \leq G$, if the restriction $\circ : H \times H \to H$ forms a group structure on $(H, \circ)$.

For any group G, G is itself a subgroup, called the **improper subgroup**. Every other subgroup is called **proper**. Also, the **trivial subgroup** $\{e\}$ is also a subgroup of G since $e \circ e = e$. Note that $\emptyset \not\leq G$ since $e \notin \emptyset$.

Example 2.4.2. With respect to regular addition, we have $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$, and with respect to regular multiplication, we have $\mathbb{Q}^* \leq \mathbb{R}^* \leq \mathbb{C}^*$. These are all infinite, Abelian subgroups.

On the other hand, the set $H = \{\pm 1, \pm i\}$ is a finite subgroup of $\mathbb{C}^*$ of order 4. Similarly, $\{\pm 1\} \leq \mathbb{R}^*$.

The dihedral group D_n is a subgroup of the symmetric group S_n. These are finite, non-Abelian subgroups. The quaternion group Q_8 is a finite, non-abelian subgroup of $\mathrm{GL}_2(\mathbb{C})$.

Example 2.4.3. Although $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\} \subseteq \{0, 1, 2, 3, 4, 5\} = \mathbb{Z}_6$, $\mathbb{Z}_5$ is NOT a subgroup of $\mathbb{Z}_6$, that is $\mathbb{Z}_5 \not\leq \mathbb{Z}_6$. Although $\mathbb{Z}_5$ is a group, this group structure is not formed by restricting the group operation from $\mathbb{Z}_6$, that is, the addition (or multiplication) modulo 5 is NOT formed by restricting addition (or multiplication) modulo 6. On the other hand, $H = \{0, 2, 4\} \leq \mathbb{Z}_6$ since H does form a group structure as illustrated in the Cayley table below. Likewise, $\{0, 3\} \leq \mathbb{Z}_6$. Conversely, $\mathbb{Z}_5$ has NO nontrivial, proper subgroups.

$+ \pmod 6$	0	2	4
0	0	2	4
2	2	4	0
4	4	0	2

Proposition 2.4.4. *Let G be a group. A subset $H \subseteq G$ is a subgroup of G if and only if*

(i) (closure) $gh \in H$ for all $g, h \in H$

(ii) (identity) $e \in H$

(iii) (inverses) $h^{-1} \in H$ for all $h \in H$

Proof. Let $H \leq G$. Since H is itself a group, the product of any two elements in H must also be an element of H. Thus, H is closure under multiplication. Also, H must contain an identity element $e' \in H$ such that $e'h = he' = h$ for all $h \in H$. In particular, this identity holds for e', that is, $e'e' = e'$ in H. Since $e' \in G$,

it has an inverse in G, namely $(e')^{-1}$. Of course, $e = e'(e')^{-1} = (e'e')(e')^{-1} = e'(e'(e'^{-1})) = e'e = e'$. Thus, the identity of H must be the *same identity* as G. Finally, each element h of H has an inverse in H, call it h', such that, $hh' = h'h = e$. But since the identity of H and G are the same, h' is an inverse of h in G. Since inverses in G are unique, we conclude that $h' = h^{-1}$ and $h^{-1} \in H$, that is, H has the *same inverses* as G.

Conversely, suppose the subset H is closed under the operation and contains the identity and inverses. Since the operation is closed, there is a well-defined binary operation on H. Likewise, there is an identity element and each element has a inverse. The last axiom to be a group for H is associativity. Let $g, h, k \in H$. Since $H \subseteq G$, we have that $g, h, k \in G$ and $g(hk) = (gh)k$ in G since it is a group, that is, associativity is *inherited* from G. Therefore, H is a group in its own right. □

In the previous proof, we mentioned that associativity from a group G is inherited by a subset H, because if the property holds for all elements of G then certainly it still remains true for the restriction to H. This is also true for commutativity, that is, every subgroup of an Abelian group is Abelian, because commutativity of the subset is inherited from the ambient commutative group. On the other hand, not every subgroup of a non-Abelian group has to be non-Abelian (why is that?). While properties like associativity and commutativity are inherited by any closed subset, issues about containment, like containing identity or inverse elements, are not guaranteed. Considering that a subset is likely a collection of G with some missing, we should be highly curious whether the elements omitted from G are products, identities, or inverses.

Example 2.4.5. Let $n\mathbb{Z} = \{kn \mid k \in \mathbb{Z}\}$, the set of all integers divisible by n (this is, of course, just the congruence class $[0]$ modulo n). Then $n\mathbb{Z} \leq \mathbb{Z}$.

Proof. To show closure of the operations, let $a, b \in n\mathbb{Z}$. Then there exists $k, \ell \in \mathbb{Z}$ such that $a = kn$ and $b = \ell n$. Then $a + b = kn + \ell n = (k + \ell)n$ where $k + \ell \in \mathbb{Z}$. Therefore, $a + b \in n\mathbb{Z}$, which shows that $n\mathbb{Z}$ is closed under addition. Of course, $0 = 0n \in n\mathbb{Z}$. This shows that $n\mathbb{Z}$ contains the identity. Finally, if $a = kn \in \mathbb{Z}$, then $-a = (-k)n \in \mathbb{Z}$. Therefore, $n\mathbb{Z}$ is closed under inverses. This then proves that $n\mathbb{Z} \leq \mathbb{Z}$ for all $n \in \mathbb{Z}$. □

Example 2.4.6. Let $\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid \det(A) = 1\}$, which is called the **special linear group**. As the name indicates, $\mathrm{SL}_n(\mathbb{R}) \leq \mathrm{GL}_n(\mathbb{R})$.

Proof. For closure, remember that the determinant as the following useful property: $\det(AB) = \det(A)\det(B)$ for all matrices A and B. So, if $A, B \in \mathrm{SL}_n(\mathbb{R})$, then $\det(A) = \det(B) = 1$ and

$$\det(AB) = \det(A)\det(B) = 1 \cdot 1 = 1.$$

Thus, $AB \in \mathrm{SL}_n(\mathbb{R})$ and it is closed under multiplication. Of course, $\det(I_n) = 1$, so $I_n \in \mathrm{SL}_n(\mathbb{R})$. Finally,

$$\det(A^{-1}) = 1 \cdot \det(A^{-1}) = \det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1.$$

Thus, $A^{-1} \in \mathrm{SL}_n(\mathbb{R})$. Therefore, $\mathrm{SL}_n(\mathbb{R})$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$. □

Proposition 2.4.7. *Let H be a subset of a group G. Then H is a subgroup of G if and only if $H \neq \emptyset$ and $gh^{-1} \in H$ for all $g, h \in H$.*

Proof. If H is a subgroup, then $H \neq \emptyset$ since $e \in H$. Likewise, if $g, h \in H$, then $h^{-1} \in H$ and $gh^{-1} \in H$.

Conversely, suppose that $H \neq \emptyset$ and $gh^{-1} \in H$ for all $g, h \in H$. Let $h \in H$, where such an element exists

since $H \neq \emptyset$. Then $e = hh^{-1} \in H$. So, H contains the identity. Likewise, $h^{-1} = eh^{-1} \in H$. Thus, H is closed under inversion. Finally, if $g, h \in H$, then $h^{-1} \in H$ and $gh = g(h^{-1})^{-1} \in H$. Thus, H is closed under multiplication. This shows that H is a subgroup. □

[i] See §3.3 Subgroups in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–6, prove the given subset is a subgroup of the given group.
(Pick 3 from Exercises 1-6)

1. $H = \{2^k \mid k \in \mathbb{Z}\} \subseteq \mathbb{Q}^*$
2. $S^1 = \{z \in \mathbb{C}^* \mid |z| = 1\} \subseteq \mathbb{C}^*\}$
3. $H = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subseteq \mathbb{R}^*$
4. $H = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \middle| \, \theta \in \mathbb{R} \right\} \subseteq \mathrm{SL}_2(\mathbb{R})$
5. $H = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \, a, b, c, d \in \mathbb{R}, a + d = 0 \right\}$ [ii] $\subseteq M_{2\times 2}(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \, a, b, c, d \in \mathbb{R} \right\}$
6. $\mathrm{SL}_2(\mathbb{Z}) \subseteq \mathrm{SL}_2(\mathbb{R})$

For Exercises 7–11, let $H, K \leq G$. Prove the given statement.
(Pick 2 from Exercises 7-11)

7. $H \cap K \leq G$, that is, the intersection of any two subgroups is itself a subgroup.
8. There exists a group G with subgroups H, K such that $H \cup K$ is not a subgroup of G, that is, the union of two subgroups is not necessarily a subgroup.
9. There exists a group G with subgroups H, K such that $HK := \{hk \mid h \in H, k \in K\}$ is not a subgroup of G, that is, the product of two subgroups is not necessarily a subgroup.

Definition 2.4.8. The **center** of a group G, denoted $Z(G)$, is defined as

$$Z(G) := \{z \in G \mid zg = gz \text{ for all } g \in G\}.$$

10. For any group G, $Z(G) \leq G$.

Definition 2.4.9. The **centralizer** of a subgroup H, denoted $C_G(H)$, is defined as

$$C_G(H) := \{z \in G \mid zh = hz \text{ for all } h \in H\}.$$ [iii]

11. For any group G and subgroup H, $C_G(H) \leq G$.

[ii] In this exercise, H is the set of **traceless matrices**. Recall that the **trace** of a square matrix is the sum of its diagonal entries.

[iii] Note that $C_G(G) = Z(G)$.

"A team will always appreciate a great individual if he's willing to sacrifice for the group."
– Kareem Abdul-Jabbar

Lecture Videos

Partially Ordered Sets

Hasse Diagrams

Direct Products

2.5 Hasse Diagrams

Definition 2.5.1. A **partial order** $\preceq$ on a set X is a relation $\preceq \subseteq X \times X$ such that

(i) (Reflexive property) $x \preceq x$ for all $x \in X$;

(ii) (Antisymmetric property) if $x \preceq y$ and $y \preceq x$ then implies $x = y$;

(iii) (Transitive property) If $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

A set X equipped with a partial order $\preceq$ is called a **partially ordered set** (or poset). For elements $x, y \in X$, we say that x and y are **comparable** if $x \preceq y$ (or $y \preceq x$).

Example 2.5.2. We say that $a \leq b$ for $a, b \in \mathbb{R}$ (or $\mathbb{Q}, \mathbb{Z}, \mathbb{N}$) if $b - a$ is nonnegative. This provides the usual ordering of $\mathbb{R}$ which is a partial ordering.

Example 2.5.3. Let X be any set. Then $\mathcal{P}(X)$ (or 2^X) denotes the **power set** of X, which is the set of all subsets of X. Then $(\mathcal{P}(X), \subseteq)$ is a partially ordered set. Let $A, B, C \subseteq X$.

Proof. (i) (Reflexive property) If $x \in A$, then $x \in A$. Therefore, $A \subseteq A$.

(ii) (Antisymmetry property) Let $A \subseteq B$ and $B \subseteq A$. Then each of the two sets contain the same elements and $A = B$.

(iii) (Transitive property) Let $A \subseteq B$ and $B \subseteq C$. Let $x \in A$. Then $x \in B$, which also implies that $x \in C$. Thus, $A \subseteq C$.

Therefore, $\subseteq$ is a partial ordering. □

The reason that this kind of ordering is called *partial* is because not all elements are comparable. For example, let $X = \{1, 2, 3\}$, $Y = \{1\}$, and $Z = \{2\}$. Then $Y, Z \subseteq X$ but $Z \not\subseteq Y$ nor $Y \not\subseteq Z$. Thus, only some of the elements of $\mathcal{P}(X)$ are ordered.

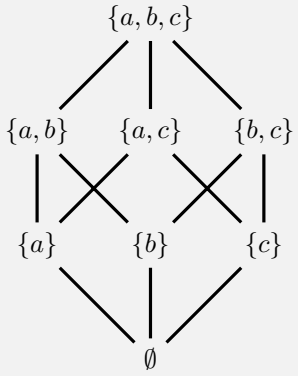
Example 2.5.4. Let G be a group. Then the set of all subgroups of G forms a partially ordered set under $\leq$. The proof of this fact is similar to the previous example.

Example 2.5.5. Let $n \in \mathbb{Z}^+$. Let X be the set of all positive divisors of n. Then $(X, |)$ is a partially ordered set with respect to divisibility. This proof is an exercise for the reader.

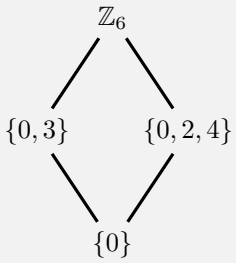
Definition 2.5.6. A **Hasse diagram** (or **lattice diagram**) of a partially ordered set $(X, \preceq)$ is a graph whose vertices are the elements of X and for which an edge (x, y) exists if $x \preceq y$ and whenever $x \preceq z \preceq y$ either $z = x$ or $z = y$.

Example 2.5.7. Let $X = \{a, b, c\}$. Then the Hasse diagram for the partially ordered set $\mathcal{P}(X)$ is given below.

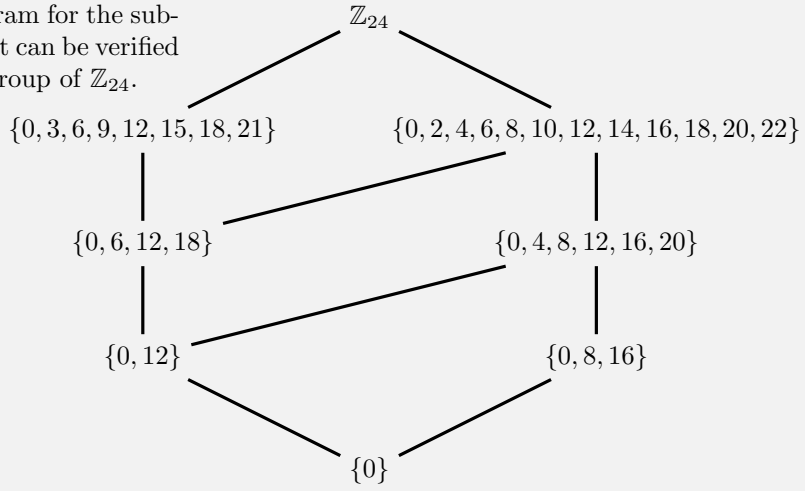
Because of the transitivity of the ordering, any path upward the graph demonstrates comparability of subsets. For example, $\{a\} \subseteq \{a, b, c\}$ since $\{a\} \subseteq \{a, b\} \subseteq \{a, b, c\}$.



Example 2.5.8. The Hasse diagram for the subgroups of $\mathbb{Z}_6$ is displayed below.



Example 2.5.9. The Hasse diagram for the subgroups of $\mathbb{Z}_{24}$ is displayed below. It can be verified that each of the subsets is a subgroup of $\mathbb{Z}_{24}$.



Proposition 2.5.10. *Let $(G, \circ)$ and $(H, *)$ be groups. Then $G \times H$ can be made into a group, called the **direct product** of G and H, using component–wise multiplication, that is,*

$$(g, h)(g', h') = (g \circ g', h * h').$$

Proof. Let $(g,h),(g',h'),(g'',h'') \in G \times H$. Then

$$\begin{aligned}(g,h)[(g',h')(g'',h'')] &= (g,h)(g' \circ g'', h' * h'') = (g \circ (g' \circ g''), h * (h' * h'')) \\ &= ((g \circ g') \circ g'', (h * h') * h'') = (g \circ g', h * h')(g'',h'') = [(g,h)(g',h')](g'',h'').\end{aligned}$$

Therefore, component-wise multiplication is associative. Next,

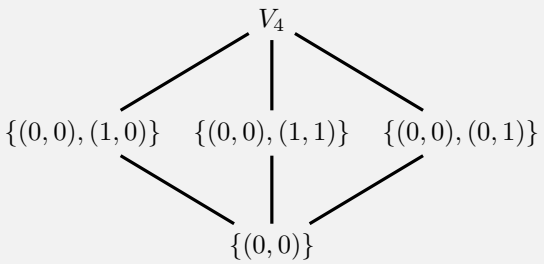
$$(e,e)(g,h) = (e \circ g, e * h) = (g,h) = (g \circ e, h * e) = (g,h)(e,e).$$

Thus, (e,e) is the identity of $G \times H$. Finally,

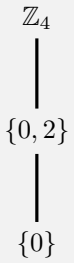
$$(g^{-1},h^{-1})(g,h) = (g^{-1} \circ g, h^{-1} * h) = (e,e) = (g \circ g^{-1}, h * h^{-1}) = (g,h)(g^{-1},h^{-1}).$$

Thus, $(g,h)^{-1} = (g^{-1},h^{-1})$ and $G \times H$ has inverses. Therefore, $G \times H$ is a group. □

Example 2.5.11. Consider the group $V_4 = \mathbb{Z}_2 \times Z_2 = \{(0,0),(1,0),(0,1),(1,1)\}$, called the **Klein 4-group**. The Hasse diagram is displayed below.



Alternatively, the group $\mathbb{Z}_4$ is also an Abelian group of order 4, but the Hasse diagram, displayed below, is very different. This demonstrates that the groups are, in fact, not the same group.



[i] See §3.3 Subgroups and §19.1 Lattices in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Let $n \in \mathbb{Z}^+$. Let X be the set of all positive divisors of n. Prove that $(X, |)$ is a partially ordered set with respect to divisibility.

For Exercises 2–7, draw the Hasse diagram for the given partially ordered set.

2. The subsets of $X = \{a, b, c, d\}$
3. The subgroups of $\mathbb{Z}_{12}$
4. The subgroups of $\mathbb{Z}_3 \times \mathbb{Z}_3$
5. The subgroups of S_3
6. The subgroups of D_4
7. The subgroups of Q_8

2.6 Supplemenatry Exercises

(Go to Solutions)

1. List all elements of $\mathbb{Z}_4 \times \mathbb{Z}_2$.
2. Given an example of two elements A and B in $\mathrm{GL}_2(\mathbb{R})$ such that $AB \neq BA$.
3. Prove that the $\det(AB) = \det(A)\det(B)$ holds for all $A, B \in \mathrm{GL}_2(\mathbb{R})$. Explain from this property that $\mathrm{GL}_2(\mathbb{R})$ is closed under multiplication and hence a group. Likewise, explain from this property that $\mathrm{SL}_2(\mathbb{R}) \leq \mathrm{GL}_2(\mathbb{R})$.
4. Given an example of two elements $g, h \in G$ for some group G and some $n \in \mathbb{N}$ such that $(gh)^n \neq g^n h^n$.
5. Let G be a group and $g, h \in G$. Prove that, for any $n \in \mathbb{Z}$, $(ghg^{-1})^n = gh^n g^{-1}$.
6. If $|G| = 2n$, prove that G must have an element of order 2.
7. Give an example of an infinite group in which every nontrivial subgroup is infinite.
8. Prove that $G \times H$ is abelian if and only if G and H are abelian.
9. Prove or disprove: Every proper subgroup of a nonabelian group is nonabelian.
10. Let G be a group with $H \leq G$. If $g \in G$, let $gHg^{-1} := \{ghg^{-1} \mid h \in H\}$. Prove that $gHg^{-1} \leq G$, called the **conjugate subgroup** of H relative to g.
11. Prove or disprove: The set of all nonzero integers is a poset, where $a \preceq b$ is defined by $a \mid b$.

    Definition 2.6.1. We say a poset $(X, \preceq)$ is a **totally ordered set** if for all $x, y \in X$ either $x \preceq y$ or $y \preceq x$.

12. Is $(\mathbb{N}, |)$ a totally ordered set? Why or why not?
13. Prove that $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ are totally ordered sets under the usual ordering $\leq$.

Chapter 3

Cyclic Groups

"Without forgiveness life is governed by... an endless cycle of resentment and retaliation."
– Roberto Assagioli

Lecture Videos


Cyclic Subgroups


Examples of Cyclic Subgroups


Cyclic Groups


Every Subgroup of a Cyclic Group is Cyclic


Order of an Element

3.1 Cyclic Groups

Often a subgroup will depend entirely on a single element of the group; that is, knowing that particular element will allow us to compute any other element in the subgroup.

Theorem 3.1.1. *Let G be a group and let $g \in G$. Then the set*

$$\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\}^{\text{i}}$$

*is an abelian subgroup of G, called the **cyclic subgroup** generated by g. In fact, this subgroup is the smallest subgroup containing g.*

Proof. Let $H = \langle g \rangle$. We will first show that H is an abelian subgroup of G. Let $h, k \in H$. Then there exists $n, m \in \mathbb{Z}$ such that $h = g^n$ and $k = g^m$. Then

$$hk = (g^n)(g^m) = g^{n+m} \in H.$$

Thus, H is closed under multiplication. Furthermore,

$$hk = g^{n+m} = g^{m+n} = (g^m)(g^n) = kh.$$

Thus, multiplication in H is commutative. Note that $1 = g^0 \in H$. Thus, H contains the identity. Likewise, if $h \in H$, then $h = g^n$ for some $n \in \mathbb{Z}$. Now, $g^{-n} \in H$ and $g^n g^{-n} = g^{n-n} = g^0 = 1$. Thus, H is closed under inversion. This shows that H is an abelian subgroup of G.

To show that H is the smallest subgroup of G containing g. This means, of course, that if $K \leq G$ and $g \in K$ then $H \subseteq K$ (which also implies that $H \leq K$). This is also the same thing as saying H is the intersection of all subgroups of G containing g. We will prove this by induction.

Let $K \leq G$ with $g \in K$. Since K is a subgroup of G, it contains $1 = g^0$. This serves as a base case. For our inductive hypothesis, suppose that $g^n \in K$ for some $n \in \mathbb{Z}$. Since K is a subgroup, it contains inverses such as $g^{-1} \in K$. Likewise, it must also be that $g^{-n} = (g^n)^{-1} \in K$. Next, since K is closed under multiplication, we have that

$$g^{n+1} = gg^n \in K \quad \text{and} \quad g^{-(n+1)} = g^{-n-1} = g^{-1}g^{-n} \in K.$$

Therefore, it follows by induction that $g^n \in K$ for all $n \in \mathbb{Z}$. Thus, $H \subseteq K$. □

Example 3.1.2. Consider $\mathbb{Q}$ under addition. Let H be the set of all rational numbers whose denominators, when in lowest terms, divide 6. Then H is a subgroup of $\mathbb{Q}$ since $H = \left\langle \frac{1}{6} \right\rangle$. In particular, we could define H similarly for any nonzero integer n and this would form the cyclic

subgroup $\left\langle \frac{1}{n} \right\rangle$.

Consider now $\mathbb{Q}^*$ under multiplication. Let H be the set of all powers of 2, that is, $\{2^n \mid n \in \mathbb{Z}\} = \left\{\ldots, \frac{1}{2}, 1, 2, 4, 8, \ldots\right\}$. This is, of course, the cyclic subgroup $\langle 2\rangle \leq \mathbb{Q}^*$. Under addition, the cyclic subgroup generated by 2 is instead $\langle 2\rangle = \{2n \mid n \in \mathbb{Z}\} = \{\ldots, -2, 0, 2, 4, 6, \ldots\} \leq \mathbb{Q}$. Thus, the cyclic subgroup of an element depends entirely on the binary operation being considered.

Definition 3.1.3. Let G be a group. We say that G is a **cyclic group** if there exists some element $g \in G$ such that $G = \langle g\rangle$. We call such an element a **generator** of G.

Example 3.1.4. Consider $\mathbb{Z}$ under addition. Then

$$n\mathbb{Z} = \langle n\rangle = \{kn \mid k \in \mathbb{Z}\}$$

is a cyclic subgroup of $\mathbb{Z}$ for all integers n. For example, $2\mathbb{Z} = \{\ldots, -2, 0, 2, 4, 6, \ldots\}$. In fact, it can be shown that every subgroup of $\mathbb{Z}$ is cyclic, even the improper subgroup. Note that $\mathbb{Z} = 1\mathbb{Z} = \langle 1\rangle$. Thus, $\mathbb{Z}$ is an infinite cyclic group. $\mathbb{Z}$ has two generators, 1 and -1. On the other hand, $\mathbb{Q}$ is not a cyclic group, as can be shown.

Consider $\mathbb{Z}_n$ under addition. Then

$$m\mathbb{Z}_n = \langle m\rangle = \{km \mid k \in \mathbb{Z}_n\}$$

is a cyclic subgroup of $\mathbb{Z}_n$ for all congruence classes m. For example, $2\mathbb{Z}_6 = \{0, 2, 4\}$ In fact, it can also be shown that every subgroup of $\mathbb{Z}_n$ is cyclic, even the improper subgroup. Note that $\mathbb{Z}_n = 1\mathbb{Z}_n = \langle 1\rangle$. Thus, $\mathbb{Z}_n$ is a finite cyclic group for all $n \in \mathbb{Z}$. Now, 1 and $n-1$ are always generators, but there can be more. For example, $\mathbb{Z}_5 = \langle 2\rangle$. In particular, the generators are all the elements of $\mathbb{Z}_n$ which are coprime p to n. This follows from the fact that $1 \in \langle p\rangle$, a consequence of the Euclidean algorithm since $1 = ap + bn \equiv ap \pmod{n}$.

Theorem 3.1.5. *Every subgroup of a cyclic group is cyclic.*

Proof. Let $H \leq G = \langle g\rangle$. Thus, every element of H is a power of g. By the Well-Ordering Principle, there is a least positive power of g in H, call it $h = g^k$. Clearly $\langle h\rangle \subseteq H$, since H is a subgroup containing h. We must show that $H \subseteq \langle h\rangle$. In this regard, let $g^n \in H$. Then by the division algorithm, we get $n = qk + r$ where $r = 0$ or $r < k$. Since k was chosen minimally, we have $r = 0$. Thus, $n = qk$. Then $g^n = g^{qk} = (g^k)^q = h^q \in \langle h\rangle$. Therefore, $H \subseteq \langle h\rangle$, which finishes the proof that $H = \langle h\rangle$. □

Corollary 3.1.6. *The subgroups of $\mathbb{Z}$ are exactly $n\mathbb{Z}$ and the subgroups of $\mathbb{Z}_n$ are exactly $m\mathbb{Z}_n$.*

Definition 3.1.7. Let G be a group and let $g \in G$. We say the **order** of g, denoted $|g|$, is the smallest positive integer n such that $g^n = e$. If no such integer exists, we say that g has infinite order.

Example 3.1.8. In the group $\mathbb{Z}_6$, we have $|0| = 1$, $|1| = 6$, $|2| = 3$, $|3| = 2|$, $|4| = 3$, and $|5| = 6$. In $\mathbb{Z}_6^*$, we have $|1| = 1$ and $|5| = 2$.

[i] For additive notation, a cyclic subgroup looks like $\langle g \rangle = \{ng \mid n \in \mathbb{Z}\}$.

[ii] See §4.1 Cyclic Subgroups in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–6, find the order of the given element in its respective group.

1. $5 \in \mathbb{Z}_{12}$
2. $\sqrt{3} \in \mathbb{R}$
3. $\sqrt{3} \in \mathbb{R}^*$
4. $-i \in \mathbb{C}^*$
5. $72 \in \mathbb{Z}_{240}$
6. $312 \in \mathbb{Z}_{471}$

For Exercises 7–10, determine the orders of each element of the given subset in the given group.
(Pick 2 from Exercises 7-10)

7. $\mathbb{Z}_{18} \subseteq \mathbb{Z}_{18}$
8. $D_4 \subseteq D_4$
9. $\mathbb{Z}_{30}^* \subseteq \mathbb{Z}_{30}^*$
10. $\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{3} \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \right\} \subseteq \mathrm{GL}_2(\mathbb{R})$

For Exercises 11–19, prove the given properties of cyclic groups or orders.
(Do Exercise 11; Pick 2 from Exercises 12-19)

11. If G has no proper, nontrivial subgroups, then G is a cyclic group.
12. If G is a group with $g \in G$, then $|g| = |g^{-1}|$.
13. If G is a group with $g, x \in G$, then $|x| = |gxg^{-1}|$.
14. If G is a group with $g, h \in G$, then $|gh| = |hg|$.
15. The cyclic group $\mathbb{Z}_p$ has no nontrivial, proper subgroups if p is prime.
16. Let G is a group and $g, h \in G$. If $|g| = n$, $|h| = m$, and $\gcd(g, h) = 1$, then $\langle g \rangle \cap \langle h \rangle = \{e\}$.
17. If G be an abelian group, then the elements of G of finite order form a subgroup, called the **torsion subgroup**, denoted $T(G)$.
18. If G is an abelian group with a pair of cyclic subgroups of order 2, then G must contain a noncyclic subgroup of order 4.
19. The order of an element in a finite cyclic group G divides $|G|$.

"I always like to look on the optimistic side of life, but I am realistic enough to know that life is a complex matter." – Walt Disney

Lecture Videos

Complex Numbers

Roots of Unity

Complex Representations of Cyclic Groups

3.2 Roots of Unity

Definition 3.2.1. Let $i = \sqrt{-1}$. Then we say z is **complex number** if it is of the form

$$z = a + bi$$

where $i = \sqrt{-1}$ and a and b are real numbers. The real number a is called the **real part** of z; the real number b is called the **imaginary part** of z.

When adding or subtracting complex numbers, follow the simple rule of "combining like-terms", that is, combine the real parts together and combine the imaginary parts together. Multiplication of complex numbers follows the "FOIL method" with $i^2 = -1$. Division of complex numbers is accomplished by "rationalizing the denominator" using the complex conjugate $\overline{a + bi} = a - bi$. Readers who need a refresher on the arithmetic and geometry of complex numbers should review Appendix A.

When considering complex numbers, it is important to remember $i^2 = -1$. Likewise, $i^3 = i^2 i = -i$ and $i^4 = i(-i) = (-1)(-1) = 1$. Higher powers of i can similarly be reduced. In particular, $i^n = i^{n \bmod 4}$. Also, $i(-i) = 1$.

Example 3.2.2.

(a) $i^{27} = i^{24+3} = i^{24} i^3 = i^{4(6)} i^3 = (i^4)^6 i^3 = (1)^6(-i) = \boxed{-i}$

(b) $i^{101} = i^{100} i = (i^4)^{25} i = 1^{25} i = \boxed{i}$

This anomaly of powers of i can be easily explained in the language of group theory. If $H = \{1, -1, i, -i\}$, then H is a cyclic subgroup of $\mathbb{C}^*$. Although this is certainly the most famous cyclic subgroup of $\mathbb{C}^*$, it is certainly not the only one.

Definition 3.2.3. Let $\zeta \in \mathbb{C}^*$. Let n be a positive integer. We say that ζ is an **nth root of unity** if $\zeta^n = 1$. We say that ζ is a **primitive nth root of unity** if n is the smallest positive integer such that $\zeta^n = 1$.

The set $\{1, -1, i, -i\}$ is the set of 4th roots of unity, that is, the four distinct 4th roots of one. The elements i and $-i$ are the two primitive 4th roots of unity. Likewise, the set $\{1, -1\}$ is the set of square roots (2nd roots) of unity, with -1 being primitive. The set $\{1\}$ is the set of first roots of unity, where 1 is primitive.

We may identify the complex numbers with points on the plane, that is, we will identify the complex number $z = a + bi$ with the point (a, b) or vector $\begin{bmatrix} a \\ b \end{bmatrix}$, called the **graph** of the complex number. When

graphing complex numbers, the x-axis becomes the **real axis** and the y-axis becomes the **imaginary axis**. The resulting plane is called the **complex plane**. This is the Cartesian representation of the complex number.

In polar coordinates, the radius r of the complex point is called the **absolute value** or **modulus**, denoted as $|z| = |x + yi| = \sqrt{x^2 + y^2}$. The angle θ of the complex number is called the **argument**, denoted $\arg(z) = \tan^{-1}\left(\frac{y}{x}\right) = \sin^{-1}\left(\frac{y}{|z|}\right) = \cos^{-1}\left(\frac{x}{|z|}\right)$.

Using basic trigonometry, we know that the components of $\boldsymbol{z} = \begin{bmatrix} a \\ b \end{bmatrix}$ are $x = r\cos\theta$ and $y = r\sin\theta$ where r and θ are the magnitude and direction of $\boldsymbol{z}$, respectively. Therefore, for the complex number $z = x + yi$, if $r = |z|$ and $\theta = \arg(z)$, then

$$z = x + yi = (r\cos\theta) + (r\sin\theta)i = r\cos\theta + ri\sin\theta = r(\cos\theta + i\sin\theta)$$

Does it make sense to consider a quantity with complex exponent, such as 2^{1-i} or e^i? Let $a > 0$ be a real number and let $z = x + yi$ be a complex number. Then by exponent rules, we have

$$a^z = a^{x+yi} = a^x a^{yi} = a^x(a^{yi}).$$

So, to compute complex exponents, we simplify need to understand imaginary powers. Furthermore,

$$a^{iy} = (e^{\ln a})^{iy} = e^{iy\ln a} = (e^{iy})^{\ln a}.$$

Thus, it suffices to explain e^{iy}. Although this might seem like a bizarre quantity, using power series, it is quite reasonable. It corresponds to the polar form of complex numbers!

Theorem 3.2.4 (Euler's Formula)**.** *Let $z = x + yi$, let $r = |z|$, and $\theta = \arg(z)$. Then*

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta}.$$

In particular,

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Example 3.2.5. Evaluate the following complex exponents.

(a) $e^{i\pi}$

When $\theta = \pi$, we get

$$e^{\pi i} = \cos\pi + i\sin\pi = \boxed{-1}$$

The previous identity shows that the three titans of mathematical constants: e, π, and i, are all related to each other in a subtle but beautiful way.[i]

(b) $e^{-1+i\pi/2}$

$$e^{-1+i\pi/2} = e^{-1}e^{i\pi/2} = \frac{1}{e}\left(\cos\left(\frac{\pi}{2}\right) + i\sin\left(\frac{\pi}{2}\right)\right) = \frac{1}{e}(0 + i) = \boxed{\frac{i}{e}}.$$

(c) $z = 2e^{\pi i/3}$

$$z = 2e^{\pi i/3} = 2\left(\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)\right) = 2\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = \boxed{1 + i\sqrt{3}}.$$

Corollary 3.2.6. *Let $r(\cos\theta + i\sin(\theta))$ and $w = s(\cos\varphi + i\sin\varphi)$. Then*

$$zw = rs(\cos(\theta+\varphi) + i\sin(\theta+\varphi)).$$

In particular, $|zw| = |z| \cdot |w|$.

Proof. This follows immediately from the usual exponent laws and Euler's identity, since $z = re^{i\theta}$ and $w = se^{i\varphi}$. □

Corollary 3.2.7. *Let $\zeta_n = e^{2\pi i/n} = \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)$. Then ζ_n is a primitive nth root of unity. In fact, ζ_n^k is a primitive root of unity for each integer k coprime to n.*

Proof. This is a special consequence of Euler's Identity called DeMoivre's identity:

$$z^n = r^n(\cos(n\theta) + i\sin(n\theta).$$

Since sine and cosine are 2π-periodic, the result follows. □

Theorem 3.2.8. *Let Z_n denote the set of all nth roots of unity in $\mathbb{C}^*$. Then Z_n is a cyclic subgroup of $\mathbb{C}^*$ generated by any primitive nth root of unity.*

Proof. Let $\zeta \in \mathbb{C}^*$ be a primitive nth root of unity. Let $G = \langle\zeta\rangle$. By exponent rules, any power of ζ is an nth root of unity since $(\zeta^m)^n = (\zeta^n)^m = 1^m = 1$. Thus, $G \subseteq Z_n$. Since ζ is primitive, $|\zeta| = n$. Thus, $|G| = n$. But each nth root of unity is a root to the polynomial $X^n - 1$, which has n complex roots by the Fundamental Theorem of Algebra. Therefore, $n = |G| \leq |Z_n| \leq n$. Thus, $|Z_n| = n$ and $G = Z_n$. □

Theorem 3.2.9. *Let $S^1 = \{z \in \mathbb{C}^* \mid |z| = 1\}$. Then $S^1 \leq \mathbb{C}^*$ which contains Z_n for each n.*

Proof. Let $\mu, \nu \in S^1$. Then $|\mu\nu| = |\mu| \cdot |\nu| = 1 \cdot 1 = 1$. Thus, $\mu\nu \in S^1$. Of course, $|1| = 1$, which implies that $1 \in S^1$. Since $\mu^{-1}\mu = 1$, we have $|\mu^{-1}| = |\mu^{-1}| \cdot |\mu| = |\mu^{-1}\mu| = |1| = 1$. Thus, $\mu^{-1} \in S^1$. Therefore $S^1 \leq \mathbb{C}^*$. If $\zeta = e^{2\pi i/n}$, then $|\zeta| = \cos^2(2\pi/n) + \sin^2(2\pi/n) = 1$. Thus, $Z_n \subseteq S^1$ for each n. □

The group S^1 is often called the **circle group**, since the set of all modulus 1 complex numbers forms the unit circle in the complex plane. Similarly, the nth roots of unity form a regular n-gon inscribed on this circle.

[i] The first time I heard this identity (I was a senior in high school), I couldn't believe it. It was too simple and these three important constants seemly had nothing to do with each. Later I plugged the expression $e^{\pi i}$ into my brother's TI-89 graphing calculator, which is capable of both symbolic algebraic manipulation and complex arithmetic. To me, it was very impressive what difficult computations a TI-89 calculator could do with effortless-easy compared to other basic graphing calculators like the TI-84. Whatever the TI-89 told me was gospel, at least mathematical gospel. After plugging in the expression $e^{\pi i}$, the TI-89 instantly replied the quantity was equal to -1. I then thought, "Oh, I guess it is true." The book was then closed on my disbelief.

[ii] See §4.2 Multiplicative Group of Complex Numbers in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–10, convert each complex number from standard form to polar form or vice versa.
(Pick 5 from Exercises 1-10)

1. $3e^{\pi i/6}$
2. $5e^{9\pi i/4}$
3. $3e^{\pi i}$
4. $\frac{1}{2}e^{7\pi i/4}$
5. $1-i$
6. -5
7. $2+2i$
8. $\sqrt{3}+i$
9. $-3i$
10. $2i+2\sqrt{3}$

For Exercises 11–23, evaluate the given complex number.
(Pick 6 from Exercises 11-23)

11. $(3-2i)+(5i-6)$
12. $(4-5i)-\overline{(4i-4)}$
13. $(5-4i)(7+2i)$
14. $(9-i)\overline{(9-i)}$
15. i^{45}
16. $(1+i)+\overline{(1+i)}$
17. $(1+i)^{-1}$
18. $(1-i)^6$
19. $(\sqrt{3}+i)^5$
20. $(-i)^{10}$
21. $\left(\frac{1-i}{2}\right)^4$
22. $(-\sqrt{2}-\sqrt{2}i)^{12}$
23. $(-2+2i)^{-5}$

For Exercises 24–25, list and graph the nth roots of unity. Which are the generators of Z_n?

24. $n=5$
25. $n=6$

For Exercises 26–31, let $z, w \in \mathbb{C}$. Prove the given statement about complex numbers.
(Pick 1 from Exercises 26-31)

26. $|z| = |\overline{z}|$
27. $z\overline{z} = |z|^2$
28. $z^{-1} = \frac{\overline{z}}{|z|^2}$
29. $|z+w| \leq |z| + |w|$
30. $|z-w| \geq ||z| - |w||$
31. $|zw| = |z||w|$

"Do not brood over your past mistakes and failures as this will only fill your mind with grief, regret and depression. Do not repeat them in the future." – Swami Sivananda

Lecture Videos

The Order of Subgroups of a Cyclic Group

The Method of Repeated Squares

3.3 Orders of Group Elements

Theorem 3.3.1. *Let $G = \langle g \rangle$ be a cyclic group. If g has finite order n, then so does G, that is, $|G| = n$. Furthermore, if $h \in G$ and $|h| = n$, then h is a generator for G, that is, $G = \langle h \rangle$.*

Proof. Since g has finite order, we have that $\langle g \rangle = \{e, g, g^2, g^3, \ldots, g^{n-1}\} = G$. Thus, $|G| = n$. If h has order n, then similarly $\langle h \rangle = \{e, h, h^2, \ldots, h^{n-1}\} = G$, since it is a finite subset with the same order. Therefore, h is a generator of G. □

Theorem 3.3.2. *Let G be a group with $g \in G$. Then $g^a = e$ if and only if $|g|$ divides a. In particular, if $h = g^m$, then the order of h divides the order of g.*

Proof. Let $|g| = n$. By the division algorithm, we have $g, r \in \mathbb{Z}$ such that $a = qn + r$ and $0 \leq r < n$. Now, $g^r = g^{a-qn} = g^a(g^n)^{-q} = g^a$. Thus, $g^a = e$ if and only if $g^r = e$. By the minimality of n as the order of g, $g^r = e$ if and only $r = 0$, that is, $a = qn$.

Note $h^n = (g^m)^n = (g^n)^m = e^m = e$. Thus, $|h|$ divides n. □

Corollary 3.3.3. *Let $G = \langle g \rangle$ be a cyclic group. Then G has exactly one subgroup of order d for each $d \mid |G|$.*

Proof. Let $|G| = n$ and $d \mid n$. Let $h = g^{n/d}$. Then clearly we have $h^d = (g^{n/d})^d = g^n = e$. On the other hand, if $1 \leq a < d$ is an integer then $1 \leq an/d < n$. Thus, $h^a = g^{an/d} \neq e$. Therefore, $|h| = d$, which implies that $|\langle h \rangle| = d$.

Let $k \in G$ such that $|k| = d$. There exists some integer m such that $k = g^m$. Then $g^{dm} = (g^m)^d = k^d = e$. By the previous theorem, we have that $n \mid dm$, that is, there is some $\ell \in \mathbb{Z}$ such that $n\ell = dm$. But this gives $(n/d)\ell = m$ and $n/d \mid m$. Therefore, $k = g^m \in \langle h \rangle$, that is, $\langle h \rangle = \langle k \rangle$. □

Corollary 3.3.4. *Let $G = \langle g \rangle$ be a cyclic group with $|G| = n$ and $h = g^m$. Then $|h| = n/\gcd(n, m)$.*

Corollary 3.3.5. *Let $H = \langle h \rangle \leq \mathbb{Z}_n$. Then the set of generators of H is $\{h^m \mid \gcd(m, |H|) = 1\}$. In particular, the set of generators of $\mathbb{Z}_n$ is $\mathbb{Z}_n^*$.*

The previous properties of cyclic groups allow for simple calculations of orders of elements. In arbitrary groups, order calculations are not so simple. A naive approach to compute the order of an element $g \in G$ is to begin computing the list

$$g, g^2, g^3, g^4, \ldots,$$

terminating when $g^n = e$. Clearly this is a very inefficient algorithm if n is sufficiently large. Lagrange's Theorem (forth coming) will provide a criterion which will greatly shrink this list, but we will still need to be able to compute sufficiently large powers of an element in the group. For example, calculating the quantity $2^{2^{1000000}}$ is probably beyond all hope.

In finite groups, such calculations are much more hopeful. Consider the calculation

$$2^{37398332} \pmod{46389}.$$

This must be an element of $\mathbb{Z}_{46389}$, which in particular will be a number between 0 and 46388. If we reduce by the modulus along the way, the quantity need never exceed the value of 46389^2. The strategy rests upon writing the exponent of $x^a \pmod n$ in binary representation:

$$a = 2^{k_1} + 2^{k_2} + \ldots + 2^{k_r}.$$

Then we compute each of the following powers recursively

$$\begin{aligned} a^{2^0} &\pmod n \\ a^{2^1} &\pmod n \\ &\vdots \\ a^{2^{k_r}} &\pmod n \end{aligned}$$

Then $x^a \equiv x^{2^{k_1}} \cdot x^{2^{k_3}} \cdots x^{2^{k_r}} \pmod n$. Thus, we can compute x^a in at most $2\log_2(a) - 1$ computations. This method is known as the **method of repeated squares** and is very important for computer algebra software.

Example 3.3.6 (Repeated Squares Algorithm)**.** We will compute $271^{321} \pmod{481}$ by "hand."[i] To begin, we write 321 in binary form. Certainly, 256 is the largest power of 2 less than 321. Then

$$321 - 256 = 65 = 64 + 1 \quad \Rightarrow \quad 321 = 256 + 64 + 1 = 2^8 + 2^6 + 2^0.$$

Next, we compute $271^{2^8} \pmod{481}$, recursively:

$$\begin{aligned}
271^{2^0} &\equiv 271^1 &&&&\equiv 271 \pmod{481} \\
271^{2^1} &\equiv 271^2 &&\equiv 73441 &&\equiv 329 \pmod{481} \\
271^{2^2} &\equiv 271^4 \equiv 329^2 &&\equiv 108241 &&\equiv 16 \pmod{481} \\
271^{2^3} &\equiv 271^8 \equiv 16^2 &&&&\equiv 256 \pmod{481} \\
271^{2^4} &\equiv 271^{16} \equiv 256^2 &&\equiv 65536 &&\equiv 120 \pmod{481} \\
271^{2^5} &\equiv 271^{32} \equiv 120^2 &&\equiv 14400 &&\equiv 451 \pmod{481} \\
271^{2^6} &\equiv 271^{64} \equiv 451^2 &&\equiv 203401 &&\equiv 419 \pmod{481} \\
271^{2^7} &\equiv 271^{128} \equiv 419^2 &&\equiv 175561 &&\equiv 477 \pmod{481} \\
271^{2^8} &\equiv 271^{256} \equiv 477^2 &&\equiv 227529 &&\equiv 16 \pmod{481}
\end{aligned}$$

Therefore,

$$271^{321} \equiv 271^{2^0+2^6+2^8} \equiv 271 \cdot 419 \cdot 16 \equiv (113549)16 \equiv 33(16) \equiv 528 \equiv \boxed{47} \pmod{481}.$$

Example 3.3.7. MAGMA[ii] is a computer algebra system which uses repeated squares to compute powers in finite arithmetic. For example, the MAGMA code

```
Modexp(271, 321, 481);
```

returns the value `47`. We may also us MAGMA to "show our work." For the binary expansion, note that command

```
Intseq(321,2);
```

returns `[ 1, 0, 0, 0, 0, 0, 1, 0, 1 ]`, which means that $271 = 1 \cdot 2^0 + 0 \cdot 2^1 + 0 \cdot 2^2 + 0 \cdot 2^3 + 0 \cdot 2^4 + 0 \cdot 2^5 + 1 \cdot 2^6 + 0 \cdot 2^7 + 1 \cdot 2^8$, as we observed earlier. Next, we can then command repeated squares, mimicking the MAGMA code from before, e.g. `329^2 mod 481;` returns `16`. Finally, we compute the product by the command `271*419*16 mod 481;` which returns `47`.

We can also use the simpler code `271^321 mod 481;` in many cases, but this does not actual utilize the repeated squares algorithm. Instead, this code computes 271^{321} then reduces modulo 481. For "small" exponents and "small" bases, this does not make much of a difference, but for larger powers, the repeated squares algorithm will be necessary. As such, to compute $a^e \pmod{n}$ using Repeated Squares, use `Modexp(a,e,n);` as above.

[i] That is, a four-function calculator. Who does basic arithmetic by hand?!

[ii] MAGMA can be accessed at http://magma.maths.usyd.edu.au/calc/. Please also see Appendix B for further details.

[iii] See §4.1 Cyclic Subgroups and §4.3 The Method of Repeated Squares in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–13, list all elements in the cyclic subgroup generated by the given element.
(Pick 7 from Exercises 1-13)

1. $7 \in \mathbb{Z}$
2. $15 \in \mathbb{Z}_{24}$
3. $8 \in \mathbb{Z}_{12}$
4. $8 \in \mathbb{Z}_{60}$
5. $8 \in \mathbb{Z}_{13}$
6. $8 \in \mathbb{Z}_{48}$
7. $3 \in \mathbb{Z}_{20}^*$
8. $5 \in \mathbb{Z}_{18}^*$
9. $7 \in \mathbb{R}^*$
10. $i \in \mathbb{C}^*$
11. $2i \in \mathbb{C}^*$
12. $\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i \in \mathbb{C}^*$
13. $\frac{1}{2} + \frac{\sqrt{3}}{2}i \in \mathbb{C}^*$

For Exercises 14–19, prove the given statement about generators of cyclic groups.
(Pick 1 from Exercises 14-19)

14. There are exactly two cyclic groups with a unique generator.
15. There are exactly four cyclic groups with a exactly two generators.
16. If p and q are distinct primes, then $\mathbb{Z}_{pq}$ has $(p-1)(q-1)$ generators.
17. If p is a prime and $r \in \mathbb{Z}$, then $\mathbb{Z}_{p^r}$ has $(p-1)p^{r-1}$ generators.
18. If $g \in G$ and $m, n \in \mathbb{Z}$, then $\langle g^m \rangle \cap \langle g^n \rangle = \langle g^{lcm(m,n)} \rangle$.
19. The cyclic group $\mathbb{Z}_n$ has an even number of generators for $n > 2$.

For Exercises 20–22, prove the given statement about cyclic groups.
(Pick 1 from Exercises 20-22)

20. Let $G = \langle g \rangle$ be a finite cyclic group of order n. If $h = g^k$ where $\gcd(n, k) = 1$, then $\langle h \rangle = G$.
21. Let G be an abelian group of order pq where $\gcd(p, q) = 1$. If $g, h \in G$ such that $|g| = p$ and $|h| = q$, then G is cyclic.
22. If G is a cyclic group of order n and $d \mid n$, then G has a subgroup of order d.

For Exercises 23–26, evaluate the given modular exponent using Repeated Squares.

23. $292^{3171} \pmod{582}$
24. $2557^{341} \pmod{5681}$
25. $2071^{9521} \pmod{4724}$
26. $971^{321} \pmod{765}$

3.4 Supplemenatry Exercises

(Go to Solutions)

For Exercises 1–5, prove or disprove the given statement about cyclic groups or orders.

1. All of the generators of $\mathbb{Z}_{60}$ are prime numbers.
2. $\mathbb{Z}_8$ is cyclic.
3. $\mathbb{Q}$ is cyclic.
4. If every proper subgroup of a group G is cyclic, then G is a cyclic group.
5. A group with a finite number of subgroups is finite.

6. Fill all elements of finite order in $\mathbb{Z}$, $\mathbb{Q}^*$, and $\mathbb{R}^*$
7. If $g^{24} = 3$ in a group G, what are the possible orders of g?
8. If g and h have orders 15 and 16, respectively, in group G, what is the order of $\langle g \rangle \cap \langle h \rangle$?
9. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$. Show that $|A|, |B| < \infty$ but $|AB| = \infty$.
10. List each generator of the unique cyclic subgroup of order 8 in $\mathbb{Z}_{32}$.
11. For $n \leq 20$, which groups $\mathbb{Z}_n^*$ are cyclic?
12. What are all the cyclic subgroups of Q_8?
13. Prove that the subgroups of $\mathbb{Z}$ are exactly $n\mathbb{Z}$ for $n \in \mathbb{N}$.
14. Prove that the generators of $\mathbb{Z}_n$ are the integers r such that $1 \leq r < n$ and $\gcd(r, n) = 1$.
15. For what integers n is -1 an nth root of unity?
16. Prove that $Z_n \leq S^1$ for all $n \geq 1$.
17. Let $z \in S^1$. Prove that $z^n = 1$ and $z^m = 1$ if and only if $z^{\gcd n,m} = 1$.
18. Let $z \in \mathbb{C}^* \setminus S^1$. Then z has infinite order in $\mathbb{C}^*$.
19. Let $e^{i\theta} \in S^1$. Prove that if $\theta \in \mathbb{Q}$, then $e^{i\theta}$ has infinite order.

Chapter 4

Permutation Groups

"Symmetry is a vast subject, significant in art and nature. Mathematics lies at its root, and it would be hard to find a better one on which to demonstrate the working of the mathematical intellect." – Hermann Weyl

Lecture Videos

Cycle Notation

Permutation Multiplication

The Inverse of a Permutation

4.1 The Symmetric Group

Recall that if X is a set then S_X is the set of permutations on X and is a group under permutation multiplication known as the **symmetry group** on X. When $X = \{1, 2, \ldots, n\}$, we denote this instead as S_n. Any subgroup of S_X is known as a **permutation group**.

The usual tableaux we have used for representing a finite permutation can be very cumbersome. We introduce a new, simpler notation for permutations, called **cycle notation**. To begin, we talk about cycles in a permutation group.

Definition 4.1.1. A permutation $\sigma \in S_X$ is called a **cycle** of length k if there exists a positive integer k and elements $a_1, a_2, \ldots, a_k \in X$ such that

$$\sigma : a_1 \mapsto a_2 \mapsto a_3 \mapsto \ldots \mapsto a_k \mapsto a_1$$

and σ fixes all other elements of X. We denote this as $\sigma = (a_1 a_2 \ldots a_k)$. We say that two cycles are **disjoint** if the two cycles have no common elements between them.

Example 4.1.2. The following two permutations are cycles in S_7 and S_6, respectively:

$$\sigma : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 3 & 5 & 1 & 4 & 2 & 7 \end{pmatrix} = (162354) \qquad \tau : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 3 & 5 & 6 \end{pmatrix} = (243)$$

is a cycle of length 6 and is a cycle of length 3.

As displayed in the previous example, one can start with $x \in X$ and can follow the images $x, \sigma(x), \sigma^2(x), \sigma^3(x), \ldots$ until the path returns to x. From this we can construct the cycle starting at x. Of course, the cycle starting at x is none other than $\{\pi(x) \mid \pi \in \langle\sigma\rangle\}$. Of course, if we had started at any other element in the cycle we would construct exactly the same cycle. In fact, two elements being contained in the same cycle generated by a permutation is an equivalence relation.

But not every permutation is a cycle. On the other hand, every permutation can be written as a product of disjoint cycles, up to reordering the cycles. We can construct the cycle decomposition, or *cycle structure*, of any permutation by starting with any element $x \in X$ and computing the cycle containing x. If there are any elements of X not contained in this cycle, we take such an element and compute the cycle containing it. We repeat this process until X has been partitioned. When a cycle of length 1 occurs, we omit it from the cycle notation. This occurs exactly when the permutation has a fixed point. If all cycles are length one, then points are fixed and the permutation is the identity. We will denote this as simply 1.

Theorem 4.1.3. *Every permutation in S_X can be written as a product of disjoint cycles.*

Example 4.1.4. We find the cycle decomposition for the following permutations:

First, $\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 1 & 3 & 6 & 5 \end{pmatrix} = (1243)(56)$ is a product of a 4-cycle and a 2-cycle. We might call ρ a 4-2-cycle. We can visualize this process as: $1 \longmapsto 2 \longmapsto 4 \longmapsto 3 \longmapsto 1; \quad 5 \longmapsto 6 \longmapsto 5.$

Next, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 3 & 1 & 5 & 2 \end{pmatrix} =$ $(1624)(3)(5) = (1624)$ is a 4-cycle.

Finally, $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 5 & 6 & 4 \end{pmatrix} =$ $(13)(2)(456) = (13)(456)$ is a 2-3-cycle.

Products of permutations can all easily be computed using cycle notation. Let σ and τ be two cycles. We first compute $\sigma(x)$. To do this, we find x in the cycle of σ. If present, $\sigma(x)$ is the element immediately to the right of x in the cycle notation. If not present, $\sigma(x)$. Since $\tau\sigma(x)$ means $\tau(\sigma(x))$, we repeat this process for the image $\sigma(x)$ on τ. If we consider all $x \in X$, we can then build the cycle structure like the previous example. With permutation products, always work right to left.

Example 4.1.5. Let $\sigma = (1352)$ and $\tau = (256)$. Then

$$\sigma\tau = (1352)(256) = (1356) \quad \text{and} \quad \tau\sigma = (256)(1352) = (1362).$$

The first product is computed in the following way:

$$\mathbf{1} \overset{\tau}{\longmapsto} 1 \overset{\sigma}{\longmapsto} \mathbf{3} \overset{\tau}{\longmapsto} 3 \overset{\sigma}{\longmapsto} \mathbf{5} \overset{\tau}{\longmapsto} 6 \overset{\sigma}{\longmapsto} \mathbf{6} \overset{\tau}{\longmapsto} 2 \overset{\sigma}{\longmapsto} \mathbf{1}; \quad \mathbf{2} \overset{\tau}{\longmapsto} 5 \overset{\sigma}{\longmapsto} \mathbf{2}; \quad \mathbf{4} \overset{\tau}{\longmapsto} 4 \overset{\sigma}{\longmapsto} \mathbf{4}$$

and the second as

$$\mathbf{1} \overset{\sigma}{\longmapsto} 3 \overset{\tau}{\longmapsto} \mathbf{3} \overset{\sigma}{\longmapsto} 5 \overset{\tau}{\longmapsto} \mathbf{6} \overset{\sigma}{\longmapsto} 6 \overset{\tau}{\longmapsto} \mathbf{2} \overset{\sigma}{\longmapsto} 1 \overset{\tau}{\longmapsto} \mathbf{1}; \quad \mathbf{4} \overset{\sigma}{\longmapsto} 4 \overset{\tau}{\longmapsto} \mathbf{4}; \quad \mathbf{5} \overset{\sigma}{\longmapsto} 2 \overset{\tau}{\longmapsto} \mathbf{5}.$$

As is expected, permutation multiplication is noncommutative.

Example 4.1.6. Let $\sigma = (1624)$ and $\tau = (13)(456)$. Then

$$\sigma\tau = (1624)(13)(456) = (136)(245) \quad \text{and} \quad \tau\sigma = (13)(456)(1624) = (143)(256).$$

Again, the products of permutations tend to be noncommutative. But this is not always the case. Let $\pi = (123)$ and $\rho = (456)$. Then

$$\pi\rho = (123)(456) = (456)(123) = \rho\pi.$$

This time the factors commute.

Proposition 4.1.7. *Let σ and τ be two disjoint cycles in S_X. Then $\sigma\tau = \tau\sigma$.*

Proof. Let $\sigma = (a_1 a_2 \dots a_k)$ and $\tau = (b_1 b_2 \dots b_\ell)$. To show that two functions are equal, namely $\sigma\tau = \tau\sigma$, it suffices to show that their images are equal for each element in the domain, that is, $\sigma\tau(x) = \tau\sigma(x)$ for all $x \in X$. In this direction, let $x \in X$. Suppose next that $x = a_i$ for some a_i. Since the cycles are disjoint, $x \neq b_j$ for any j. Then $\tau(x) = x$. We then get

$$\sigma\tau(x) = \sigma\tau(a_i) = \sigma(a_i) = a_{i+1} = \tau(a_{i+1}) = \tau\sigma(a_i).$$

The case that $x = b_j$ is handled similarly. Finally, suppose that $x \neq a_i, b_j$ for any i and j. Then x is fixed by both σ and τ. Then

$$\sigma\tau(x) = \sigma(x) = x = \tau(x) = \tau\sigma(x).$$

Therefore, $\sigma\tau = \tau\sigma$. □

Inverses of permutations in cycle notation is very convenient. If $\sigma = (a_1 a_2 \dots a_k)$ is a cycle then $\sigma^{-1} = (a_k a_{k-1} \dots a_1) = (a_1 a_k a_{k-1} \dots a_2)$.

Example 4.1.8. Let $\sigma = (1624)$ and $\tau = (13)(456)$. Then $\sigma^{-1} = (1624)^{-1} = (4261) = (1426)$ and $\tau^{-1} = \big((13)(456)\big)^{-1} = (456)^{-1}(13)^{-1} = (654)(31) = (13)(465)$.

[i] See §5.1 Definitions and Notation in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, write the given permutation in cycle notation.

1. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}$ 2. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix}$ 3. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 4 & 2 \end{pmatrix}$ 4. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{pmatrix}$

For Exercises 5–19, write the given permutation in cycle notation.
(Pick 8 from Exercises 5-19)

5. $(1245)(234)$
6. $(12)(1253)$
7. $(143)(23)(24)$
8. $(1423)(34)(56)(1324)$
9. $(1254)(13)(25)$
10. $(1254)(13)(25)^2$
11. $(1254)^{-1}(123)(45)(1254)$
12. $(1254)^2(123)(45)$
13. $(123)(45)(1254)^{-2}$
14. $(1254)^{100}$
15. $|(1254)|$
16. $|(1254)^2|$
17. $(12)^{-1}$
18. $(12537)^{-1}$
19. $[(1235)(467)]^{-1}$

For Exercises 20–21, let $\tau = (a_1, a_2, \ldots, a_k)$ be a cycle of length k. Prove the given statement.

20. If $\sigma \in S_n$, then $\sigma\tau\sigma^{-1} = (\sigma(a_1), \sigma(a_2), \ldots, \sigma(a_k))$ is a cycle of length k.
21. Let μ be a cycle of length k. There exists a permutation σ such that $\sigma\tau\sigma^{-1} = \mu$.

For Exercises 22–28, prove the given statement about permutations.
(Pick 2 from Exercises 22-28)

22. Let $\sigma = \sigma_1 \cdots \sigma_m \in S_n$ be the product of disjoint cycles. Then $|\sigma| = \text{lcm}(|\sigma_1|, \ldots, |\sigma_m|)$.
23. If σ is a cycle of odd length, then σ^2 is also a cycle.
24. Any element in S_n can be written as a finite product of transpositions from $\{(12), (13), \ldots, (1n)\}$.
25. Any element in S_n can be written as a finite product of transpositions from $\{(12), (23), \ldots, (n-1, n)\}$.
26. Any element in S_n can be written as a finite product of permutations from $\{(12), (12\ldots n)\}$.
27. Let G be a group and define a map $\lambda_g : G \to G$ by $\lambda_g(x) = gx$. Then λ_g is a permutation of G.
28. Let $\alpha \in S_n$ for $n \geq 3$. If $\alpha\beta = \beta\alpha$ for all $\beta \in S_n$, then $\alpha = 1$. This implies that $Z(S_n) = \{1\}$.

"There's no point in living in an alternate reality." – Jessica Cutler

Lecture Videos

Transpositions

Even and Odd Permutations

The Alternating Group

4.2 The Alternating Group

We will continue our discussion of permutations on the set X. For this lecture, let X be a set containing more than one element. If $|X| = n$, then we may assume without the loss of generality that $X = \{1, 2, \ldots, n\}$.

Definition 4.2.1. A **transposition** is a permutation of length 2.

Theorem 4.2.2. *Any permutation can be decomposed as a product of transpositions.*

Proof. Note that the identity permutation can be expressed as $1 = (a_1a_2)(a_1a_2)$. Thus, we can assume that σ is a nontrivial permutation. Since σ can be expressed as a product of disjoint cycles, it suffices that cycles can be decomposed as a product of transpositions. Suppose that $\sigma = (a_1a_2a_3 \ldots a_k)$ is a k-cycle. But then

$$\sigma = (a_1a_2a_3 \ldots a_k) = (a_1a_k)(a_1a_{k-1}) \ldots (a_1a_4)(a_1a_3)(a_1a_2). \qquad \square$$

Example 4.2.3. Consider the permutation:

$$\sigma = (16)(253) = (16)(23)(25).$$

Thus, σ can be written of a product of 3 transpositions. Unlike cycle decompositions, transposition decompositions are not unique. For example,

$$\sigma = (16)(25)(35) = (16)(35)(23) = (16)(45)(23)(45)(25).$$

Theorem 4.2.4. *Every permutation can be expressed as a product of an even number of transpositions or an odd number of transpositions but not both.*

Proof. Let $\boldsymbol{e}_i \in \mathbb{R}^n$ be the vector with a one in the ith position and zeros elsewhere. Then $I_n = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \ldots & \boldsymbol{e}_n \end{bmatrix}$. Let $\sigma \in S_n$. Then define $[\sigma] = \begin{bmatrix} \boldsymbol{e}_{\sigma(1)} & \boldsymbol{e}_{\sigma(2)} & \ldots & \boldsymbol{e}_{\sigma(n)} \end{bmatrix}$. Clearly, the matrix $[\sigma]$ is formed by permuting the columns of I_n and as such is called a **permutation matrix**. It is a fact from linear algebra that interchanging any two rows or columns in a square matrix changes the sign of the determinant. Thus, $\det([\sigma]) = \pm 1$. Since σ can be expressed as a product of transpositions, then $[\sigma]$ can be factored into a product of transposition matrices, that is, a square matrix where only two columns of I_n are interchanged[i]. Each transposition matrix has a determinant equal to -1. Thus, $\det([\sigma]) = (-1)^k$, where k is the number transpositions in the product. If k is odd, then $\det([\sigma]) = -1$. If k is even, then $\det([\sigma]) = 1$. As the determinant of a matrix cannot be 1 and -1, no permutation can be expressed as a product of an even and an odd number of transpositions. $\square$

Definition 4.2.5. We say that a permutation is **even** if it can be expressed as a product of an even number of transpositions. We say that a permutation is **odd** if it can be expressed as a product of an odd number of transpositions. Let the set A_n denote the set of all even permutations in S_n, which is called the **alternating group** on n letters.

Theorem 4.2.6. *The subset A_n is a subgroup of S_n.*

Proof. Let $\sigma, \tau \in A_n$. Since σ and τ are even permutations, we can be expressed as a product of $2s$ and $2t$ transpositions., respectively, where $s, t \in \mathbb{N}$. Then $\sigma\tau$ can be expressed as a product of $(2s)+(2t) = 2(s+t)$ transpositions. Thus, $\sigma\tau \in A_n$. Like shown in the proof of Theorem (i), the identity permutations is even. For $\sigma \in A_n$, suppose that

$$\sigma = \sigma_1\sigma_2 \cdots \sigma_{2s}$$

where each σ_i is a transposition. Then

$$\sigma^{-1} = \sigma_{2s}^{-1}\sigma_{2s-1}^{-1} \cdots \sigma_1^{-1} = \sigma_{2s}\sigma_{2s-1} \cdots \sigma_1 \in A_n$$

since every transposition is equal to its own inverse, that is, $(a_1a_2)^{-1} = (a_2a_1) = (a_1a_2)$. Therefore, $A_n \leq S_n$. □

Recall that we have shown previously that $|S_n| = n!$. A similar result holds for the alternating group.

Theorem 4.2.7. *The order of A_n is $n!/2$.*

Proof. Let B_n denote the set of odd permutations. Let $f : A_n \to B_n$ be given as $f(\sigma) = \sigma(12)$. Since σ is an even permutation, we get that $f(\sigma)$ is odd. Suppose that $f(\sigma) = f(\tau)$ for some $\sigma, \tau \in S_n$. Then $\sigma(12) = \tau(12)$, but right cancellation in S_n guarantees that $\sigma = \tau$. Thus, f is injective (one-to-one). Let $\rho \in B_n$. Then $\rho(12) \in A_n$ and $f(\rho(12)) = \rho(12)(12) = \rho$. Thus, f is surjective (onto). We conclude that f is a bijection. Therefore, $|A_n| = |B_n|$. But $S_n = A_n \cup B_n$ and $A_n \cap B_n = \emptyset$ by Theorem (ii). This implies that $|S_n| = |A_n| + |B_n|$. Combining this equations then gives $|A_n| = |S_n|/2$. □

Example 4.2.8. The group A_4 consists of $12 = 4!/2$ permutations listed below, consisting of the identity, all 2-2-cycles, and all 3-cycles:

1	(12)(34)	(13)(24)	(14)(23)
(123)	(124)	(134)	(234)
(132)	(142)	(143)	(243)

The other twelve elements of S_n are odd and consist of $\binom{4}{2} = 6$ transposition and $3! = 6$ 4-cycles. The alternating group A_4 has cyclic subgroups of order 1, 2, and 3 generated by the identity, a 2-2-cycle, or a 3-cycle, respectively. Also, $V_4 = \{1, (12)(34), (13)(24), (14)(23)\}$ is a non-cyclic subgroup of order 4 and another representation of the Klein 4-group. Interestingly enough, A_4 has no subgroup of order 6. Thus, in contrast to cyclic groups, there need not be a subgroup of order d for every divisor d of the group's order.

[i] Such as matrix is one of three types of **elementary matrices** which correspond to the three elementary row (or column) operations. These transposition matrices correspond to the interchange elementary operation which interchanges the positions of any two rows (or columns). The alternating group consists of the permutation matrices with determinant equal to 1.

[ii] See §5.1 Definitions and Notation in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–5, express the given permutation as products of transpositions and identify whether it is even or odd.

1. (14356)
2. (156)(234)
3. (1426)(142)
4. (17254)(1423)(154632)
5. 142637
6. What are the possible cycle structures of elements of A_5? Of A_6?

(Pick 1 from Exercises 7-10)

7. Find all of the subgroups of A_4 and their orders.
8. Find all possible orders of elements in S_7 and A_7.
9. Does A_8 contain an element of order 26? Why or why not?
10. Find an element of A_{10} of maximum order.

For Exercises 11–15, prove the given statement about transpositions.
(Pick 2 from Exercises 11-15)

11. The alternating group A_n is nonabelian for $n \geq 4$.
12. A 3-cycle is an even permutation.
13. Every even permutation is a product of 3-cycles.
14. If σ is even, then σ^{-1} is likewise even.
15. If $\sigma \in S_n$ and $\tau \in A_n$, then $\sigma\tau\sigma^{-1} \in A_n$

"If the Earth could be made to rotate twice as fast, managers would get twice as much done. If the Earth could be made to rotate twenty times as fast, everyone else would get twice as much done since all the managers would fly off." – Norman Ralph Augustine

Lecture Videos

Standard Notation for the Dihedral Group

The Generators of the Dihedral Group

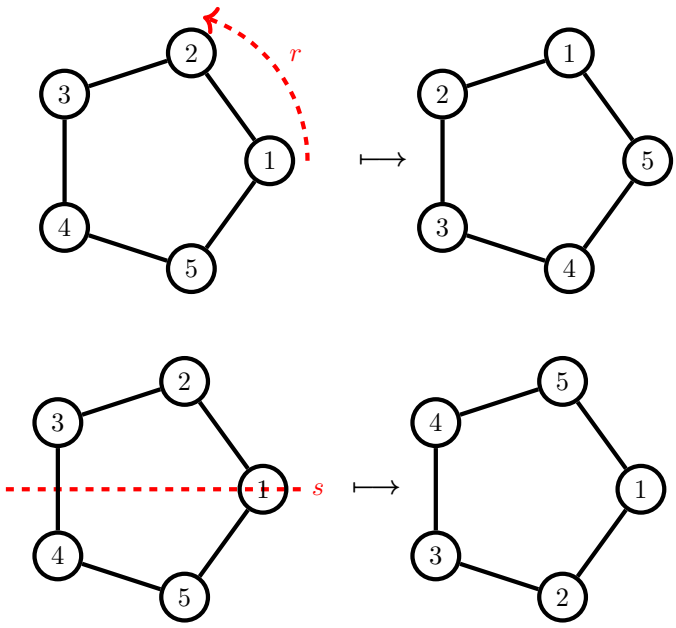
Normal Forms in the Dihedral Group

4.3 The Dihedral Group

Recall that D_n is the dihedral group of symmetries of the regular n-gon. We have seen this group in detail for D_3 and D_4, the symmetries of an equilateral triangle and square, respectively. Now, any regular n-gon can be inscribed into a circle whose center coincides with the center of the n-gon, that is, the common intersection point of each of the altitudes. As the side lengths of the polygon are irrelevant in terms of symmetry, we may assume that our polygon is inscribed in the unit circle. Using the unit circle in the complex plane, we can use complex numbers to denote the vertices of the n-gon, namely, $1, e^{2\pi/n}, e^{2(2\pi i/n)}, e^{3(2\pi i/n)}, \ldots, e^{(n-1)(2\pi i/n)}$.

One symmetry of the n-gon is counterclockwise rotation by $\theta = \dfrac{2\pi}{n}$. In terms of the complex roots of unity, this corresponds to multiplication by the primitive root $e^{2\pi i/n}$. Clearly this symmetry has order n. Another symmetry is reflection across the real axis. This can also be expressed as complex conjugation. This reflection clearly has order 2.

We will prove below that all symmetries of the n-gon are **generated** by these two symmetries, meaning that every symmetry can be expressed as a product using only r, s and their inverses. For example, we proved last time that S_n is generated by the set of transpositions. A cyclic group is exactly a group generated by a single element.



Lemma 4.3.1. *Let D_n be the dihedral group with r and s defined as above. Then $sr^k s^{-1} = r^{-k}$. In particular, $sr^k = r^{-k}s$.*

Proof. Consider the symmetry $sr^k s^{-1}$. Since $s^2 = 1$, we have $s^{-1} = s$ and $sr^k s^{-1} = sr^k s$. In terms of the geometry transformations, $sr^k s$ means to reflect across the real axis, rotation $\theta = \dfrac{2\pi k}{n}$ counterclockwise, and reflect back. The net effect to the polygon is then just rotation $\theta = \dfrac{2\pi k}{n}$ clockwise, which is exactly $(r^{-1})^k = r^{-k}$. This establishes $sr^k s^{-1} = r^{-k}$. Multiplying both sides on the right by s then gives $sr^k = r^{-k}s$. □

Theorem 4.3.2. *The group D_n is generated by the symmetries r and s defined above. In particular, every symmetry σ can be expressed in the form $\sigma = r^k s^\ell$ for $k, \ell \in \mathbb{Z}$.*

Proof. From geometry, the only symmetries of the unit circle are rotations and reflections. These are the only types of symmetries for the regular n-gon which is a subset of the unit circle. All counterclockwise rotational symmetries of the n-gon are powers of r. All clockwise rotational symmetries of the n-gon are powers of r^{-1}. Let ρ be a reflection across the line $\mathcal{L}$. If ρ is a symmetry of the n-gon, then $\mathcal{L}$ must pass through the origin and we can measure and represent $\mathcal{L}$ by the angle θ it makes with the positive real axis. In order to be a symmetry, $\mathcal{L}$ must pass through a vertex or the midpoint of an edge of the polygon. In either case, $\theta = k\left(\frac{\pi}{n}\right)$. Let r_θ be the rotation counterclockwise by θ. The $\rho = r_\theta s r_\theta^{-1}$. Then Lemma 8.1.6 shows that $\rho = r_\theta r_\theta s = r_{2\theta} s = r^k s$. Therefore, every symmetry is generated by r and s.

Each product of r's and s's is called a **word**. We can simplify words by collapsing down adjacent r's together and adjacent s's together. Since $s^2 = 1$, a generic word in D_n looks like

$$\sigma = r^{k_0} s r^{k_1} s r^{k_2} \cdots s r^{k_m},$$

where k_0 and k_m may be zero. We will induct on the number of s's in σ. For our base case, suppose that s appears zero times. Then $\sigma = r^{k_0} = r^{k_0} s^0$. For our inductive hypothesis, suppose that if σ involves $m-1$ many s's, then it can be expressed as $\sigma = r^k s^\ell$. Suppose that

$$\sigma = r^{k_0} s r^{k_1} s r^{k_2} \cdots s r^{k_m}.$$

Then $\sigma = \rho s r^{k_m}$ for some permutation ρ which can be expressed with $m-1$ many s's. By our inductive hypothesis, we get

$$\sigma = \rho s r^{k_m} = (r^k s^\ell) s r^{k_m} = r^k s^{\ell+1} r^{k_m} = r^k r^{(-1)^{\ell+1} k_m} s^{\ell+1},$$

where the last equality comes from Lemma 8.1.6. Therefore, the result follows by induction. □

We call the expression $\sigma = r^k s^\ell$ the **normal form** of σ, which is unique up to modulo n on k and modulo 2 on ℓ.

Corollary 4.3.3. *The order of D_n is $2n$.*

Proof. Since each symmetry in D_n can be expressed in the normal form $r^k s^\ell$ for some $k, \ell \in \mathbb{Z}$, we count the possible normal forms in D_n. Since $r^n = 1$ and $s^2 = 1$, there exists n choices for k and 2 choices for ℓ. This gives $|D_n| = 2n$. □

Example 4.3.4. Labeling the vertices of triangle clockwise as $1, 2, 3$ (or $1, e^{4/3\pi i}, e^{2/3\pi i}$), we get the 6 symmetries of the triangle:

$$\begin{array}{lll} 1 & r = (123) & r^2 = (132) \\ s = (23) & rs = (12) & r^2 s = (13) \end{array}$$

Similarly, if we label the vertices of the square clockwise as $1, 2, 3, 4$, we get the 8 symmetries of the square:

$$\begin{array}{llll} 1 & r = (1234) & r^2 = (13)(24) & r^3 = (1432) \\ s = (24) & rs = (12)(34) & r^2 s = (13) & r^3 s = (14)(23) \end{array}$$

Symmetry groups can also be discussed for three-dimensional objects. For example, the groups of symmetries of the tetrahedron and cube are A_4 and S_4, respectively. For the tetrahedron, each symmetry permutes the four vertices, so naturally this permutation group is a subgroup of S_4. As for the cube, symmetries of

the cube necessarily permute the four interior diagonals of the cube, thus giving a permutation group in S_4.

[i]See §5.2 Dihedral Groups in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

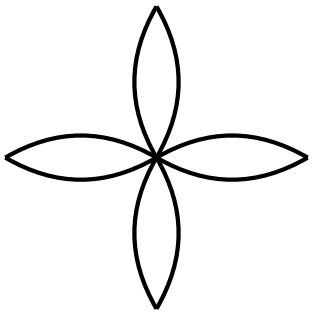
1. Write element in D_5 in cycle notation and in normal form.
2. Prove that D_n is nonabelian for $n \geq 3$.
3. Recall that the **center** of a group is $Z(G) = \{z \in G \mid \forall g \in G, gz = zg\}$. Find $Z(D_8)$, $Z(D_{10})$, and $Z(D_{2n})$.

Definition 4.3.5. We say that a **rigid motion** if a function $f : \mathbb{R}^n \to \mathbb{R}^n$, for $n \in \mathbb{N}$, if $\|f(\boldsymbol{x}) - f(\boldsymbol{y})\| = \|\boldsymbol{x} - \boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, that is, f preserves the distance between vectors in the domain and vectors in the image. Let $X \subseteq \mathbb{R}^n$. Then **the symmetry group** of X is the set of rigid motions $f : \mathbb{R}^n \to \mathbb{R}^n$ for which $f(X) = X$.

In other words, a symmetry is a rigid motion that moves the shape back onto itself. Note that if we view the regular n-gon as a subset of $\mathbb{R}^2$, then D_n is the symmetry group of the regular n-gon.

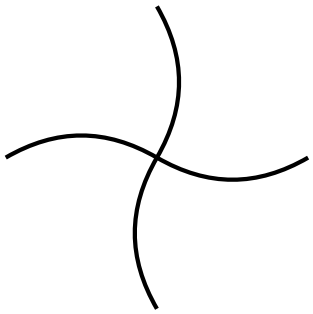
For Exercises 4–12, determine the symmetry group of the given subset of $\mathbb{R}^2$. Note that symmetries of colored shapes must preserve the coloring.

4.

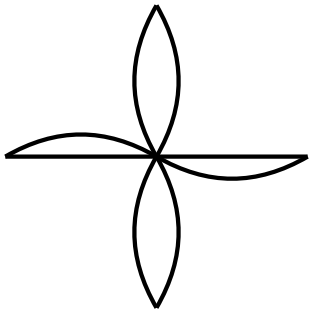


5.

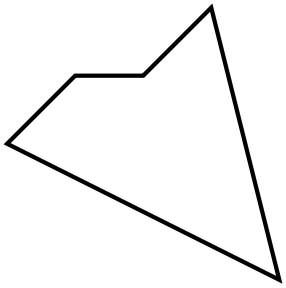
6.



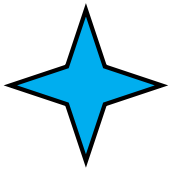
7.



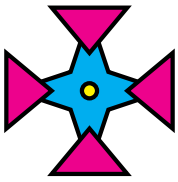
8.



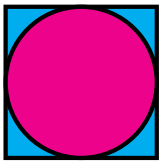
9.



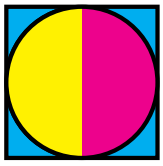
10.



11.



12.



4.4 Supplemenatry Exercises

(Go to Solutions)

1. Find $(a_1 a_2 \dots a_n)^{-1}$.
2. List all of the subgroups of S_4.

For Exercises 3–5, compute the given subset of S_4. Is it a subgroup?

3. $\{\sigma \in S_4 \mid \sigma(1) = 3\}$
4. $\{\sigma \in S_4 \mid \sigma(2) = 2\}$
5. $\{\sigma \in S_4 \mid \sigma(1) = 3, \sigma(2) = 2\}$

6. Find an element of largest order in S_n for $n = 3, \dots, 10$.
7. Let $\sigma \in S_n$ have order n. Show that for all integers i and j, $\sigma^i = \sigma^j$ if and only if $i \equiv j \pmod{n}$.
8. Prove that S_n is nonabelian for $n \geq 3$.
9. If $\sigma \in S_n$ is not a cycle, then σ can be written as the product of at most $n - 2$ transpositions.
10. If $\sigma \in S_n$ can be expressed as an odd number of transpositions, show that any other product of transpositions equaling σ must also be odd.
11. Show that $\alpha^{-1}\beta^{-1}\alpha\beta \in A_n$ for all $\alpha, \beta \in S_n$.
12. Show that, in D_n, $srs = r^{-1}$.
13. Show that, in D_n, $r^k s = sr^{-k}$ for all $k \in \mathbb{Z}$.
14. Show that, in D_n, $|r^k| = \dfrac{n}{\gcd(k, n)}$ for all $k \in \mathbb{Z}$.
15. Show that the symmetry group of the tetrahedron is A_4.
16. Show that the symmetry group of the cube is S_4.
17. For $\alpha, \beta \in S_n$, define $\alpha \sim \beta$ if there exists some $\sigma \in S_n$ such that $\sigma\alpha\sigma^{-1} = \beta$. Show that $\sim$ is an equivalence relation on S_n. We say that α and β are **conjugates** if $\alpha \sim \beta$.
18. Show that $\alpha, \beta \in S_n$ have the same cycle structure if and only if α and β are conjugates.

For Exercises 19–23, for a fixed permutation $\sigma \in S_X$, define a relation σ on X by the rule $x \sim_\sigma y$ if and only if there exists a $n \in \mathbb{Z}$ such that $\sigma^n(x) = y$, for $x, y \in X$.

19. Prove that $\sim_\sigma$ is an equivalence relation.
20. Compute the equivalence classes of $\sim_\sigma$ for each element of $X = \{1, 2, 3, 4, 5\}$ is $\sigma = (1254)$.
21. Compute the equivalence classes of $\sim_\sigma$ for each element of $X = \{1, 2, 3, 4, 5\}$ is $\sigma = (123)(45)$.
22. Compute the equivalence classes of $\sim_\sigma$ for each element of $X = \{1, 2, 3, 4, 5\}$ is $\sigma = (13)(25)$.

> **Definition 4.4.1.** Let $H \leq S_X$. We say that H is **transitive** if for all $x, y \in X$ there exists $\sigma \in H$ such that $\sigma(x) = y$.

23. Prove that H is transitive if and only if $\sim_\sigma$ has a single equivalence class for each $\sigma \in H$.

Chapter 5

Cosets

"War does not determine who is right – only who is left." – Bertrand Russell

Lecture Videos

Cosets

Cosets Partition a Group

Counting Cosets

5.1 Cosets

Definition 5.1.1. Let G be a group and let $H \leq G$. Define a **left coset** of H with representative $g \in G$ to be the set

$$gH = \{gh \mid h \in H\}.$$

Let G/H denote the set of left cosets of G with respect to H.

Similarly, we define the **right coset** as

$$Hg = \{hg \mid h \in H\}.$$

Let $H\backslash G$ denote the set of right cosets of G with respect to H.

It should be mentioned that although cosets are subsets of G constructed from a subgroup H of G, the cosets gH and Hg are not necessarily subgroups themselves, nor are they necessarily equal.

Example 5.1.2. Let $H = \langle 3 \rangle \leq \mathbb{Z}_6$, that is, $H = \{0, 3\}$. There are then three left cosets:

$$\begin{aligned} 0 + H &= 3 + H = \{0, 3\} \\ 1 + H &= 4 + H = \{1, 4\} \\ 2 + H &= 5 + H = \{2, 5\} \end{aligned}$$

Because $\mathbb{Z}_6$ is an additive group, the cosets are written with additive notation as well. When the identity is used to represent a coset, such as $0 + H$, it will be more convenient to suppress the identity and express the coset as just H. After all, $0 + H = H$. Because Z_6 is abelian, the right cosets are the same as the left cosets.

Example 5.1.3. Let $H = A_3 = \langle (123) \rangle \leq S_3$, that is, $H = \{1, (123), (132)\}$. There are two distinct left cosets of H:

$$\begin{aligned} H &= (123)H = (132)H = \{1, (123), (132)\} \\ (12)H &= (13)H = (23)H = \{(12), (13), (23)\} \end{aligned}$$

Interesting enough, one coset is the set of even permutations (A_3) and the other is the set of odd permutations (B_3). Even though S_3 is nonabelian, it is still the case that the right cosets coincide with the left cosets:

$$\begin{aligned} H &= H(123) = H(132) = \{1, (123), (132)\} \\ H(12) &= H(13) = H(23) = \{(12), (13), (23)\} \end{aligned}$$

Before we begin to believe this is always that case, let $K = \langle (12) \rangle = \{1, (12)\}$. Then the left and right cosets are listed below:

$$\begin{aligned} K &= (12)K = \{1,(12)\} & K &= K(12) = \{1,(12)\} \\ (123)K &= (13)K = \{(123),(13)\} & K(123) &= K(23) = \{(123),(23)\} \\ (132)K &= (23)K = \{(132),(23)\} & K(132) &= K(13) = \{(132),(13)\} \end{aligned}$$

Thus, it is possible that $gK \neq Kg$. This is actually a very frequent occurrence in nonabelian groups.

Theorem 5.1.4. *Let $H \leq G$. We define a relation on G by the rule $x \sim y$ if and only if $x^{-1}y \in H$. Then this is an equivalence relation where the equivalence classes are the left cosets gH. In particular, the left cosets form a partition on the group G.*

Proof. Let $x \in G$. Then $x^{-1}x = e \in H$, since H is a subgroup of G. Thus, $x \sim x$ and $\sim$ is reflexive. Next, let $x \sim y$ for some $x, y \in G$. Then $x^{-1}y \in H$. Since H is a subgroup, $y^{-1}x = (x^{-1}y)^{-1} \in H$. Thus, $y \sim x$ and $\sim$ is symmetric. Finally, let $x \sim y$ and $y \sim z$ for some $x, y, z \in H$. Then $x^{-1}y, y^{-1}z \in H$. Since H is a subgroup, $x^{-1}z = (x^{-1}y)(y^{-1}z) \in H$. Thus, $x \sim z$ and $\sim$ is transitive. Therefore, $\sim$ is an equivalence relation on G.

We next will show that $[g] = gH$. Let $x \in gH$. Then there exists some $h \in H$ such that $x = gh$. Then $g^{-1}x = g^{-1}(gh) = h \in H$, which shows that $x \in [g]$. Thus, $gH \subseteq [g]$. The containment $[g] \subseteq gH$ is handled similarly. Therefore, $[g] = gH$, that is, the left cosets are the equivalence classes of the relation $\sim$. □

An analogous statement about right cosets can be made, with similar proof, using the relation $x \sim y$ if and only if $xy^{-1} \in H$.

Theorem 5.1.5. *Let G be a group with $H \leq G$ and $g \in G$. Then $|H| = |gH|$.*

Proof. We define a mapping $\varphi : H \to gH$ by the rule $\varphi(h) = gh$ for all $h \in H$, which is clearly well-defined. We claim that this mapping is bijective. For injectivity, let $\varphi(h) = \varphi(h')$ for $h, h' \in H$. Then $gh = gh'$. Cancelling g on the left, we get $h = h'$. Thus, φ is injective. For surjectivity, let $x \in gH$. Then there exists some $h \in H$ such that $x = gh$. So, $\varphi(h) = gh = x$. Thus, φ is surjective, proving the claim. Since φ is bijective, we conclude that $|H| = |gH|$. □

In particular, all left cosets have the same cardinality as H and have the same cardinality as each other. A similar proof extends the analogous statement for right cosets.

Theorem 5.1.6. *Let H be a subgroup of a group G. Then $|G/H| = |H\backslash G|$.*

Proof. Like the previous proof, it suffices to construct a bijection between the two sets. Let $\varphi : G/H \to H\backslash G$ be defined by the rule $\varphi(gH) = Hg^{-1}$. Now, we need to check that this definition of φ is well-defined since it is defined by a representative of gH, that is, we need to check that different representatives of gH give the same image. Suppose that $aH = bH$. We claim that $Ha^{-1} = Hb^{-1}$. Let $x \in Ha^{-1}$. So there exists some $h \in H$ such that $x = ha^{-1}$. Then $x^{-1} = (ha^{-1})^{-1} = ah^{-1} \in aH$. Since $aH = bH$, this implies that there is some $k \in H$ such that $x^{-1} = bk$. Then $x = (bk)^{-1} = k^{-1}b^{-1} \in Hb^{-1}$. Therefore, $Ha^{-1} \subseteq Hb^{-1}$. The containment $Hb^{-1} \subseteq Ha^{-1}$ is handled similarly. Since we have established $Ha^{-1} = Hb^{-1}$ if $aH = bH$, we have proven that φ is well-defined.

Define $\psi : H\backslash G \to G/H$ by the rule $\psi(Hg) = g^{-1}H$. By similarly reasoning, this mapping is also well-defined and is the inverse of φ. Since φ is invertible, we conclude that φ is bijective. □

Proposition 5.1.7. *Let H be a subgroup of a group G and let $x, y \in G$. Then $xH = yH$ if and only if $y \in xH$.*

The proof is left as an exercise. As usual, as there is nothing particularly special about left cosets, the analogous result holds for right cosets also.

[i] See §6.1 Cosets in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–8, list all the left and right cosets of the given subgroup inside the given group.

1. $\langle 8 \rangle \leq \mathbb{Z}_{24}$
2. $\langle 3 \rangle \leq \mathbb{Z}_8^*$
3. $3\mathbb{Z} \leq \mathbb{Z}$
4. $A_4 \leq S_4$
5. $A_n \leq S_n$
6. $D_4 \leq S_4$
7. $S^1 \leq \mathbb{C}^*$
8. $\langle (123) \rangle \leq S_4$
9. In the proof of Theorem 5.1.6, we defined a function $\varphi : G/H \to H\backslash G$ be defined by the rule $\varphi(gH) = Hg^{-1}$. We then argued this is a well-defined bijection. Why could we not use instead $\varphi(gH) = Hg$?

For Exercises 10–12, prove the given statement about cosets.

10. Let $H \leq G$ and suppose $g \in G$ such that $ghg^{-1} \in H$ for all $h \in H$. Then $gH = Hg$.
11. Let $H, K \leq G$. Then, for any $g \in G$, $gH \cap gK$ is a left coset of $H \cap K$ inside of G.
12. Let $H \leq G$ and $x, y \in G$. The following are equivalent:

    (a) $xH = yH$ (b) $Hx^{-1} = Hy^{-1}$ (c) $xH \subseteq yH$

    (d) $y \in xH$ (e) $x^{-1}y \in H$

"Think left and think right and think low and think high. Oh, the thinks you can think up if only you try!"
– Dr. Seuss

Lecture Videos

The Index of a Subgroup | Lagrange's Theorem | Euler's Theorem and Fermat's Little Theorem

5.2 Lagrange's Theorem

Definition 5.2.1. Let G be a group with subgroup H. Then we define the **index** of H in G, denoted $[G : H]$, as the number of left cosets of H in G, that is, $[G : H] = |G/H|$.

Since $|G/H| = |H\backslash G|$, we could equally have defined the index as the number of right cosets.

Example 5.2.2. Let $G = \mathbb{Z}_6$ and $H = \langle 3 \rangle = \{0, 3\}$. Then $[G : H] = 3$ since there were three (left) cosets H, $1 + H$, and $2 + H$.

Example 5.2.3. Let $G = S_3$, $H = A_3 = \{1, (123), (132)\}$, and $K = \{1, (12)\}$. Then $[G : H] = 2$, since H has two left cosets: H and $(12)H$, and $[G : K] = 3$, since K has three left cosets K, $(123)K$, and $(132)K$.

Theorem 5.2.4 (Lagrange's Theorem)**.** *Let G be a finite group and let $H \leq G$. Then $[G : H] = |G|/|H|$. In particular, the order of H divides the order of G.*

Proof. Since G is partitioned by the left cosets of G, two cosets are either equal or disjoint. In particular, G is the disjoint union of all the distinct cosets of H. So, $|G|$ is the sum of each distinct coset's cardinality, which is always just $|H|$. As there are $[G : H]$ many cosets, we get $|G| = [G : H] \cdot |H|$, which proves the result. □

Corollary 5.2.5. *Suppose that G is a finite group and $g \in G$. Then $|g|$ must divide $|G|$. In particular, $g^{|G|} = e$.*

Proof. Let $|G| = n$ and $|g| = m$. The cyclic subgroup $\langle g \rangle$ is clearly a subgroup of G and $|\langle g \rangle| = m$. By Lagrange's Theorem, $m \mid n$, that is, there exists some $k \in \mathbb{Z}$ such that $km = n$. Then $g^n = g^{km} = (g^m)^k = e^k = e$. □

Corollary 5.2.6. *Let $|G| = p$ a prime. Then G is cyclic and any $g \in G$ such that $g \neq e$ is a generator of G.*

Proof. Let $g \in G$. Let $|g| = n$. By Lagrange's Theorem, $n \mid p$, but since p is prime either $n = 1$ or $n = p$. If $n = 1$, then $g = e$. If $n = p = |G|$, then $\langle g \rangle = G$. □

Corollary 5.2.7. *Let H and K be subgroups of a finite group G such that $K \subseteq H \subseteq G$. Then*

$$[G : K] = [G : H][H : K].$$

Proof. Note that

$$[G : K] = \frac{|G|}{|K|} = \frac{|G|}{|K|}\left(\frac{|H|}{|H|}\right) = \frac{|G|}{|H|}\left(\frac{|H|}{|K|}\right) = [G : H][H : K].$$ □

The previous result is also true for infinite groups with subgroups of finite index but much more involved to prove.

It is important to recognize that the converse of Lagrange's Theorem is not true in general, that is, if $d \mid |G|$ then there is not necessarily a subgroup of order d. For example, $|A_4| = 24$ and $6 \mid 24$ but A_4 has no subgroup of order 6. To see this, assume to the contrary that $H \leq A_4$ and $|H| = 6$. Clearly, $1 \in H$. There are only three 2-2 cycles in A_4. Thus, H must contain a 3-cycle and its inverse. As these come in inverse pairs, H must also contain a 2-2 cycle. But any 3-cycle and 2-2-cycle generate all of A_4 together, which can be proven. Therefore, A_4 has no subgroup of order 6.

The **Euler ϕ-function**, also called the **totient function**, is a map $\phi : \mathbb{Z}^+ \to \mathbb{R}$ such that $\phi(1) = 1$ and $\phi(n)$ is equal to the number of positive integers coprime to n for $n > 1$. In particular, we have shown that $|\mathbb{Z}_n^*| = \phi(n)$. The following two important theorems from number theory are more corollaries of Lagrange's Theorem.

Theorem 5.2.8 (Euler's Theorem)**.** *Let a and n be integers such that $n > 0$ and $\gcd(a, n) = 1$. Then*

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

Proof. The result follows from Corollary 5.2.5 since $a \in \mathbb{Z}_n^*$ and $|\mathbb{Z}_n^*| = \phi(n)$. □

Theorem 5.2.9 (Fermat's Little Theorem)**.** *Let p be prime. If $p \nmid a$, then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Furthermore, $b^p \equiv b \pmod{p}$ for any integer b.

Proof. Since $\phi(p) = p - 1$, the first equation follows immediately from Euler's Theorem. As for the second equation, if $p \nmid b$, then $b^{p-1} \equiv 1 \pmod{p}$. Multiplying both sides by b then gives the identity. If $p \mid b$, then $b \equiv 0 \pmod{p}$ and $b^p \equiv 0^p \equiv 0 \equiv b \pmod{p}$. □

Fermat's Little Theorem leads to a very basic primality test in number theory as illustrated in Example 5.2.10. If we are uncertain if n is prime, we can choose some positive integer $a < n$. We first can compute $d = \gcd(a, n)$ (using the Euclidean algorithm). If $d \neq 1$, then $d \mid n$ and n is not prime. If $d = 1$, compute $a^{n-1} \pmod{n}$ (using the Repeated Squares algorithm). If this value is not congruent to 1, then we can conclude that n is not prime even though you do not necessarily know a factorization of n. If $a^{n-1} \equiv 1 \pmod{n}$) for a "large" list of integers a, then we can conclude that n is probably prime.

[i] See §6.1 Cosets, §6.2 Lagrange's Theorem, and §6.3 Fermat's and Euler's Theorems in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Suppose that $|G| = 60$ is a group. What are the possible orders of subgroups of G?

   For Exercises 2–4, prove the given statement about orders and indices.

2. Suppose that G is a finite group with $g, h \in G$ such that $|g| = 5$ and $|h| = 7$. Then $|G| \geq 35$.
3. If $p = 4n + 3$ is prime, then there is no solution to $x^2 \equiv -1 \pmod{p}$.
4. As additive groups, $[\mathbb{Q} : \mathbb{Z}] = \infty$.

Example 5.2.10 (Fermat's Little Primality Test)**.** By Fermat's Little Theorem, if p is prime and $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$. Note that by this identity 15 cannot be a prime since

$$2^{15-1} \equiv 2^1 4 \equiv 4 \not\equiv 1 \pmod{15}.$$

Note that we can observe that 15 is composite without determining its factorization (of course, we know $15 = 3(5)$). Conversely,

$$2^{17-1} \equiv 2^{16} \equiv 1 \pmod{17},$$

so by Fermat's Little Theorem, 17 could be prime (which we know it is).

Definition 5.2.11. Let $n \in \mathbb{Z}$. We say that n is a **pseudoprime** if $2^{n-1} \equiv 1 \pmod{n}$ but n is composite.

Hence, pseudoprimes appear to be prime from Fermat's primality test but actuarlly are not. Hence, *Fermat's Little Test* can determine whether a number is *certainly composite* (such as 15) or *maybe prime* (such as 17). It cannot, by itself, determine whether a number is prime but it can determine when a number is composite but provides no clue to its factorization.

For Exercises 5–10, for the given n, compute $2^{n-1} \pmod{n}$. Determine whether n is certainly composite or maybe prime. MAGMA is recommended for these exercises.

5. 342 6. 811 7. 601 8. 561 9. 771 10. 631

Example 5.2.12. The primality test of Example 5.2.10 was conducted using $a = 2$, but any base a such that $\gcd(a, n) = 1$ would work (note that if $d = \gcd(a, n) > 1$, then we know that n is composite since $d \mid n$). Consider $n = 341$. Note that

$$2^{341-1} \equiv 2^{340} \equiv 1 \pmod{341},$$

but

$$3^{341-1} \equiv 3^{340} \equiv 56 \pmod{341}.$$

So, while Fermat's Little Test using $a = 2$ determined that 341 is maybe prime, using $a = 3$ gives that 341 is composite. Thus, 341 is a pseudoprime.

Definition 5.2.13. Let $n, a \in \mathbb{Z}$. We say that n is a **pseudoprime base** a if $\gcd(a, n) = 1$ and $a^{n-1} \equiv 1 \pmod{n}$ but n is composite.

Pseudoprimes of various bases do exist, but if n is determined to be composite by some base a, regardless if it is pseudoprime for many other bases, then n is not prime. Hence, Fermat's Little Test can be greatly improved by testing various bases a. If n again and again is determined to be maybe prime for several different bases, then we can conclude that n is probably prime. Unfortunately,

this test, with several different bases, cannot determine whether n is certianly prime because of **Carmichael numbers**, that is, numbers which are pseudoprime of base a for all integers a coprime to n. For example, $n = 561$ is a Carmichael number. Clearly, $561 = 3(11)(17)$, but if you compute $a^{560} \pmod{561}$, then you always get 1 if $3, 11, 17 \nmid a$. Carmichael numbers with no small prime could trick this test, but fortunately Carmichael numbers are rare among the integers even though there are infinitely many of them.

For Exercises 11–16, for the given n, compute $3^{n-1} \pmod{n}$. Comparing results from Exercises 5–10, determine whether n is certainly composite or probably prime. MAGMA is recommended for these exercises.

11. 342 12. 811 13. 601 14. 561 15. 771 16. 631

5.3 Supplemenatry Exercises

(Go to Solutions)

1. Describe the left cosets of $\mathrm{SL}_2(\mathbb{R}) \leq \mathrm{GL}_2(\mathbb{R})$. Compute $[\mathrm{GL}_2(\mathbb{R}) : \mathrm{SL}_2(\mathbb{R})]$.
2. Verify Euler's Theorem for $n = 15$ and $a = 4$.
3. Suppose that $[G : H] = 2$. If $a, b \in G \setminus H$, show that $ab \in H$.
4. Suppose that $[G : H] = 2$. Show that $gH = Hg$ for all $g \in G$.
5. Let G be a cyclic group of order n. Show that there are $\phi(n)$ generators of G.
6. Let $n = \prod_{i=1}^{k} p_i^{e_i}$, where $p_1, \ldots, p_k$ are distince primes. Prove that
$$\phi(n) = n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right).$$
7. For all $n \in \mathbb{Z}^+$, prove $n = \sum_{d|n} \phi(d)$.
8. For $H, K \leq G$, define $a \sim b$ if there exists some $h \in H$ and $k \in K$ such that $hak = b$, for $a, b \in G$. Show that $\sim$ is an equivalence relation on G. The equivalence classes associated to $\sim$ are called **double cosets** relative to H and K. Compute the double cosets $H = K = \langle (123) \rangle \leq A_4$.

   For Exercises 9–12, prove or disprove the given statement about cosets, orders, or indices.

9. If $g^n = e$, then $|g| \mid n$.
10. Every subgroup of the integers has finite index.
11. Every subgroup of the integers has finite order.
12. As additive groups, $[\mathbb{C} : \mathbb{R}] = \infty$.

Chapter 6

Applied Algebra

"There are no secrets that time does not reveal." – Jean Racine

Lecture Videos

Symmetric Key Cryptography

Caesar Cipher

Affine Cipher

6.1 Symmetric Key Cryptography

When people need to secretly store or communicate messages, they turn to cryptography. Cryptography involves using techniques to obscure a message so outsiders cannot read the message. It is typically split into two steps: encryption, in which the message is obscured, and decryption, in which the original message is recovered from the obscured form.

Definition 6.1.1. A **cipher** replaces each letter (or number) in the message with a different letter, following some established mapping, called the **encryption mapping**. The pre-image of a message with respect to the encryption map is called the **plaintext**. The image of a message with respect to the encryption map is called the **ciphertext**. The mapping is necessarily bijective so ciphertexts can be converted back to the plaintext. This inverse mapping is called the **decryption mapping**.

The strength of any cryptosystem depends on the difficulty of an eavesdropper (Eve) from discovery the decryption map. Although much (if not all) of the encryption mapping is public information, some information about the decryption mapping is kept secret, called the **key**, which without it the decryption mapping is too difficult to compute.

Symmetric-Key Cryptography is a method of cryptography where both the sender (Alice) and the receiver (Bob) use the same key for encryption and decryption.

The mappings commonly used in cryptography are algebraic and require the message to numerical. As such, we will use the simple rule $A = 0$, $B = 1$, $C = 2$, etc. to digitize a string of alphabetic characters as an array of numerical digits for cryptographic purposes.

Example 6.1.2. Below is MAGMA code[i] that can be used to implement the above encoding process. For simplicity, lower and upper case will be considered identical. Also, no numerical digits or special characters will be allowed.

```
Digitize := function(message);
E := [];
m := 0;
for i:= 1 to #message do
    case message[i]:
    when "a","A": E cat:=[0]; when "b","B": E cat:=[1]; when "c","C": E cat:=[2];
    when "d","D": E cat:=[3]; when "e","E": E cat:=[4]; when "f","F": E cat:=[5];
    when "g","G": E cat:=[6]; when "h","H": E cat:=[7]; when "i","I": E cat:=[8];
    when "j","J": E cat:=[9]; when "k","K": E cat:=[10]; when "l","L": E cat:=[11];
    when "m","M": E cat:=[12]; when "n","N": E cat:=[13]; when "o","O": E cat:=[14];
    when "p","P": E cat:=[15]; when "q","Q": E cat:=[16]; when "r","R": E cat:=[17];
    when "s","S": E cat:=[18]; when "t","T": E cat:=[19]; when "u","U": E cat:=[20];
    when "v","V": E cat:=[21]; when "w","W": E cat:=[22]; when "x","X": E cat:=[23];
    when "y","Y": E cat:=[24]; when "z","Z": E cat:=[25];
    end case;
end for;
return E;
end function;
```

For example, the message `"Cryptography"` is digitized as `[ 2, 17, 24, 15, 19, 14, 6, 17, 0,`

`15, 7, 24 ].`

To implement this in MAGMA, first copy and paste the above code verbatim. The previous code will be required each time you want to run `Digitize` since the MAGMA website will not store this code for you. Then follow it on the next line with: `Digitize("Cryptography");` If you want to store this information for future reference, you can store as a variable such as `E` use the line instead `E := Digitize("Cryptography");` In this context, MAGMA will not automatically output the value of function `Digitize`, so follow this line with `E;`

Example 6.1.3. Decoding an array of digits is also required. Below is MAGMA code that can be used to implement the above decoding process.

```
Alphabetize := function(code);
D := "";
m := 0;
for i:= 1 to #code do
    case code[i]:
    when 0: D cat:="A"; when 1: D cat:="B"; when 2: D cat:="C";
    when 3: D cat:="D"; when 4 : D cat:="E"; when 5: D cat:="F";
    when 6: D cat:="G"; when 7: D cat:="H"; when 8: D cat:="I";
    when 9: D cat:="J"; when 10: D cat:="K"; when 11: D cat:="L";
    when 12: D cat:="M"; when 13: D cat:="N"; when 14: D cat:="O";
    when 15: D cat:="P"; when 16: D cat:="Q"; when 17: D cat:="R";
    when 18: D cat:="S"; when 19: D cat:="T"; when 20: D cat:="U";
    when 21: D cat:="V"; when 22: D cat:="W"; when 23: D cat:="X";
    when 24: D cat:="Y"; when 25: D cat:="Z";
    end case;
end for;
return D;
end function;
```

For example, the array `[ 12, 8, 18, 18, 4, 11, 3, 8, 13, 4 ]` is alphabetized as `"MISSELDINE"`.

Like the encoding process, first copy and paste the above code verbatim and follow it with `Alphabetize([ 12, 8, 18, 18, 4, 11, 3, 8, 13, 4 ]);` Again, it might be useful to store this as a variable for future reference.

A simple example of a cipher is called the **Caesar cipher**, named after Julius Caesar who used this method of encryption in ancient Rome. In this approach, each letter is replaced with a letter some fixed number of positions later in the alphabet.[ii] This fixed number is the key of a Caesar cipher. When at the end of the alphabet, the letters will wrap around to the beginning. In particular, the Caesar cipher works addition modulo 26!

Example 6.1.4. Alice wants to use the Caesar cipher with shift of 3 (the actual value used by Caesar) to encrypt the message: `"We ride at noon"`. She uses the MAGMA code below for encryption, but she needs to make sure the message is digitized first. Making sure that the code for `Digitize`, `Alphabetize`, and `CaesarEncrypt` are copied above, she runs the code `CaesarEncrypt("We ride at noon", 1, 3);` can gets the encrypted message `"ZH ULGH DW QRRQ"`.

```
CaesarEncrypt := function(plaintext, a, b);
P := Digitize(plaintext);
C := [];
for i := 1 to #P do
    C cat:= [(a*P[i]+b) mod 26];
end for;
return Alphabetize(C);
end function;
```

Alice is worried that the spacing my give clues about the plaintext, such as the presence of 1-letter

or 2-letter words. Using standard blocks makes it more protected from attacks, such as putting ciphertext in blocks of 4 letters: `"ZHUL GHDW QRRQ"`.

Also is also worried that Bob might have a hard time reading the message without spacing since MAGMA will ignore it, so she changes the message to `"WEXRIDEXATXNOON"`. As `"X"` is an uncommon letter, Bob will know to ignore those letters. Such a letter is called a **null character**. Null characters could also be appended to the end of a message to make all the blocks have the same length. Since the length of the plaintext is one less than a multiple of 4, Alice addes the null character `"X"` at the end to complete the last block, that is, `"WEXRIDEXATXNOONX"`. The method of adding null characters, removing spacing, and otherwise preparing the plaintext to be encrypted is called **padding**. The padded ciphertext is then `"ZHAU LGHA DWAQ RRQA"`.

Example 6.1.5. Later Bob receives the encrypted message `"DOJHEUD"` from Alice. He recognizes that Alice encrypted it using their agreed Caesar shift with key $b = 3$. He runs `CaesarDecrypt("DOJHEUD", 1, 3);` to decrypt the message. We read the plaintext `"ALGEBRA"`. He used the MAGMA code for `CaesarDecrypt` below.

```
CaesarDecrypt := function(ciphertext, a, b);
C := Digitize(ciphertext);
P := [];
ainv := InverseMod(a,26);
binv := ainv*b;
for i := 1 to #C do
    P cat:= [(ainv*C[i]-binv) mod 26];
end for;
return Alphabetize(P);
end function;
```

For the Caesar ciphers that we have considered above, our encryption mapping is given as $x \mapsto x + b \pmod{n}$ and decryption mapping $y \mapsto y - b \pmod{n}$.

Symmetry-key cryptography has the benefit of being easier to use compared to other cryptosystems but has the weakness that both parties must keep the key secure. The more parties involved in the communication the harder it will be to keep the secret key safe.

If an encryption process is compromised, for example if an enemy captures an officer and learns the details of the encryption through interrogation (and torture, yikes!) all past encrypted messages are vulnerable to be read. All future messages can be protected by adjusting the encryption process. This is best done by switching the encryption key. Switching an encryption method hastily is impractical since the encryption method will likely be part of the communication infrastructure. Typically, a new key can easily be created and distributed. For this reason, the security of an reliable cryptosystem must not rest on the secret of the encryption method but instead on the secret of the encryption key.

For example, with the Caesar shift, the encryption method has only 26 encryption keys (although the shift by 26 is rather a silly encryption). Thus, if an enemy receives the encrypted message and knows it was encrypted using a Caesar cipher, then the enemy need only try all 26 possibilities, which is a trivial feat for a computer and a moderate feat for a person. This is called a **brute force attack**. The encryption can be improved by increasing the modulus 26 to a much greater number n by adapting the `Digitize` function, but this only improves the security marginally. Therefore, a Caesar cipher is a weak encryption method.

An improvement of the Caesar cipher is an **affine cipher** with encryption mapping $x \mapsto ax + b \pmod{n}$ and decryption $y \mapsto a^{-1}y - a^{-1}b \pmod{n}$. Of course, $\gcd(a, n) = 1$.

Example 6.1.6. Using $a = 5$ and $b = 3$ Alice encrypts the message `"ALGEBRA"` as `"DGHXIKD"` using the code `CaesarEncrypt("ALGEBRA", 5, 3);` She sends the message to Bob who decrypts it using

the code `CaesarDecrypt("DGHXIKD",5,3);`

For a large modulus, this is a great improvement over the Caesar cipher because the number of possible keys can be too large to brute force. This is an example of a **monoalphabetic encoding**, that is, each individual character is encoded with an individual character. The Caesar cipher is an example of a monoalphabetic encryption, as each letter in the message is replaced with an individual cipher letter. Unfortunately, monoalphabetic ciphers are considered weak cryptosystems for the following reason, despite having 26! possible keys (too big to brute force). Since different characters in a language appear with different frequencies, for example in English 12.702% of all letters used is an `"e"`, while `"x"` comprising only 0.150% of all letters used, using these frequencies and a little bit of guess-and-check, a computer can easily decrypt any substitution cipher. Our `Digitize` function can be improved to try to avoid these frequencies such as using a number to represent a pair or triple of numbers, putting the letters into blocks. The usage of blocks does improve the security from a frequency attack and also allows us to use large moduli, but it is still not immune. For example, `"th"` shows up frequently but `"qx"` hardly ever. Fortunately, there do exist much more sophisticated digitization processes to protect the type of ciphers discussed above but are beyond the scope of our class.

[ii] MAGMA can be accessed at http://magma.maths.usyd.edu.au/calc/. Please also see Appendix B for further details. The MAGMA code for this section can be accessed at https://github.com/emisseldine/Math4220/blob/main/Symmetric_Key_Cryptography.

[ii] This same encryption method is used in the holiday classic "A Christmas Story" when Ralphie uses his Little Orphan Annie Decoder Ring to decrypt the message `"Be Sure to Drink Your Ovaltine"`, although in this story the encrypted message was with numbers and the decrypted message was with letters.

[iii] See §7.1 Private Key Cryptography in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–2, use the affine cipher implemented in the MAGMA code found here:
https://github.com/emisseldine/Math4220/blob/main/Symmetric_Key_Cryptography.

1. Encrypt `"IxLOVExMATH"`.
2. Decrypt `"ZLOOA WKLVA EHARQ WKHA ILQDO"`.

"The man who can keep a secret may be wise, but he is not half as wise as the man with no secrets to keep." – E. W. Howe

Lecture Videos

Diffie-Hellman Key Exchange

RSA

6.2 Public Key Cryptography

Suppose that you are connecting to your bank's website. It is possible that someone could intercept any communication between you and your bank, so you will want to encrypt the communication. Even though symmetric key cryptography can be modified to make it much more secure than the model we used, the problem is that symmetric key encryption methods require that both parties have already agreed on a shared secret encryption key. How can you and your bank agree on a key if you have not already?

This becomes the goal of public key cryptography – to provide a way for two parties to agree on a key without a snooping third party being able to determine the key. The method relies on a one-way function; something that is easy to do one way, but hard to reverse.

To explain, consider the following analogy. Alice wants to send Bob, who lived far away, a very valuable treasure. Since she is not able to meet Bob in person to deliver the treasure, she hires a courier service to deliver the treasure to Bob. Not sure if she can trust the courier or not sure if he will try to steal the treasure in transit, Alice places the treasure in a box and places a lock on the box for which she only has a key to open.

The treasure is then sent to Bob. He is unable to open the box to get the treasure since it was locked by Alice and he does not have a key. Instead of unlocking the box, Bob places a second lock on the box for which only he has a key to. Bob then sends the box back to Alice.

When the box returns to Alice with now two locks on it, one from Alice and one from Bob, Alice removes her lock from the box. After which, Alice sends the box back to Bob. When Bob receives the box a second time, it only has one lock on it for which he has the key. Bob then removes the final lock and receives the valuable treasure from Alice. Although the treasure bounced back and forth between Alice and Bob, Alice was able to securely send the treasure to Bob since at all times the box was locked by a lock belonging to Alice or Bob. This method is known as the Diffie-Hellman Key Exchange and has been applied using group theory as the first public key cryptosystem. It is called a key exchange because Alice and Bob will exchange a common key to use in future symmetric key cryptosystems.

If we return to the narrative that Alice wants to share a great treasure with Bob that she sends to Bob in a locked treasure box, the back-and-forth nature of the Diffie-Hellman exchange can be very slow. Is there a faster way?

Also, what if the courier has bad intentions and instead of bringing the box to Bob for him to put a lock on the treasure box the courier himself places a lock on the treasure box, claiming it is Bob's lock, and returns the box to Alice. Thinking that Bob's lock is on the box, she removes her lock leaving the courier's lock on the box only. Upon leaving with the box, the courier will be able to unlock the box and steal the treasure. Is there some way to guarantee Bob's authenticity before Alice removes her lock? Enter RSA.

To explain RSA (Rivest, Adelman, Shamir) simply, instead of Alice locking her box and sending it to Bob for him to place a lock on it, Alice will use an open lock of Bob's that he left with Alice the last time they were together. She places the lock on the chest and the courier brings the locked chest to Bob, the only one able to open Bob's lock.

In this public-key cryptosystem, Bob shares with Alice (and the whole world) his open locks (his public key) and she can use as many as she needs. Then when she sends a message to Bob locked with his lock, only his private key is capable of unlocking his locks.

To implement RSA in $\mathbb{Z}_n$, Alice must choose a public key e and private key d so that $m^{de} \equiv m \pmod{n}$. If Bob wants to send a secure plaintext m to Alice, he encrypts m by the ciphertext $S \equiv m^e \pmod{n}$. He then sends S to Alice. To decrypt, Alice computes $S^d \equiv (m^e)^d \equiv m^{de} \equiv m \pmod{n}$, the plaintext message from Bob. Since e is public knowledge, anybody can send a secure message to Alice. Since d is secret knowledge, only Alice can decrypt her messages.

To create a public and private key for RSA, Alice selects an integer e such that $\gcd(e, \phi(n)) = 1$. Then using the Euclidean algorithm, or the MAGMA function `InverseMod(e, phi)`, Alice computes her private key d. Since $de + k\phi(n) = 1$ for some $k \in \mathbb{Z}$, Alice has that

$$x^{de} \equiv x^{1-k\phi(n)} \equiv x(x^{\phi(n)})^{-k} \equiv x(1)^{-k} \equiv x \pmod{n},$$

by Euler's Theorem. She then publishes her public key (e, n) and secures her private key $(d, \phi(n))$.

Since Eve knows e and n, can she compute d? Of course, if she knows $\phi(n)$, then yes, she can do it in the same way that Alice computed d. Without $\phi(n)$, it is very hard to compute d other than brute force. So, the security of RSA rests on the difficulty of computing $\phi(n)$. The totient function is **multiplicative**, meaning that $\phi(1) = 1$ and if $\gcd(a, b) = 1$ then $\phi(ab) = \phi(a)\phi(b)$. Essentially, $\phi(n)$ is easy to compute if one has a prime factorization of n and nearly impossible otherwise–other than by brute force. Therefore, the security of RSA rests upon the difficulty of factoring large integers, which is hard to do even for computers.[i]

To create a secure $\phi(n)$, Alice will pick two large primes p and q. In this class, we will pick primes with 45 digits or more. Then $n = pq$, called a **semiprime**. Then

$$\phi(n) = \phi(pq) = \phi(p)\phi(q) = (p-1)(q-1).$$

Thus, Alice can compute $\phi(n)$ and subsequently compute d given e and $n = pq$. Alice's secret essentially then is the factoring $n = pq$. It is important that she keeps her primes secret.

Example 6.2.1. Suppose that Alice has computed $n = 3127 = 53 \cdot 59$, $e = 3$, and $d = 2011$. Show how Bob would encrypt the message $m = 50$ and how Alice would then decrypt it.

Bob would only know his message, $m = 50$ and Alice's public key, $n = 3127$ and $e = 3$. He would encrypt the message by computing

$$m^e \equiv 50^3 \equiv 3047 \pmod{3127}.$$

Alice can then decrypt this message using her private key d which she already computed by the Euclidean algorithm: note that $\phi(3127) = (53-1)(59-1) = 52(58) = 3016$ and

$$\begin{aligned} 3016 &= 1005 \cdot 3 + 1 \\ 1 &= 3016 - (1005)3 \end{aligned}$$

Thus, $d \equiv -1005 \equiv 2011 \pmod{3016}$. Then

$$S^d \equiv S^{2011} \equiv 3047^{2011} \equiv 50 \pmod{3127}.$$

Example 6.2.2. This time Alice decides to use the primes

```
p := 581443228694704049027262871;
q := 422739459734865292638810361;
```

Alice computed these by using the code

```
p := NextPrime(Random(10^25, 10^27)); p;
q := NextPrime(Random(10^25, 10^27)); q;
```

which chooses a random prime between 25 to 27 digits. Then

```
n := p*q; n;
phi := (p-1)*(q-1); phi;
```

which gibes

$$n = pq = 245798996364894914258483011002798087005440825465406431$$

and

$$\phi = (p-1)(q-1) = 24579899636489491425848300999861539857587148379933200.$$

Then she choose $e = 19$ for her public key. For her private key, she computes

```
e := 19;
d := InverseMod(e, phi); d;
```

giving the value $d = 25873578564725780448261369473538463007986471978877179$.

Bob encrypts the message `"Two if by sea"` using `EncryptRSA("Two if by sea", e, N)` and gets `[ 20098610805067933171020340071756787863136016934669840 9 ]`, call it S, which he sends to Alice.

To decrypt, Alice runs `DecryptRSA(S, d, N)`, which gives her the message `"Two if by sea"`.

The code for all of these MAGMA functions is included at the end of this section.

RSA differs from Diffie-Hellman because no exchange process is needed; Bob could send Alice an encrypted message using Bob's public key without having to communicate with Alice beforehand to determine a shared secret key. This is especially handy for applications like encrypting email, where both parties might not be online at the same time to perform a Diffie-Hellman style key exchange. In practice, online secure communications begin with a RSA communication to decide upon a symmetric-key cipher and private key. Then future communications are performed with the symmetric-key cryptosystem. This is because the symmetric-key communication is much faster to encrypt and decrypt. Unfortunately, symmetric-key cryptosystems are less secure than public-key systems, since new symmetric-key must be exchanged via public-key encryption every so often. If no communication occurs between the two servers for a prolonged period of time (maybe 20 minutes) then the symmetric-key encryption will become vulnerable and the computers will be logged out of the secure channel for safety.

[i] MAGMA can be accessed at http://magma.maths.usyd.edu.au/calc/. Please also see Appendix B for further details. The MAGMA code for this section can be accessed at https://github.com/emisseldine/Math4220/blob/main/Public_Key_Cryptography.

[ii] This is to say digital computers. Factoring large numbers on a quantum computer is easy. As such other group-based cryptosystems are needed to retain security in the quantum age.

[iii] See §7.2 Public Key Cryptography in Judson's Abstract Algebra: Theory and Applications for additional reading.

```
PadMessage:= function(message, N);
PM := [];
E := "";
for i:= 1 to #message do
case message[i]:
when "a": E cat:="10"; when "b": E cat:="11"; when "c": E cat:="12"; when "d": E cat:="13";
when "e": E cat:="14";  when "f": E cat:="15"; when "g": E cat:="16"; when "h": E cat:="17";
when "i": E cat:="18"; when "j": E cat:="19"; when "k": E cat:="20"; when "l": E cat:="21";
when "m": E cat:="22"; when "n": E cat:="23"; when "o": E cat:="24"; when "p": E cat:="25";
when "q": E cat:="26"; when "r": E cat:="27"; when "s": E cat:="28"; when "t": E cat:="29";
when "u": E cat:="30"; when "v": E cat:="31"; when "w": E cat:="32"; when "x": E cat:="33";
when "y": E cat:="34"; when "z": E cat:="35"; when "A": E cat:="36"; when "B": E cat:="37";
when "C": E cat:="38"; when "D": E cat:="39"; when "E": E cat:="40"; when "F": E cat:="41";
when "G": E cat:="42"; when "H": E cat:="43"; when "I": E cat:="44"; when "J": E cat:="45";
when "K": E cat:="46"; when "L": E cat:="47"; when "M": E cat:="48"; when "N": E cat:="49";
when "O": E cat:="50"; when "P": E cat:="51"; when "Q": E cat:="52"; when "R": E cat:="53";
when "S": E cat:="54"; when "T": E cat:="55"; when "U": E cat:="56"; when "V": E cat:="57";
when "W": E cat:="58"; when "X": E cat:="59"; when "Y": E cat:="60"; when "Z": E cat:="61";
when " ": E cat:="62"; when ".": E cat:="63"; when ",": E cat:="64"; when "?": E cat:="65";
when "!": E cat:="66"; when "$": E cat:="67"; when "%": E cat:="68"; when "&": E cat:="69";
when "0": E cat:="70"; when "1": E cat:="71"; when "2": E cat:="72"; when "3": E cat:="73";
when "4": E cat:="74"; when "5": E cat:="75"; when "6": E cat:="76"; when "7": E cat:="77";
when "8": E cat:="78"; when "9": E cat:="79"; when "+": E cat:="80"; when "-": E cat:="81";
when "*": E cat:="82"; when "/": E cat:="83"; when "=": E cat:="84"; when "^": E cat:="85";
when "@": E cat:="86"; when "#": E cat:="87"; when "(": E cat:="88"; when ")": E cat:="89";
when ";": E cat:="90"; when ":": E cat:="91"; when "<": E cat:="92"; when ">": E cat:="93";
when "'": E cat:="94"; when "[": E cat:="95"; when "]": E cat:="96"; when "{": E cat:="97";
when "}": E cat:="98"; when "_": E cat:="99";
end case;

if StringToInteger(E, 10) gt N then
    PM cat:= [StringToInteger(Substring(E, 1, #E-2), 10)];
    E := Substring(E, #E-1, 2);
end if;
end for;
PM cat:= [StringToInteger(E, 10)];
return PM;
end function;

DepadMessage := function(code);
D := &cat [IntegerToString(code[i]) : i in [1..#code]];
M := "";

for i:= 1 to #D div 2 do
case Substring(D, 2*i-1, 2):
when "10": M cat:="a"; when "11": M cat:="b"; when "12": M cat:="c"; when "13": M cat:="d";
when "14": M cat:="e"; when "15": M cat:="f"; when "16": M cat:="g"; when "17": M cat:="h";
when "18": M cat:="i"; when "19": M cat:="j"; when "20": M cat:="k"; when "21": M cat:="l";
when "22": M cat:="m"; when "23": M cat:="n"; when "24": M cat:="o"; when "25": M cat:="p";
when "26": M cat:="q"; when "27": M cat:="r"; when "28": M cat:="s"; when "29": M cat:="t";
when "30": M cat:="u"; when "31": M cat:="v"; when "32": M cat:="w"; when "33": M cat:="x";
when "34": M cat:="y"; when "35": M cat:="z"; when "36": M cat:="A"; when "37": M cat:="B";
when "38": M cat:="C"; when "39": M cat:="D"; when "40": M cat:="E"; when "41": M cat:="F";
when "42": M cat:="G"; when "43": M cat:="H"; when "44": M cat:="I"; when "45": M cat:="J";
when "46": M cat:="K"; when "47": M cat:="L"; when "48": M cat:="M"; when "49": M cat:="N";
when "50": M cat:="O"; when "51": M cat:="P"; when "52": M cat:="Q"; when "53": M cat:="R";
when "54": M cat:="S"; when "55": M cat:="T"; when "56": M cat:="U"; when "57": M cat:="V";
when "58": M cat:="W"; when "59": M cat:="X"; when "60": M cat:="Y"; when "61": M cat:="Z";
when "62": M cat:=" "; when "63": M cat:="."; when "64": M cat:=","; when "65": M cat:="?";
when "66": M cat:="!"; when "67": M cat:="$"; when "68": M cat:="%"; when "69": M cat:="&";
when "70": M cat:="0"; when "71": M cat:="1"; when "72": M cat:="2"; when "73": M cat:="3";
when "74": M cat:="4"; when "75": M cat:="5"; when "76": M cat:="6"; when "77": M cat:="7";
when "78": M cat:="8"; when "79": M cat:="9"; when "80": M cat:="+"; when "81": M cat:="-";
when "82": M cat:="*"; when "83": M cat:="/"; when "84": M cat:="="; when "85": M cat:="^";
when "86": M cat:="@"; when "87": M cat:="#"; when "88": M cat:="("; when "89": M cat:=")";
when "90": M cat:=";"; when "91": M cat:=":"; when "92": M cat:="<"; when "93": M cat:=">";
when "94": M cat:="'"; when "95": M cat:="["; when "96": M cat:="]"; when "97": M cat:="{";
when "98": M cat:="}"; when "99": M cat:="_";
end case;
end for;

return M;
```

```
end function;

ExpRSA := function(plaintext, e, N);
ciphertext := [];

for i:=1 to #plaintext do
    ciphertext cat:= [Modexp(plaintext[i], e, N)];
end for;
return ciphertext;
end function;

EncryptRSA := function(plaintext, e, N);
return ExpRSA(PadMessage(plaintext, N), e, N);
end function;

DecryptRSA := function(ciphertext, d, N);
return DepadMessage(ExpRSA(ciphertext, d, N));
end function;
```

Exercises

(Go to Solutions)

For Exercises 1–4, encrypt the plaintext m using the RSA public key (n, e).

1. 31, $(3551, 629)$
2. 23, $(2257, 47)$
3. 142371, $(120979, 13251)$
4. 231561, $(45629, 781)$

For Exercises 5–8, decrypt the ciphertext S using the RSA private key (n, d).

5. 2791, $(3551, 1997)$
6. 34, $(5893, 81)$
7. 112135, $(120979, 27331)$
8. 129381, $(79403, 671)$

For Exercises 9–16, determine the RSA private key d for the given RSA public key (n, e). In MAGMA, you may use the command `Factorization(n)`.

9. $(3551, 629)$
10. $(2257, 47)$
11. $(120979, 13251)$
12. $(45629, 781)$
13. $(451, 231)$
14. $(3053, 1921)$
15. $(37986733, 12371)$
16. $(16394854313, 34578451)$

17. Use the RSA cipher implemented in the MAGMA code found here:
https://github.com/emisseldine/Math4220/blob/main/Public_Key_Cryptography
to encrypt the plaintext `"Abstract Algebra is the best!"`“” using $e = 17$ and

$$n = 1030939719288293437227911877452767233915886422910927826089912725109886136705892109592272121542938809416816702749278385596820061415954841542352923327267600354767.$$

"From the errors of others, a wise man corrects his own." – Publilius Syrus

Lecture Videos

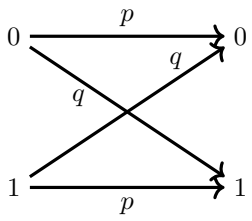
Binary Symmetric Channel

(n,m)-Block Codes

6.3 Algebraic Coding Theory

Digital communications are often transmitted as a sequence of binary numbers (1's and 0's), called a *bit*. Unfortunately, during transmission sometimes "noise" can interfere with the original message and alter the bits in the message. As the significance of each bit is important, a signal error in a bit can make the transmitted message useless or dangerous. As such, it will be important to be able to detect and, if possible, correct erroneous transmission. Group theory will be a valuable tool for this.

A **binary symmetric channel** T is a model that consists of a transmitter capable of sending a binary signal together with a receiver. Let p be the probability that a bit is transmitted correctly. Let $q = 1-p$ then be the probability that a bit is transmitted incorrectly. This model is an example of a **Bernoulli trial** (or **binomial trial**) from Probability Theory, which which is a random, independent[i] experiment with two possible outcomes: "success" and "failure."



Let X be the random variable which counts the number of failures. Then by standard probability reasoning, the probability that k errors will occur if n bits are transmitted through T is

$$\mathcal{P}(X = k) = \binom{n}{k} p^{n-k} q^k.$$

Example 6.3.1. Suppose that $p = 0.995$ and a 500-bit message is sent. The probability that the message is error-free ($k = 0$) is

$$\mathcal{P}(X = 0) = \binom{500}{0} (0.995)^{500} (0.005)^0 \approx 0.08157186$$

Thus, the probability that at least one error will be in the transmission is then $\mathcal{P}(X > 0) = 1 - \mathcal{P}(X = 0) \approx 0.91842814$. Continuing, the probability that exactly one error will occur is

$$\mathcal{P}(X = 1) = \binom{500}{1} (0.995)^{499} (0.005)^1 \approx 0.20495443$$

The probability that exactly two errors will occur is

$$\mathcal{P}(X = 2) = \binom{500}{2} (0.995)^{498} (0.005)^2 \approx 0.25696547$$

In particular, the probability of the message having one or two errors is $\mathcal{P}(1 \leq X \leq 2) = \mathcal{P}(X = 1) + \mathcal{P}(X = 2) \approx 0.46191990$. Furthermore, the probability of having more than two errors in the transmission is $\mathcal{P}(X > 2) = 1 - \mathcal{P}(X = 0) - \mathcal{P}(X = 1) - \mathcal{P}(X = 2) \approx 0.45650824$. This hopefully illustrates the fact that even with a low probability of failure, we must expect frequent errors when transmitting binary messages.

A sequence of m-bits (read left to right) $b_1 b_2 \dots b_m$ can be identified naturally with the k-tuple $(b_1, b_2, \dots, b_m) \in \mathbb{Z}_2^m$.

Definition 6.3.2. An **(n, m)-block code** C is a subset of $\mathbb{Z}_2^n$ paired with two functions $E : \mathbb{Z}_2^m \to \mathbb{Z}_2^n$ and $D : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$, called the **encoding function** and **decoding function** respectively, such that $D \circ E = Id_{\mathbb{Z}_2^m}$ and $E(\mathbb{Z}_2^m) = C$. Necessarily, E must be injective. We call the elements of C **codewords**.

Below illustrates the typical model of communication systems.

Message $w \in \mathbb{Z}_2^m$ $\xmapsto{E}$ Associated code $E(w) = c \in \mathbb{Z}_2^n$ $\xmapsto{T}$ Transmitted code $T(c) \in \mathbb{Z}_2^n$ $\xmapsto{D}$ Decoded message $D(y) \in \mathbb{Z}_2^m$

Ideally, we want $D \circ T \circ E = Id_{\mathbb{Z}_2^m}$, but it will not be possible since T is not a function as $T(c)$ is determined by chance. A **k-error detecting code** is a code which reports an error message when $D \circ T \circ E \neq Id_{\mathbb{Z}_2^m}$ for k or fewer errors. An **ℓ-error correcting code** is a code for which $D \circ T \circ E = Id_{\mathbb{Z}_2^m}$ even in the presence of ℓ or fewer errors.

Example 6.3.3. The ASCII (American Standard Code for Information Interchange) coding system is an $(8, 7)$-block code, yielding $2^8 = 256$ possible codewords (or two hexadecimal digits). However, only seven bits are needed since there are only $2^7 = 128$ ASCII characters. In particular, we need to encode $\mathbb{Z}_2^7$. The eighth bit, placed on the left, is the sum of the other seven bits modulo 2. For example, the letters A, B, C are represented below in decimal, binary, and in ASCII.

Character	Decimal	Binary	ASCII
"A"	65	1000001	0100 0001 (41)
"B"	66	1000010	0100 0010 (42)
"C"	67	1000011	1100 0011 (43)

In particular, the mapping $E : \mathbb{Z}_2^7 \to \mathbb{Z}_2^8$ by the rule $E(b_1, b_2, \ldots, b_7) = (b_1 + b_2 + \ldots + b_7, b_1, b_2, \ldots, b_7)$ is an encoding mapping and the mapping $D : \mathbb{Z}_2^8 \to \mathbb{Z}_2^7$ by the rule $D(c_1, c_2, \ldots, c_8) = (c_2, c_3, \ldots, c_8)$ is the corresponding decoding function. This forms a code which can detect a single error in transmission. To see this, if a codeword had exactly one binary digit toggle during transmission, then the sum of the seven bits will not add up to be the same value as the first bit, called the **parity check bit**. Therefore, the code can detect the change in a single bit. Unfortunately, this **parity checking code** cannot detect the presence of two errors. For example, if $T(00000000) = 00000011$, then the code would accept the transmitted codeword as correct. This parity checking code and its variation are probably the most used type of code.

Example 6.3.4. Consider $\mathbb{Z}_2^m$ for encoding. Then we can construct the **triple repetition code** C with $E : \mathbb{Z}_2^m \to \mathbb{Z}_2^{3m}$ given by $E(b_1, \ldots, b_m) = (b_1, \ldots, b_m, b_1, \ldots, b_m, b_1, \ldots, b_m)$ and $D : \mathbb{Z}_2^{3m} \to \mathbb{Z}_2^m$ given by $D(c_1, \ldots, c_{3m}) = (c_1, \ldots, c_m)$. This $(3m, m)$- block code is, in fact, error correcting, because every bit is essentially sent three times and if there is an error with a bit transmission the code can correct this by outputting the most common value of the three transmission. This error-correcting code still has limits. First, it requires sending three times as much data than the original message. Also, the code cannot detect nor correct if two or more errors occur for the "same" bit.

[i] We assume that errors occurring in different bits are independent of one another, meaning the presence of an error in one bit has no influence on the presense of an error in a different bit.

[ii] See §8.1 Error-Detecting and Correcting Codes in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Why is the following code unacceptable?

Message	0	1	2	3	4	5	6	7	8	9
Codeword	000	001	010	011	100	101	110	111	000	001

For Exercises 2–5, suppose that a 1000-bit message is transmitted across of binary symmetric channel T, with a probably of a single error occuring in a bit is q. Given q, determine the probably of k errors occuring in transmission across T.

2. $q = 0.01$, $k = 2$
3. $q = 0.01$, $k \geq 2$
4. $q = 0.001$, $k = 2$
5. $q = 0.001$, $k \geq 2$

For Exercises 6–11, we discuss UPC symbols, as defined in Definition 6.3.5.

Definition 6.3.5. Universal Product Code (or **UPC**) **symbols** are found on most products in grocery and retail stores (commonly refered to as "bar codes"). The UPC symbol is a 12-digit code identifying the manufacturer of a product and the product itself. The first 11 digits contain information about the product; the twelfth digit is used for error detection. If $d_1 d_2 \dots d_{12}$ is a valid UPC number, then

$$(3, 1, 3, \dots, 1) \cdot (d_1, d_2, d_3, \dots, d_{12}) = 3 \cdot d_1 + 1 \cdot d_2 + 3 \cdot d_3 + \dots + 3 \cdot d_{11} + 1 \cdot d_{12} \equiv 0 \pmod{10}.$$

6. Is 0-50000-30042-6 is a valid UPC number?
7. Is 0-50000-30043-6 a valid UPC number?
8. Write a formula to calculate the check digit, d_{12}, in the UPC number.

Definition 6.3.6. Given a code, a **transposition error** is when order of two digits have been interchanged with each other.

9. The UPC error detection scheme can detect many transposition errors. Show that the transposition error 0-50000-03042-6 is detected.
10. Show that the transposition error 0-05000-30042-6 is not detected.
11. Can you find a general rule for the types of transposition errors that can be detected by UPC?

For Exercises 12–16, we discuss ISBN codes, as defined in Definition 6.3.7.

Definition 6.3.7. Every book has an **International Standard Book Number** (or **ISBN**) **code**. This is a 10-digit code indicating the book's publisher and title. The tenth digit is a check digit satisfying

$$(10, 9, 8, \dots, 1) \cdot (d_1, d_2, d_3, \dots, d_{10}) \equiv 0 \pmod{11}.$$[iii]

12. Is ISBN 0-534-91500-0 a valid ISBN code?
13. Is ISBN 0-534-91700-0 a valid ISBN code?
14. Is ISBN 0-534-19500-0 a valid ISBN code?

[iii] One problem is that d_{10} might have to be a 10 to make the inner product zero; in this case, 11 digits would be needed to make this scheme work. Therefore, the character X is used for the eleventh digit, that is, X means 10. So ISBN 3-540-96035-X is a valid ISBN code.

15. Does ISBN detect all single-digit errors?

16. Does ISBN detect all transposition errors?

17. Inner products often prove useful in implementing error-detecting codes, such as UPC and ISBN, as above. Suppose that

$$(d_1, d_2, \ldots, d_k) \cdot (w_1, w_2, \ldots, w_k) \equiv 0 \pmod{n}$$

is an error-detecting code. Prove that all single-digit errors are detected if and only if $\gcd(w_i, n) = 1$ for $1 \leq i \leq k$.

"An error doesn't become a mistake until you refuse to correct it." – Orlando Aloysius Battista

Lecture Videos

The Hamming Metric

Error-Detecting and Error-Correcting Codes

Group Codes

6.4 The Hamming Metric

Definition 6.4.1. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}_2^n$. The **Hamming distance** between $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{Z}_2^n$, denoted $d(\boldsymbol{x}, \boldsymbol{y})$, is the number of bits in which $\boldsymbol{x}$ and $\boldsymbol{y}$ differ. The **norm** (or **weight**) of $\boldsymbol{x}$, denoted $\|\boldsymbol{x}\|$ is the total number of 1's in $\boldsymbol{x}$.

The **minimum distance** of a code C, denoted $d_{\min}$ is the minimum of all Hamming distances $d(\boldsymbol{x}, \boldsymbol{y})$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are distinct codewords in C.

Example 6.4.2. Let $\boldsymbol{x} = 10101$, $\boldsymbol{y} = 11010$, and $\boldsymbol{z} = 00011$ be the set of codewords for a code C. Then

$$d(\boldsymbol{x}, \boldsymbol{y}) = 4, \quad d(\boldsymbol{x}, \boldsymbol{z}) = 3, \quad d(\boldsymbol{y}, \boldsymbol{z}) = 3.$$

Thus, the minimum distance of C is 3. Also,

$$\|\boldsymbol{x}\| = 3, \quad \|\boldsymbol{y}\| = 3, \quad \|\boldsymbol{z}\| = 2.$$

Proposition 6.4.3. *Let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{Z}_2^n$. Then*

(i) (positive definite) $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$ and $d(\boldsymbol{x}, \boldsymbol{y}) = 0$ if and only if $\boldsymbol{x} = \boldsymbol{y}$;

(ii) (symmetric) $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$;

(iii) (triangle inequality) $d(\boldsymbol{x}, \boldsymbol{y}) \leq d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})$

Any function $d : X \times X \to \mathbb{R}$ which satisfies the three axioms in Proposition 6.4.3 is called a **metric**.

Proposition 6.4.4. *Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}_2^n$. Then $\|\boldsymbol{x} + \boldsymbol{y}\| = d(\boldsymbol{x}, \boldsymbol{y})$. In particular, $\|\boldsymbol{x}\| = d(\boldsymbol{x}, \boldsymbol{0})$.*

Proof. Note that $\|\boldsymbol{x} + \boldsymbol{y}\|$ counts the number of 1's in $\boldsymbol{x} + \boldsymbol{y}$, which is exactly the number of terms where $\boldsymbol{x}$ and $\boldsymbol{y}$ differ since $0 + 0 = 1 + 1 = 0$ and $1 + 0 = 0 + 1 = 1$. In particular, $\|\boldsymbol{x}\| = \|\boldsymbol{x} + \boldsymbol{0}\| = d(\boldsymbol{x}, \boldsymbol{0})$. □

Suppose that $\boldsymbol{x} = 1101$ and $\boldsymbol{y} = 1100$ are codewords in some code. If we transmit 1101 and an error occurs in the rightmost bit, then 1100 will be received. Since 1100 is a codeword, the decoder will decode 1100 as the transmitted message. This code is clearly not very appropriate for error detection. The problem is that $d(\boldsymbol{x}, \boldsymbol{y}) = 1$.

If $\boldsymbol{x} = 1100$ and $\boldsymbol{y} = 1010$ are codewords, then $d(\boldsymbol{x}, \boldsymbol{y}) = 2$. If $\boldsymbol{x}$ is transmitted and a single error occurs, then $\boldsymbol{y}$ can never be received. In particular, if $d(\boldsymbol{x}, \boldsymbol{y}) = k + 1$ and T transmits $\boldsymbol{x}$ with k errors, then the decoder D cannot confuse $T(\boldsymbol{x})$ with $\boldsymbol{y}$. This proves the following result.

Theorem 6.4.5. *Let C be a code with minimum distance $k+1$. Then C can error detect up to k errors.*

Given that the Hamming metric actually gives us a distance function, it makes sense to ask what the closest codeword to a binary message is. In fact, the decoder D can accomplish error correction by always mapping a binary message to its closest codeword. Of course, this would be undefined if the binary message was the midpoint between two codewords. For example, if $d(\boldsymbol{x}, \boldsymbol{y}) = 2\ell+1$ (or 2ℓ) and $T(\boldsymbol{x})$ has less than $\ell+1$ errors, then the decoder can correct $T(\boldsymbol{x})$ by mapping it back to $\boldsymbol{x}$. On the other hand, if $d(\boldsymbol{x}, \boldsymbol{y}) = 2\ell+1$ (or 2ℓ) and $T(\boldsymbol{x})$ has greater than $\ell+1$ errors, then the the closest codeword is $\boldsymbol{y}$ instead of $\boldsymbol{x}$. We summarize this in the next theorem.

Theorem 6.4.6. *Let C be a code with minimum distance $2\ell+1$. Then C can error correct up to ℓ errors.*

Example 6.4.7. The following table shows the eight codewords of a code and their respective Hamming distances between each other. We can clearly see that $d_{\min} = 2$. Therefore, code has 1-error detection and no error-correction.

	0000	0011	0101	0110	1001	1010	1100	1111
0000	0	2	2	2	2	2	2	4
0011	2	0	2	2	2	2	4	2
0101	2	2	0	2	2	4	2	2
0110	2	2	2	0	4	2	2	2
1001	2	2	2	4	0	2	2	2
1010	2	2	4	2	2	0	2	2
1100	2	4	2	2	2	2	0	2
1111	4	2	2	2	2	2	2	0

Example 6.4.8. The following table shows the four codewords of a code and their respective Hamming distances between each other. We can clearly see that $d_{\min} = 3$. Therefore, code has 2-error detection and 1-error correction.

	00000	00111	11100	11011
00000	0	3	3	4
00111	3	0	4	3
11100	3	4	0	3
11011	4	3	3	0

Although the codes we have discussed so far have utilized the binary operation of $+$ on $\mathbb{Z}_2^n$, we have not utilized that $(\mathbb{Z}_2^n, +)$ is a group, until now.

Definition 6.4.9. A **group code** is a code C which is also a subgroup of $\mathbb{Z}_2^n$.

Because $C \leq \mathbb{Z}_2^n$, the code C contains the zero codeword $\mathbf{0}$ and the sum of two codewords is a codeword,[i] As a reminder, it is not necessary to check that C contains $\mathbf{0}$, but instead one can check that $C \neq \emptyset$ because if $\boldsymbol{x} \in C$ then $\boldsymbol{x} + \boldsymbol{x} = \mathbf{0} \in C$.

Example 6.4.10. Consider the following code C in $\mathbb{Z}_2^7$ is a group code has order 16:

0000000	0001111	0010101	0011010
0100110	0101001	0110011	0111100
1000011	1001100	1010110	1011001
1100101	1101010	1110000	1111111

One of the great advantages of group codes is the simplicity of computing the minimum distance.

Theorem 6.4.11. *Let C be a group code. Then $d_{\min}$ is the minimum of all the norms of nonzero codewords in C, that is,*

$$d_{\min} = \min\{\|\boldsymbol{x}\| \mid \boldsymbol{x} \in C, \boldsymbol{x} \neq \mathbf{0}\}.$$

Proof. By definition, $d_{\min} = \min\{d(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{x}, \boldsymbol{y} \in C, \boldsymbol{x} \neq \boldsymbol{y}\}$. By Proposition 6.4.4, $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} + \boldsymbol{y}\|$. Note also that $\boldsymbol{x} \neq \boldsymbol{y}$ if and only if $\boldsymbol{x} + \boldsymbol{y} \neq 0$. Since C is a group code, if $\boldsymbol{x}, \boldsymbol{y} \in C$ then $\boldsymbol{x} + \boldsymbol{y} \in C$. Therefore, the two sets are equal as they span the same elements. □

Example 6.4.12. In the previous exercise, the minimum norm was 3. Therefore, $d_{\min} = 3$ and the code C can error-correct one error. This $(7, 4)$-block code can encode four bits ($2^4 = 16$) with 1-error correction. This is a great improvement over the triple repetition $(12, 4)$ which also has 1-error correction.

[i] One might wonder why the inverse axiom is not mentioned here. Since every element in $\mathbb{Z}_2^n$ is its own inverse, every subset of $\mathbb{Z}_2^n$ is closed under inverses.

[ii] See §8.1 Error-Detecting and Correcting Codes and §8.2 Linear Codes in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, compute the Hamming norm of the vector.

1. 1011
2. 01111
3. 011-010
4. 1111-0101

For Exercises 5–8, compute the Hamming distance between the two vectors.

5. 00110, 01111
6. 1001, 0111
7. 011-010, 011-100
8. 1111-0101, 0101-0100

For Exercises 9–12, find the minimum distance of the given code. What is the error-detection and error-correction efficacy?

9. $\{011\text{-}010, 011\text{-}100, 110\text{-}111, 110\text{-}000\}$
10. $\{000\text{-}000, 011\text{-}100, 110\text{-}101, 110\text{-}001\}$
11. $\{011\text{-}100, 011\text{-}011, 111\text{-}011, 100\text{-}011, 000\text{-}000, 010\text{-}101, 110\text{-}100, 110\text{-}011\}$
12. $\{011\text{-}0\text{-}110, 011\text{-}1\text{-}100, 111\text{-}0\text{-}000, 111\text{-}1\text{-}111, 100\text{-}1\text{-}001, 100\text{-}0\text{-}011, 000\text{-}1\text{-}111, 000\text{-}0\text{-}000\}$
13. Why is the following code is not a group code?

Message	00	01	10	11
Codeword	0110	1010	1100	1001

14. For all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}_2^n$, prove that $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$.
15. For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{Z}_2^n$, prove that $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + \boldsymbol{z})$.

Definition 6.4.13. Let $d : X \times X \to \mathbb{R}$ be a function for some set X. We call d a **metric** (or **distance**) function if it satisfies the three following metric axioms:
We say that d is a **metric** (or **distance**) function if the following three axioms:

(i) (Triangle Inequality): For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{Z}_2^n$, it holds that

$$d(\boldsymbol{x}, \boldsymbol{y}) \leq d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y}).$$

(ii) (Summetric): For all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}_2^n$, it holds that

$$d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x}).$$

(iii) (Positive Definite): For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{Z}_2^n$, it holds that

$$d(\boldsymbol{x}, \boldsymbol{y}) \geq 0,$$

where equality holds if and only if $\boldsymbol{x} = \boldsymbol{y}$.

We say that (X, d) is a **metric space**.

16. Prove that $\mathbb{Z}_2^n$ equiped with the Hamming metric is a metric space.[iii]

[iii] Hence, decoding a message reduces to deciding which is the closest codeword in terms of this distance d.

"It is easier to criticize than to correct our past errors." – Livy

Lecture Videos

Coding Theory and Linear Algebra

Linear Codes

Decoding a Linear Codeword

6.5 Linear Codes

A binary sequence can naturally be viewed as a column vector. This perspective allows us to aid our group theory with the power of linear algebra, a discipline of abstract algebra known as **representation theory**. One example of this is the parity-check from the ASCII code. Let $\boldsymbol{x} \in \mathbb{Z}_2^7$. Then the check bit at the beginning of the ASCII codeword is equal to the *inner product* (or *dot product*) of $\boldsymbol{x}$ and the all 1's vector $\mathbf{1}$:

$$\mathbf{1} \cdot \boldsymbol{x} = \mathbf{1}^T \boldsymbol{x} = 1 \cdot x_1 + 1 \cdot x_2 + \ldots + 1 \cdot x_7 = x_1 + x_2 + \ldots + x_7.$$

Likewise, we can compute the norm of a codeword using inner products, that is, $\|\boldsymbol{x}\| = \boldsymbol{x} \cdot \boldsymbol{x}$, where the dot product is considered in $\mathbb{R}^n$, not $\mathbb{Z}_2^n$.

Let $M_{mn}(\mathbb{Z}_2)$ be the set of $m \times n$ matrices with entries from $\mathbb{Z}_2$. This forms an abelian group under matrix addition and forms a vector space when coupled with scalar multiplication. Let $H \in M_{mn}(\mathbb{Z}_2)$. Then the **null space** of H, denoted $\text{nul}(H)$, is the set of all vectors in $\mathbb{Z}_2^n$ whose product with H on the left is the zero vector, that is,

$$\text{nul}(H) = \{\boldsymbol{y} \mid H\boldsymbol{y} = \mathbf{0}\}.$$

It is a standard result from Linear Algebra that $\text{nul}(H)$ is a subspace of $\mathbb{Z}_2^n$ and hence an additive subgroup of $\mathbb{Z}_2^n$.

Theorem 6.5.1. *Let $H \in M_{m\times n}(\mathbb{Z}_2)$. Then $C = \text{nul}(H)$ is a group code, called a **linear code**.*

Example 6.5.2. Let $H = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$. Let $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)^T \in \text{nul}(H)$. Then $\boldsymbol{x}$ is a solution to the homogeneous linear system below:

$$\begin{cases} x_2 + x_4 = 0 \\ x_1 + x_2 + x_3 + x_4 = 0 \\ x_3 + x_4 + x_5 = 0 \end{cases} \sim \begin{cases} x_1 + x_4 + x_5 = 0 \\ x_2 + x_4 = 0 \\ x_3 + x_4 + x_5 = 0 \end{cases}$$

Gaussian Elimination can be used to row reduce the matrix/linear system, performing row operations

in $\mathbb{Z}_2$, as displayed above. Solving the system gives the general solution

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_4 + x_5 \\ x_4 \\ x_4 + x_5 \\ x_4 \\ x_5 \end{pmatrix} = x_4 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Therefore, the null space of H is given as

$$C = \text{Span}\{11110, 10101\} = \{00000, 11110, 10101, 01011\},$$

which is a $(5, 2)$-block group code. With $d_{\min} = 3$, the code can 2-error detect and 1-error correct.

Example 6.5.3. Let C be the linear code given by the matrix $H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$.

It is very easy for the decoder to check if a transmitted message is a codeword without computing the null space of H since H is the set of vectors $\boldsymbol{x}$ such that $H\boldsymbol{x} = \mathbf{0}$. Suppose that $\boldsymbol{x} = 010\ 011$ is received. Then

$$H\boldsymbol{x} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0+0+0+0+1+1 \\ 0+1+0+0+1+1 \\ 0+0+0+0+0+1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore, the code has detected an error. Now the decoder must decide to request the message be sent again or correct it itself.

Definition 6.5.4. Let $A \in M_{(n-m)\times m}(\mathbb{Z}_2)$. The the matrix $H = (A \mid I_{n-m})$ is an $(n-m) \times n$ matrix called a **canonical parity-check matrix**. Associated to this matrix H is the $n \times m$ matrix $G = \begin{pmatrix} I_m \\ \hline A \end{pmatrix}$, called the **standard generator matrix**.

These two matrices have a special relationship, specifically

$$HG = \left(\begin{array}{c|c} A & I_{n-m} \end{array} \right) \left(\begin{array}{c} I_m \\ \hline A \end{array} \right) = AI_m + I_{n-m}A = A + A = 0,$$

where here 0 denotes the $(n-m) \times m$ zero matrix. Of course, the last equality in the above equation follows from working modulo 2, every element is its own inverse. In particular, let $\boldsymbol{y} = G\boldsymbol{x}$ for some $\boldsymbol{x} \in \mathbb{Z}_2^m$. Then

$$H\boldsymbol{y} = H(G\boldsymbol{x}) = (HG)\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0}.$$

Thus, $\boldsymbol{y} \in \text{nul}(H)$, that is, $\text{im}(G) \leq \text{nul}(H)$. On the other hand, because G has I_m in its top m rows, $\text{rank}(G) = m$. So, $\dim(\text{im}(G)) = m$. On the other hand, H also contains a copy of I_{n-m}. This means that the $\text{nullity}(H) \leq n - (n-m) = m$. Finally,

$$m = \dim(\text{im}(G)) \leq \dim(\text{nul}(H)) \leq m.$$

Thus, $\dim(\text{im}(G)) = \dim(\text{nul}(H))$ and $\text{nul}(H) = \text{im}(G)$. Therefore, the matrix G generates the linear code associated to the matrix H.

Example 6.5.5. Suppose that we have the following eight words to encode (all of $\mathbb{Z}_2^3$):

$$000, 001, 010, 011, 100, 101, 110, 111.$$

Using $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, we construct the canonical parity-check matrix H and its generator G:

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

One possible encoding process is to compute $\text{nul}(H)$, which corresponds to the solution set of the homogeneous system

$$\begin{cases} \quad x_2 + x_3 + x_4 \quad\quad\quad = 0 \\ x_1 + x_2 \quad\quad + x_5 \quad\quad = 0 \\ x_1 \quad\quad + x_3 \quad\quad\quad + x_6 = 0 \end{cases} \sim \begin{cases} \quad x_2 + x_3 = x_4 \\ x_1 + x_2 \quad\quad = x_5 \\ x_1 \quad\quad + x_3 = x_6 \end{cases}$$

This gives

$$\text{nul}(H) = \text{Span}\{100\ 011, 010\ 110, 001\ 101\} = \left\{ \begin{array}{l} 000\ 000,\ 100\ 011,\ 010\ 110,\ 110\ 101, \\ 001\ 101,\ 101\ 110,\ 011\ 011,\ 111\ 000 \end{array} \right\}.$$

Considering the reduced linear system above, each of the last three bits serves as parity check, x_4 can check if there is an error with x_2 and x_3, x_5 can check if there is an error with x_1 and x_2, and x_6 can check if there is an error with x_1 and x_3. Thus, the code can correct a single error.

Alternatively, each of these codewords could be computed as an image of G:

$$\left\{ \begin{array}{l} G(000) = 000\ 000,\ G(100) = 100\ 011,\ G(010) = 010\ 110,\ G(110) = 110\ 101, \\ G(001) = 001\ 101,\ G(101) = 101\ 110,\ G(011) = 011\ 011,\ G(111) = 111\ 000 \end{array} \right\}$$

Notice that the first three bits of each codeword is just the decoded message.

All these results are summarized in the following theorem.

Theorem 6.5.6. *If $H \in M_{(n-m)\times n}(\mathbb{Z}_2)$ is a canonical parity-check matrix, then* $\mathrm{nul}(H)$ *consists of all $\boldsymbol{y} \in \mathbb{Z}_2^n$ whose first m bits are arbitrary, called the* ***information bits****, but whose last $n-m$ bits are determined by $H\boldsymbol{x} = \boldsymbol{0}$. Each of the last $n-m$ bits, called the* ***check bits****, serves as a parity check for some of the first m bits. Hence, H gives rise to an (n, m)-block code. Furthermore, if G is the standard generator matrix, then* $\mathrm{nul}(H) = \mathrm{im}(G)$.

Hence, every group code can be realized as a linear code for some matrix H.

[i] See §8.2 Linear Codes and §8.3 Parity-Check and Generator Matrices in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, compute the null space for the given binary matrix H. What type of (n, k)-linear code is it? Find a generator matrix G associated to this linear code.

1. $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$

5. Construct a $(5, 2)$-linear code. What is the error-detection and error-correction efficacy?

6. Let C be a linear code. Prove that either every codeword has even norm or exactly half of the codewords have even norm.

7. Prove that the codewords of even norm in a linear code C form a subcode.

"An error does not become truth by reason of multiplied propagation, nor does truth become error because nobody sees it." – Mahatma Gandhi

Lecture Videos

Linear Codes and Minimum Dependency Relations

Error Correction in Linear Codes

6.6 Decoding

The efficacy of linear codes comes from balance between linear independence and dependence in the column vectors of H. For example, to form an (n, m)-block code, H must be a $p \times n$ matrix with $\text{nullity}(H) = m$. The nullity determines the size of code C, and as a consequence the message set $\mathbb{Z}_2^m$. For parity-check matrices, the rank of H, which is $\text{rank}(H) = n - m$, will be the size of the augmented identity matrix I_{n-m}. In particular, we have that $p = n - m$. But still, how does one choose the matrix A? More importantly, how large does one set n? The larger n is, the more room there is possible between codewords, thus increasing the minimum distance. This allows for larger levels of error detection and correction. Thus, a large n allows better codes, but a large n does not guarantee better detection/correction. This depends on the choice of A.

Let $\boldsymbol{e}_i$ denote vector of all 0's except for a 1 in the ith position. These are very important vectors in Linear Algebra. As i ranges from 1 to n, the set of the $\boldsymbol{e}_i$'s form the standard basis on $\mathbb{Z}_2^n$. Also, the product $H\boldsymbol{e}_i$ is equal to the ith column of the matrix H. Let $\boldsymbol{x} \in \mathbb{Z}_2^n$. Then there exists some index set I such that $\boldsymbol{x} = \sum_{i \in I} \boldsymbol{e}_i$. Thus, $H\boldsymbol{x} = \sum_{i \in I} H\boldsymbol{e}_i$, that is, $H\boldsymbol{x}$ is a sum of columns of H. If $H\boldsymbol{x} = \boldsymbol{0}$, then the corresponding sum of columns is zero. Let $\boldsymbol{x} \in \text{nul}(H)$ such that $\boldsymbol{x}$ has minimum nonzero weight. Of course, $w(\boldsymbol{x}) = |I|$. Then $\sum_{i \in I} H\boldsymbol{e}_i$ is the smallest nontrivial linear combination of column vectors of H which combine to zero. We call this a **minimum dependency relation** on the columns of H. Then $d_{\min} = w(\boldsymbol{x}) = |I|$, the size of a minimum dependency relation. We have then shown the following.

Theorem 6.6.1. *Let H be a matrix with $\mathbb{Z}_2$-entries. Let d be the size of a minimum dependency relation on the columns of H. Then the linear code $\text{nul}(H)$ can detect $d-1$ errors and correct $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors.*

Corollary 6.6.2. *Let H be an $m \times n$ binary matrix. Then the null space of H is a single error-detecting code if and only if no column of H consists entirely of zeros. Furthermore, the null space of H is a single error-correcting code if and only if H does not contain any zero columns and no two columns of H are identical.*

Proof. Let d denote the size of a minimum dependency relation. For the first equivalence, note that both statements imply that $d > 1$. For the second equivalence, note that both statements imply that $d > 2$. □

Example 6.6.3. The linear code associated with matrix H_1, below, can detect 1 error since it has no column of zeros but cannot correct the error since the first two columns are identical. The linear code associated with matrix H_2, also below, cannot detect an error since it has a column of zeros.

Finally, the linear code associated with matrix H_3 can both correct and detect a single error.

$$H_1 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad H_3 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Although a code CAN correct an error, how can this be done efficiently? Consider the linear code C associated with the binary matrix H. If $\boldsymbol{x} \in \mathbb{Z}_2^n$ was received and has no errors, then no correction is necessary. As mentioned before, we can detect errors by computing $H\boldsymbol{x}$, called the **syndrome** of $\boldsymbol{x}$. If the syndrome is zero, then $\boldsymbol{x}$ has no errors. If H is a parity-check matrix, then decoding just means finding the information bits and discarding the parity bits.

In the presence of errors, let $\boldsymbol{\mu}, \boldsymbol{\varepsilon} \in \mathbb{Z}_2^n$ such that $\boldsymbol{\mu}$ was the original codeword, that is, $T(\boldsymbol{\mu}) = \boldsymbol{x}$, and $\boldsymbol{\varepsilon}$ is the error added to $\boldsymbol{\mu}$ from noise, that is, $\boldsymbol{x} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}$. Then

$$H\boldsymbol{x} = H(\boldsymbol{\mu} + \boldsymbol{\varepsilon}) = H\boldsymbol{\mu} + H\boldsymbol{\varepsilon} = H\boldsymbol{\varepsilon},$$

where the last equality follows since $\boldsymbol{\mu}$ is a codeword, that is, $\boldsymbol{\mu} \in \text{nul}(H)$. Therefore, $\boldsymbol{x}$ has the same syndrome as its error. If $\boldsymbol{\varepsilon} = \boldsymbol{e}_i$, then $H\boldsymbol{x} = H\boldsymbol{\varepsilon} = H\boldsymbol{e}_i$, which the the ith column of H. Therefore, the error of $\boldsymbol{x}$ can be found in the ith bit. Therefore, using syndromes, we can both detect and correct a single error.

For multiple errors, the detection process remains the same, use the syndrome. For error correction, we will solve the linear system associated with the augmented matrix $(H \mid H\boldsymbol{x})$. This will express $H\boldsymbol{x}$ as a linear combination of the columns of H. Since the columns are linearly dependent, there will be multiple solutions. We assume, of course, that the numbers of errors is few and choose the vector in this solution set of minimum weight. On the other hand, this solution set is none other than the coset of $C = \text{nul}(H)$ in $\mathbb{Z}_2^n$ represented by $\boldsymbol{x}$, that is, $\boldsymbol{x} + C$. Thus, to correct errors we select the minimum-weighted vector in the coset $\boldsymbol{x} + C$. This is known as **Coset Decoding**.

Example 6.6.4. Consider the linear code associated with $H = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$. Suppose that the codes $\boldsymbol{x} = 001\text{-}111$, $\boldsymbol{y} = 111\text{-}110$, $\boldsymbol{z} = 010\text{-}111$, and $\boldsymbol{w} = 111\text{-}111$ are received. We compute their syndromes below:

$$H\boldsymbol{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad H\boldsymbol{y} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad H\boldsymbol{z} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad H\boldsymbol{w} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Considering the syndromes, the first message was sent without errors and is $\boxed{\textbf{001}\text{-}111}$. For the second message, the syndrome is the third column of H and thus $\boldsymbol{y}$ has an error in the third bit. Thus, the original message was $\boxed{\textbf{110}\text{-}110}$. For the third message, the syndrome is the fourth column of H. Thus, the error is in the fourth bit and the sent message was $\boxed{\textbf{010}\text{-}011}$.

The last one is more difficult. The syndrome does not correspond to a column of H and hence represents more than one error in transmission. We can compute C using the generator matrix G:

$$C = \{000\text{-}000, 001\text{-}111, 010\text{-}011, 011\text{-}100, 100\text{-}101, 101\text{-}010, 110\text{-}110, 111\text{-}001\},$$

and we get the coset

$$\boldsymbol{x} + C = \{000\text{-}110, 001\text{-}001, 010\text{-}101, 011\text{-}010, 100\text{-}011, 101\text{-}100, 110\text{-}000, 111\text{-}111\}.$$

Unfortunately, $\boldsymbol{x}+C$ has three elements of weight 2! This is because $d_{\min} = 3$ and C can only correct a single error not two. But C can detect two errors, so the receiver should request the message be sent again.

[i] See §8.2 Linear Codes and §8.4 Efficient Decoding in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, determine whether the matrix H is a canonical parity-check matrix. If so, find the corresponding standard generator matrix G. What is the error-detection and error-correction effecacy of the code generated by H?

1. $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$

2. $\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$

For Exercises 5–8, let

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Compute the syndrome caused by the transmission error in the given bit.

5. 1st bit
6. 3rd bit
7. 5th bit
8. 3rd and 4th bits

For Exercises 9–12, let C be the linear code associated to the matrix H

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Decode the given message, if possible.

9. 01111
10. 10101
11. 01110
12. 00011

6.7 Supplemenatry Exercises

(Go to Solutions)

1. The following cipher has been encrypted using a monoalphabetic cipher. Use Frequency Analysis and the website https://www.101computing.net/frequency-analysis/ to decrypt this message.

   ```
   DJ DK C QLXDWI WF SDGDU PCX. XLRLU KQCSLKBDQK, KJXDHDET FXWZ C BDIILE RCKL,
   BCGL PWE JBLDX FDXKJ GDSJWXO CTCDEKJ JBL LGDU TCUCSJDS LZQDXL. IYXDET JBL
   RCJJUL, XLRLU KQDLK ZCECTLI JW KJLCU KLSXLJ QUCEK JW JBL LZQDXL'K YUJDZCJL
   PLCQWE, JBL ILCJB KJCX, CE CXZWXLI KQCSL KJCJDWE PDJB LEWYTB QWPLX JW ILKJXWO
   CE LEJDXL QUCELJ. QYXKYLI RO JBL LZQDXL'K KDEDKJLX CTLEJK, QXDESLKK ULDC
   XCSLK BWZL CRWCXI BLX KJCXKBDQ, SYKJWIDCE WF JBL KJWULE QUCEK JBCJ SCE KCGL
   BLX QLWQUL CEI XLKJWXL FXLLIWZ JW JBL TCUCVO...
   ```

2. Suppose that
   $$(d_1, d_2, \ldots, d_k) \cdot (w_1, w_2, \ldots, w_k) \equiv 0 \pmod{n}$$
   is an error-detecting code. Prove that all transposition errors are detected if and only if $\gcd(w_i - w_j, n) = 1$ for $1 \leq i < j \leq k$.

3. Suppose that a code C has a minimum weight of 7. What is the error-detection and error-correction efficacy?

4. Let C be a linear code. Prove that either the ith coordinates in the codewords of C are all zero or exactly half of them are zero.

5. If we are to use an error-correcting linear code to transmit the 128 ASCII characters, what size matrix must be used? What size matrix must be used to transmit the extended ASCII character set of 256 characters? What if we require only error detection in both cases?

6. Find the canonical parity-check matrix that gives the even parity check bit code with three information positions. What is the matrix for seven information positions? What are the corresponding standard generator matrices?

7. How many check positions are needed for a single error-correcting code with 20 information positions? With 32 information positions?

> **Definition 6.7.1.** Let C be an (n, m)-linear code. The orthogonal code (or **dual**) of C, denoted $C^\perp$, is
> $$C^\perp = \{\boldsymbol{x} \in \mathbb{Z}_2^n \mid \boldsymbol{x} \cdot \boldsymbol{y} = 0, \boldsymbol{y} \in C\}.$$

8. If C is the linear code given by H below, find $C^\perp$.
   $$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

9. Prove that $C^\perp$ is an $(n, n-m)$-linear code.

10. If H and G are the parity-check and generator matrices, respectively, of C, prove that $G^\top$ and $H^\top$ are the parity-check and generator matrices, respectively, of $C^\perp$.

> **Definition 6.7.2.** Let H be the $m \times n$ binary matrix whose jth column is the binary expansion of j. Then the linear code associated to H is called a **Hamming code**.

11. Determine that

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

generates a Hamming code. What are the error-correcting properties of this Hamming code?

Chapter 7

Isomorphisms

"Whether I'm wearing lots of makeup or no makeup, I'm always the same person inside." – Lady Gaga

Lecture Videos

Group Isomorphisms

Properties of Group Isomorphisms

Group Invariants

7.1 Isomorphisms

Definition 7.1.1. Two groups $(G, *)$ and $(H, \circ)$ are **isomorphic** if there exists a bijective map $\varphi : G \to H$ such that the group operation is *preserved* by φ, that is, for all $a, b \in G$,

$$\varphi(a * b) = \varphi(a) \circ \varphi(b).$$

The above equation is often referred to as the **homomorphic property**. If G is isomorphic to H, we write $G \cong H$. The map φ is called an **isomorphism**.

Another way of expressing the homomorphic property is to say that the image of a product is equal to a product of images. In a way, the product in the image is the same as the product in the domain. The word isomorphism derives from Greek which means *"same shape."* When two groups are isomorphic, they are essentially the same group, although the appearance of the groups may differ. Of course, the relation $\cong$ forms an equivalence relation on the class of all groups.

Example 7.1.2. Consider the map $\varphi : \mathbb{R} \to \mathbb{R}^+$ given by exponentiation: $\varphi(x) = e^x$ for all $x \in \mathbb{R}$. This forms an isomorphism between $(\mathbb{R}, +)$ and $(\mathbb{R}^+, \cdot)$. The map φ is bijective since it has an inverse $\psi : \mathbb{R}^+ \to \mathbb{R}$ given by the natural logarithm: $\psi(x) = \ln(x)$. Note that $\varphi \circ \psi(x) = e^{\ln x} = x$ and $\psi \circ \varphi(x) = \ln(e^x) = x$. Let $x, y \in \mathbb{R}$. Then

$$\varphi(x + y) = e^{x+y} = e^x \cdot e^y = \varphi(x) \cdot \varphi(y).$$

This proves the homomorphic property. Thus, φ is an isomorphism and $(\mathbb{R}, +) \cong (\mathbb{R}^+, \cdot)$.

Example 7.1.3. Let $\varphi : \mathbb{Z}_4 \to \langle i \rangle$ be given by the rule $\varphi(n) = i^n$. We first check that this map is well-defined since $\mathbb{Z}_4$ is technically a set of equivalence classes and we need to verify if different representatives gives the same image. Let $m \equiv n \pmod 4$. Then there exists some $k \in \mathbb{Z}$ such that $m = n + 4k$. Then $\varphi(m) = i^m = i^{n+4k} = i^n \cdot (i^4)^k = i^n \cdot 1 = i^n = \varphi(n)$. Therefore, φ is well-defined.

Let $m, n \in \mathbb{Z}_4$ and suppose that $\varphi(m) = \varphi(n)$. Then $i^m = i^n$, that is, $i^{m-n} = 1$. Then $4 \mid (m - n)$ since $|i| = 4$ in $\mathbb{C}^*$. Thus, $m \equiv n \pmod 4$. This shows that φ is one-to-one. Of course, if $i^n \in \langle i \rangle$ then $\varphi(n) = i^n$ and φ is surjective. This shows that φ is a bijection.

To show the homomorphic property, let $m, n \in \mathbb{Z}_n$ and consider

$$\varphi(m + n) = i^{m+n} = i^m \cdot i^n = \varphi(m) \cdot \varphi(n).$$

Therefore, φ is an isomorphism and $\mathbb{Z}_4 \cong \langle i \rangle$. We mention that this prove can be modified to show that $\mathbb{Z}_n \cong \langle e^{2\pi i/n} \rangle$.

Isomorphism preserve more than just multiplication.

Theorem 7.1.4. *Let $\varphi : G \to H$ be a group isomorphism. Then the following hold.*

(i) $|G| = |H|$;

(ii) If G is abelian, then H is abelain;

(iii) If G is cyclic, then H is cyclic;

(iv) If G has a subgroup of order n, then H has a subgroup of order n.

Proof. Since φ is a bijective, it preserves the cardinality of sets, which proves ((i)). Suppose that G is abelian. Then if $a, b \in H$, then there exists $x, y \in G$ such that $\varphi(x) = a$ and $\varphi(y) = b$. Then

$$ab = \varphi(x)\varphi(y) = \varphi(xy) = \varphi(yx) = \varphi(y)\varphi(x) = ba.$$

Thus, H is abelian also, proving ((ii)). The remaining properties are exercises for the reader. □

To show that two groups are not isomorphic we would have to show that no isomorphism exists. The proof of non-existence is typically done by proof-by-contradiction, that is, show that a contradiction occurs if we suppose there is an isomorphism. We can do this by showing there is some property of a group that should be preserved by an isomorphism, called an **invariant** but is not preserved. Theorem 7.1.4 provides four invariants of groups: group order, commutativity, cyclic generation, and subgroup/element order. If two groups are isomorphic, then they either both have the invariant property or neither of them have the invariant property.

Example 7.1.5. The dihedral groups D_4 and D_5 are not isomorphic since $|D_4| = 8 \neq 10 = |D_5|$ and group order is an invariant. Thus, $D_4 \ncong D_5$. In fact, $D_n \cong D_m$ if and only if $n = m$.

We next will show that S_3 and $\mathbb{Z}_6$ are not isomorphic. This time $|S_3| = 6 = |\mathbb{Z}_6|$, so we need to compare a different invariant. We note that S_3 is nonabelian and $\mathbb{Z}_6$ is abelian. Since commutativity is an invariant, $S_3 \ncong \mathbb{Z}_6$.

Example 7.1.6. The groups $\mathbb{Z}_8$ and $\mathbb{Z}_{12}$ cannot be isomorphic since there is no bijections between a set of cardinality 8 and a set of cardinality 12. On the other hand, $\mathbb{Z}_8^* \cong \mathbb{Z}_{12}^*$, which can be seen by using the isomorphism $\varphi : \mathbb{Z}_8^* \to \mathbb{Z}_{12}^*$ given as $\varphi = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 1 & 5 & 7 & 11 \end{pmatrix}$. This clearly gives a well-defined bijection. The homomorphic property is a little more difficult to prove. We can accomplish by comparing the Cayley tables of these two groups, as below.

$\mathbb{Z}_8^*$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

$$\begin{array}{c} \varphi \\ 1 \mapsto 1 \\ 3 \mapsto 5 \\ 5 \mapsto 7 \\ 7 \mapsto 11 \end{array}$$

$\mathbb{Z}_{12}^*$	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

After comparing all possible products, we see that φ preserves multiplication and is an isomorphism. More interestingly, we say that $\mathbb{Z}_8^*$ and $\mathbb{Z}_{12}^*$ have the same Cayley table, after re-labeling, of course, provided by the map φ. This is what was meant by saying the groups were essentially the same but the appearance of the groups differs. By comparing Cayley tables, we also see that $\mathbb{Z}_8^* \cong \mathbb{Z}_{12}^* \cong \mathbb{Z}_2 \times \mathbb{Z}_2$,

the Klein 4-group.

[i]See §9.1 Definition and Examples in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–13, prove the given groups are isomorphic.
(Pick 3 from Exercises 1-8; Pick 1 from Exercise 9-13)

1. $(\mathbb{R}^*, \cdot) \cong (\mathbb{R} \setminus \{-1\}, *)$, where $a * b = a + b + ab$
2. $\mathbb{Z}_8^* \cong V_4$
3. $\mathbb{Z}_5^* \cong \mathbb{Z}_{10}^*$
4. $\mathbb{Z}_5^* \cong \mathbb{Z}_4$
5. $S_3 \times \mathbb{Z}_2 \cong D_6$
6. $\mathbb{C}^* \cong \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in \mathbb{R} \right\} \leq \mathrm{GL}_2(\mathbb{R})$
7. $\mathbb{Z}_8^8 \cong \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \leq \mathrm{GL}_2(\mathbb{R})$
8. $S_3 \cong \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \right\} \leq \mathrm{GL}_3(\mathbb{R})$
9. $\mathbb{Z} \cong n\mathbb{Z}$ for $n \neq 0$
10. $\mathbb{Z}_n \cong Z_n := \{\zeta_n^k \mid k \in \mathbb{Z}\}$, where $\zeta_n = e^{2\pi i/n}$
11. $\mathbb{Z}_n \cong \langle g \rangle$, for any $|g| = n$
12. $D_n \cong \left\{ \begin{pmatrix} \pm 1 & a \\ 0 & 1 \end{pmatrix} \middle| a \in \mathbb{Z}_n \right\} \leq \mathrm{GL}_2(\mathbb{Z}_n)$
13. $D_n \cong \left\langle \begin{pmatrix} \zeta_n & 0 \\ 0 & \overline{\zeta_n} \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle \leq \mathrm{GL}_2(\mathbb{C})$

For Exercises 14–16, prove the given groups are NOT isomorphic.
(Pick 2 from Exercises 14-16)

14. $\mathbb{Q} \ncong \mathbb{Z}$
15. $\mathbb{Z}_5^* \cong \mathbb{Z}_{12}^*$
16. $S_4 \ncong D_{12}$

17. Suppose that $\varphi : G \to H$ and $\psi : H \to K$ are two group isomorphisms. Prove that φ^{-1} and $\psi \circ \varphi$ are also group isomorphisms. Show that $\cong$ forms an equivalence relation on the category of groups.

"We may have all come on different ships, but we're in the same boat now." – Martin Luther King, Jr.

Lecture Videos

Cyclic Groups are Isomorphic up to their Order

Cayley's Theorem

7.2 Cayley's Theorem

Given that isomorphisms create an equivalence relation on the class of groups, when categorizing groups, it suffices to consider all groups up to isomorphism. It is very simple to classify all cyclic groups up to isomorphism.

Theorem 7.2.1. *Two cyclic groups are isomorphic if and only if they have the same order. In particular, the order of a cyclic group determines the group up to isomorphism.*

Proof. As mentioned earlier, if two groups are isomorphic then they have the same order. To prove converse for cyclic groups it suffices to prove that $G \cong \mathbb{Z}_n$ if $|G| = n$ since $\cong$ is an equivalence relation.

Let $G = \langle g \rangle$ be a cyclic group of order $|G| = n < \infty$. Define $\varphi : \mathbb{Z}_n \to G$ by the rule $\varphi(k) = g^k$. Suppose $\ell \equiv k \pmod{n}$. Then there exists an integer a such that $k = \ell + an$. Then $\varphi(k) = g^k = g^{\ell+an} = g^\ell \cdot (g^n)^a = g^\ell \cdot 1 = g^\ell = \varphi(\ell)$. Therefore, φ is well-defined.

Let $k, \ell \in \mathbb{Z}_n$ and suppose that $\varphi(k) = \varphi(\ell)$. Then $g^k = g^\ell$, that is, $g^{k-\ell} = 1$. Then $n \mid (k - \ell)$ since $|g| = n$ in G. Thus, $k \equiv \ell \pmod{n}$. This shows that φ is one-to-one. Of course, if $g^k \in G$ then $\varphi(k) = g^k$ and φ is surjective. This shows that φ is a bijection.

To show the homomorphic property, let $k, \ell \in \mathbb{Z}_n$ and consider

$$\varphi(k + \ell) = g^{k+\ell} = g^k \cdot g^\ell = \varphi(k) \cdot \varphi(\ell).$$

Therefore, φ is an isomorphism and $\mathbb{Z}_n \cong G$.

If $|G|$ is infinite, the only part of the above proof that need be adapted is the injectivity of φ. Let $k, \ell \in \mathbb{Z}$ and $\varphi(k) = \varphi(\ell)$. Then $g^k = g^\ell$ or $g^{k-\ell} = 1$. Since the order of g is infinite, it must be that $k - \ell = 0$ or $k = \ell$. Thus, φ is injective when $|G|$ is infinite. □

We can also classify all finite abelian groups up to isomorphism using direct products of cyclic groups, which will be discussed in more detail in the next section.

The classification of finite nonabelian groups up to isomorphism has been a much greater task. Although such a discussion is far beyond the topics taught in a first semester of algebra, many important topics related to this will be introduced in forth coming lectures.

In the meanwhile, we will prove the first representation theoretic theorem of group theory, known as Cayley's Theorem. Representation Theory, a research-active branch of Group Theory, aims to represent abstract groups as more familiar groups, such as cyclic groups, permutation groups, and matrix groups. The advantage of representing a group as a group of matrices allows one to employ Linear Algebra in the study of groups. Similar benefits exist for representing abstract groups as other concrete groups. Cayley's Theorem allows us to represent EVERY group as a permutation group.

Lemma 7.2.2. *Let G be a group with $g \in G$. Let $\lambda_g : G \to G$ be a map given by the rule $\lambda_g(x) = gx$. Then $\lambda_g : G \to G$ is a permutation.*

Proof. We must show that λ_g is bijective for each $g \in G$. Let $x, y \in G$ and suppose that $\lambda_g(x) = \lambda_g(y)$. Then $gx = gy$. By the left cancellation law, we conclude that $x = y$. Thus, λ_g is injective. Likewise, let $x \in G$. Then $g^{-1}x \in G$ and $\lambda_g(g^{-1}x) = g(g^{-1}x) = (gg^{-1})x = x$. Thus, λ_g is surjective. Therefore, λ_g is a permutation. □

Similarly, the map $\rho_g : G \to G$ given by $\rho_g(x) = xg^{-1}$ is likewise a permutation on G.

Example 7.2.3. Consider the Cayley table of $\mathbb{Z}_4$:

$\mathbb{Z}_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Each row of the Cayley table is just a permutation of the elements of group. This is exactly the permutations λ_g. Namely,

$$\lambda_0 = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix} = 1, \quad \lambda_1 = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \end{pmatrix} = (0123),$$

$$\lambda_2 = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 2 & 3 & 0 & 1 \end{pmatrix} = (02)(13), \quad \lambda_3 = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 3 & 0 & 1 & 2 \end{pmatrix} = (0321)$$

Similarly, the columns of the Cayley table are also permutations of G. These lead to ρ_g.

Lemma 7.2.4. *Let G be a group. Let $\overline{G} = \{\lambda_g \mid g \in G\}$. Then $\overline{G} \leq S_{|G|}$.*

Proof. Note that for all $x \in G$, we have $\lambda_e(x) = ex = x = Id(x)$. Thus, $\lambda_e = Id$, the identity of $S_{|G|}$. Let $\lambda_g, \lambda_h \in \overline{G}$. Then, for all $x \in G$,

$$(\lambda_g \circ \lambda_h)(x) = \lambda_g(\lambda_h(x)) = \lambda_g(hx) = g(hx) = (gh)x = \lambda_{gh}(x).$$

Thus, $\lambda_g \circ \lambda_h = \lambda_{gh} \in \overline{G}$. Thus, $\overline{G}$ is closed under multiplication. Finally, for all $x \in G$, $\lambda_g^{-1}(x) = g^{-1}x$ since $\varphi_g(g^{-1}x) = x$. But $g^{-1}x = \lambda_{g^{-1}}(x)$. Thus, $\lambda_g^{-1} = \lambda_{g^{-1}} \in \overline{G}$. This shows that $\overline{G}$ is closed under inverses. Therefore, $\overline{G}$ is a subgroup. □

Theorem 7.2.5 (Cayley's Theorem)**.** *Let G be a group. Let $\varphi : G \to \overline{G}$ a be map given by the rule $\varphi(g) = \lambda_g$. Then φ is an isomorphism.*

Proof. Let $g, h \in G$ such that $\varphi(g) = \varphi(h)$, that is, $\lambda_g = \lambda_h$. Let $x \in G$. Then $\lambda_g(x) = \lambda_h(x)$, that is, $gx = hx$. Canceling x on the right, we conclude that $g = h$. Thus, φ is injective. Clearly, $g \mapsto \lambda_g$ is surjective. Therefore, φ is bijective. Finally,

$$\varphi(gh) = \lambda_{gh} = \lambda_g \circ \lambda_h = \varphi(g) \circ \varphi(h).$$

Therefore, φ is an isomorphism. □

The map $g \mapsto \lambda_g$ is known as the **left regular representation** of G and the map $g \mapsto \rho_g$ is known as the **right regular representation**.

[i] See §9.1 Definition and Examples in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Let G and H be groups with identities e_G and e_H, respectively. Let $\varphi : G \to H$ be a group isomorphism. Prove that $\varphi(g) = e_H$ if and only if $g = e_G$.
2. Let $G \cong H$. Show that if G is cyclic, then so is H.
3. Let $G \cong H$. Show that if G has a subgroup of order n, then so does H.
4. Let $|G| = p$ for some prime. Prove that $G \cong \mathbb{Z}_p$.
5. Prove that S_n is isomorphic to a subgroup of A_{n+2}.

"There are certain things whose number is unknown. If we count them by threes, we have two left over; by fives, we have three left over; and by sevens, two are left over. How many things are there?" – Sun Zi

Lecture Videos

Direct Products of Cyclic Groups

The Chinese Remainder Theorem

The Fundamental Theorem of Finite Abelian Groups

7.3 The Chinese Remainder Theorem

Theorem 7.3.1. *Let $(g,h) \in G \times H$. If $|g| = r$ and $|h| = s$ for $r, s \in \mathbb{Z}$, then $|(g,h)| = \text{lcm}(r,s)$.*

Proof. Let $m = \text{lcm}(r,s)$. Since $r \mid m$ we have $g^m = e$ in G, by Lagrange's Theorem. Likewise, $h^m = e$ in H. Thus, in $G \times H$, we have $(g,h)^m = (g^m, h^m) = (e,e)$. Suppose that $k < m$. Then either $r \nmid k$ or $s \nmid k$. Without the loss of generality, assume $r \nmid k$. Then $g^k \neq e$. Thus, $(g,h)^k = (g^k, h^k) \neq (e,e)$ since $g^k \neq e$. Thus, m is the smallest positive power of (g,h) which is equal to the identity of $G \times H$. Therefore, $|(g,h)| = m$. □

Corollary 7.3.2. *Let $(g_1, \ldots, g_n) \in \prod_{i=1}^{n} G_i$. Then $|(g_1, \ldots, g_n)| = \text{lcm}(|g_1|, \ldots, |g_n|)$.*

Proof. The proof follows by induction. □

Example 7.3.3. Let $(8, 56) \in \mathbb{Z}_{12} \times \mathbb{Z}_{60}$. Since $\gcd(8,12) = 4$, we have that $|8| = 12/4 = 3$ in $\mathbb{Z}_{12}$. Likewise, since $\gcd(56,60) = 4$, we have that $|56| = 60/4 = 15$. Thus, $\text{lcm}(3,15) = 15$ and $|(8,56)| = 15$ in $\mathbb{Z}_{12} \times \mathbb{Z}_{60}$.

Example 7.3.4. Let $(1,1) \in \mathbb{Z}_2 \times \mathbb{Z}_3$. Since $\gcd(1,2) = 1$, the order of 1 in $\mathbb{Z}_2$ is $|1| = 2/1 = 2$. Likewise, $\gcd(1,3) = 1$ and 1 has order 3 in $\mathbb{Z}_3$. Then $|(1,1)| = \text{lcm}(2,3) = 6$. But $|\mathbb{Z}_2 \times \mathbb{Z}_3| = 2 \cdot 3 = 6$. Since $\mathbb{Z}_2 \times \mathbb{Z}_3$ has an element with the same order as the group, it is cyclic. Since the order is 6, we can conclude that $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$. In fact, $n \mapsto (n,n)$ gives such an isomorphism from $\mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$.

Theorem 7.3.5 (The Chinese Remainder Theorem)**.**[i] *The group $\mathbb{Z}_n \times \mathbb{Z}_m$ is isomorphic to $\mathbb{Z}_{mn}$ if and only if $\gcd(n,m) = 1$.*

Proof. Since $|\mathbb{Z}_{nm}| = nm = |\mathbb{Z}_n \times \mathbb{Z}_m|$, it suffices to prove that $\mathbb{Z}_n \times \mathbb{Z}_m$ is cyclic if and only if $\gcd(n,m) = 1$ since cyclic groups of the same order are isomorphic. Let $G = \mathbb{Z}_n \times \mathbb{Z}_m$ and let $(k,\ell) \in G$. By Theorem 7.3.1, if $|k| = r$ and $|\ell| = s$ then $|(k,\ell)| = \text{lcm}(r,s)$. Since $r \mid n$ and $s \mid m$ by Lagrange's Theorem, $\text{lcm}(r,s) \mid \text{lcm}(n,m)$. Thus, the order of every element in G divides $\text{lcm}(n,m)$. If $\text{lcm}(n,m) < nm$, then G cannot have an element of order nm and hence cannot be cyclic. Since $\text{lcm}(n,m) = \dfrac{nm}{\gcd(n,m)}$,

$\operatorname{lcm}(n,m) < nm$ if and only if $\gcd(n,m) \neq 1$. If $\gcd(n,m) = 1$, then $\operatorname{lcm}(n,m) = nm$ and the order of $|(1,1)| = nm$ by Theorem 7.3.1. □

Corollary 7.3.6. *Let $n_1, \ldots, n_k$ be positive integers. Then*

$$\prod_{i=1}^{k} \mathbb{Z}_{n_i} \cong \mathbb{Z}_{n_1 \cdots n_k}$$

if and only if $\gcd(n_1, \ldots, n_k) = 1$.

Proof. The proof follows by induction. □

Corollary 7.3.7. *If n is a positive integer with prime factorization $n = \prod_{i=1}^{r} p_i^{e_i}$, where each p_i is a distinct prime, then*

$$\mathbb{Z}_n \cong \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}.$$

Corollary 7.3.8. *Let $n_1, \ldots, n_k$ be positive integers such that $\gcd(n_i, n_j) = 1$ for $i \neq j$. Then for any integers $a_1, \ldots, a_k$, the system of equations*

$$\begin{cases} x \equiv a_1 \pmod{n_1} \\ x \equiv a_2 \pmod{n_2} \\ \quad \vdots \\ x \equiv a_k \pmod{n_k} \end{cases}$$

has a unique solution modulo $n_1 \cdots n_k$.

Proof. Let $\varphi : \mathbb{Z}_{n_1 \cdots n_k} \to \prod_{i=1}^{k} \mathbb{Z}_{n_i}$ be an isomorphism. Since φ is surjective, there exists some $x \in \mathbb{Z}_{n_1 \cdots n_k}$ such that $\varphi(x) = (a_1, \ldots, a_k)$. Since φ is injective, x is the unique pre-image of $(a_1, \ldots, a_k)$. □

In number theory, the Chinese Remainder Theorem typically denotes the previous corollary, but in algebra, the Chinese Remainder Theorem refers to Theorem 7.3.5 and all its corollaries. The Chinese mathematician Sun Zi wrote about the theorem in the first century A.D., which is where the name derives.

Example 7.3.9. Solve the system of equations

$$\begin{cases} x \equiv 3 \pmod{4} \\ x \equiv 4 \pmod{5} \end{cases}$$

Since $x \equiv 3 \pmod 4$, there exists some $k \in \mathbb{Z}$ such that $x = 3 + 4k$. Substituting this into the second equation, we get

$$3 + 4k \equiv 4 \pmod 5.$$

Solving for k, we get

$$4k \equiv 1 \pmod 5 \quad \Rightarrow \quad k \equiv 4(1) = 4 \pmod 5$$

since $4^{-1} \equiv 4 \pmod 5$. Thus, let $x = 3+4(4+5\ell) = 3+16+20\ell = \boxed{19}+20\ell$. Checking our solution, note that $19 = 3+16 \equiv 3 \pmod 4$ and $19 = 4+15 \equiv 4 \pmod 5$.

Among other applications, the Chinese Remainder Theorem can be helpful in computer design in regards to calculations of large integers and parallel processing.

Theorem 7.3.10 (The Fundamental Theorem of Finite Abelian Groups)**.** *Let G be a finite abelian group. Then*

$$G \cong \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}},$$

where each p_i is a prime number not necessarily distinct.

Example 7.3.11. There are six abelian groups of order $540 = 2^2 \cdot 3^3 \cdot 5$, up to isomorphism and are isomorphic to one of the following:

$$(\mathbb{Z}_2 \times \mathbb{Z}_2) \times (\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3) \times \mathbb{Z}_5, \quad \mathbb{Z}_4 \times (\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3) \times \mathbb{Z}_5, \quad (\mathbb{Z}_2 \times \mathbb{Z}_2) \times (\mathbb{Z}_9 \times \mathbb{Z}_3) \times \mathbb{Z}_5,$$

$$\mathbb{Z}_4 \times (\mathbb{Z}_9 \times \mathbb{Z}_3) \times \mathbb{Z}_5, \quad (\mathbb{Z}_2 \times \mathbb{Z}_2) \times \mathbb{Z}_{27} \times \mathbb{Z}_5, \quad \mathbb{Z}_4 \times \mathbb{Z}_{27} \times \mathbb{Z}_5.$$

None of the above groups are isomorphic.

[i] It is common tradition in mathematics to name a new result by the author(s) who first discovered it. Unfortunately, our records of history are not always perfect and sometimes the name a theorem is given is instead attributed to someone (or something) else, perhaps to someone who first conjectured or popularized the theorem. Sometimes the original mathematician is only discovered after another name has been coined. As such, renaming a theorem to its proper name is rarely an effort that is truly successful because the traditional name is so entrenched in the history and tradition of mathematics. The name "the Chinese Remainder Theorem" is a victim of this entrenchment. The first historical record of the Chinese Remainder Theorem is found in the Chinese manuscript Sunzi Suanjing and is believed to be authored by the 3rd century Chinese mathematician Sun Zi. Very little of Sun Zi other than the authorship of Sunzi Suanjing (for sometime Sun Zi was erroneously identified with the Chinese general Sun Tsu who authored the Art of War). It is not known whether Sunzi Suanjing is an original scholarship of Sun Zi or just a compilation of important mathematical work at the time of Sun Zi, much like a college textbook. It is also not known whether Sunzi Suanjing was written all at once or if Sun Zi added to some previous version drafted by someone else. In the presence of these historical ambiguities, it would be most proper to name the Chinese Remainder Theorem instead "**Sun Zi's Theorem**" in the absence of a more precise attribution to the original scholar (if it was not, in fact, Sun Zi). The name "Chinese Remainder Theorem" as coined by L E Dickson, an American number theorist , in the early 20th century. He obtained an English translation brought from China by a missionary. Instead of attributing the author of book (who may have been unknown to Dickson at the time), it appears the result is attributed to its Chinese origin.

[ii] See §9.2 Direct Products, §13.1 Finite Abelian Groups, and §16.5 An Application to Software Design in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, find the order of the given element.

1. $(3, 4) \in \mathbb{Z}_4 \times \mathbb{Z}_6$
2. $(6, 15, 4) \in \mathbb{Z}_{30} \times \mathbb{Z}_{45} \times \mathbb{Z}_{24}$
3. $(5, 10, 15) \in \mathbb{Z}_{25} \times \mathbb{Z}_{25} \times \mathbb{Z}_{25}$
4. $(8, 8, 8,) \in \mathbb{Z}_{10} \times \mathbb{Z}_{24} \times \mathbb{Z}_{80}$

For Exercises 5–6, find all of the abelian groups of the given order, up to isomorphism.

5. 200
6. 720

For Exercises 7–8, determine whether there is a noncyclic abelian group of the given order.

7. 51
8. 52

For Exercises 9–12, solve the system of congruences.

9. $\begin{cases} x \equiv 2 \pmod 5 \\ x \equiv 6 \pmod{11} \end{cases}$
10. $\begin{cases} x \equiv 3 \pmod 7 \\ x \equiv 0 \pmod 8 \\ x \equiv 5 \pmod{15} \end{cases}$
11. $\begin{cases} x \equiv 2 \pmod 4 \\ x \equiv 4 \pmod 7 \\ x \equiv 7 \pmod 9 \\ x \equiv 5 \pmod{11} \end{cases}$
12. $\begin{cases} x \equiv 3 \pmod 5 \\ x \equiv 0 \pmod 8 \\ x \equiv 1 \pmod{11} \\ x \equiv 5 \pmod{13} \end{cases}$

"You can look for external sources of motivation and that can catalyze a change, but it won't sustain one. It has to be from an internal desire." – Jillian Michaels

Lecture Videos


The Frobenius Product of Subsets of a Group


The Counting Formula of Products of Subgroups


Internal Direct Products

7.4 Products of Subsets

Definition 7.4.1. Let H, K be subsets of a group G. Then define the **product** of the subsets H and K as

$$HK = \{hk \mid h \in H, k \in K\}.$$

This will be a subset of G. Note that cosets are a special case of this subset product, namely $gH = \{g\}H$.

As the multiplication of sets is defined element-wise in G, this multiplication is likewise associative. If G is abelian, the multiplication of subsets will likewise be commutative.

Example 7.4.2. Let $H = \langle s \rangle, K = \langle rs \rangle \leq D_4$. Then

$$HK = \{1, s\}\{1, rs\} = \{1, rs, s, srs\} = \{1, rs, s, r^3\}.$$

Conversely,

$$KH = \{1, rs\}\{1, s\} = \{1, s, rs, rss\} = \{1, s, rs, r\}.$$

Because multiplication of elements is noncommutative multiplication of sets is also likely to be noncommutative. It should also be mentioned that even if H and K is a subgroup, the subset HK is not necessarily a subgroup of G.

Example 7.4.3. Let $H = \{i, j, k\}, K = \{-i, -j, -k\} \subseteq Q_8$. Then

$$HK = \{1, -k, j, k, 1, -i, -j, i, 1\} = \mathbb{Q}_8 \setminus \{-1\} = KH.$$

Even though the group is nonabelian, it is still possible that the product of two sets commutes.

Example 7.4.4. Let $H = \{1, (123), (132)\}, K = \{1, (12)\} \leq S_3$. Then

$$HK = \{1, (12), (123), (123)(12), (132), (132)(12)\} = \{1, (12), (123), (13), (132), (23)\} = S_3 = KH.$$

Thus, it is even possible that a product of subsets can equal the entire group.

Theorem 7.4.5. *Let $H, K \leq G$ and $|H|, |K| < \infty$. Then*

$$|HK| = \frac{|H||K|}{|H \cap K|}.$$

Proof. Consider the map $\varphi : H \times K \to HK$ given by the rule $\varphi(h,k) = hk$. Clearly, φ is a well-defined, surjective function. It is not necessarily injective. Let $x \in H \cap K$. Then $\varphi(h,k) = hk = hek = h(xx^{-1})k = (hx)(x^{-1}k) = \varphi(hx, x^{-1}k)$ since $hx \in H$ and $x^{-1}k \in K$ for any $h \in H$ and $k \in K$. Furthermore, if $\varphi(h,k) = \varphi(h',k')$ for some $h, h' \in H$ and $k, k' \in K$ then $hk = h'k'$ and $(h'^{-1})h = k'k^{-1}$. Let $x = (h'^{-1})h = k'k^{-1}$. Then $x = (h'^{-1})h \in H$ and $x = k'k^{-1} \in K$, that is, $x \in H \cap K$. Then

$$(h,k) = (h'(h')^{-1}h, k(k')^{-1}k') = (h'(h')^{-1}h, (k'k^{-1})^{-1}k') = (h'x, x^{-1}k').$$

Therefore, the pre-image of hk is given as $\varphi^{-1}(\{hk\}) = \{(hx, x^{-1}k) \mid x \in H \cap K\}$. Therefore, $H \times K$ is partitioned by the pre-images of φ and each has the same cardinality $|H \cap K|$. This proves that

$$|HK| = \frac{|H \times K|}{|H \cap K|} = \frac{|H||K|}{|H \cap K|}.$$

□

Definition 7.4.6. Let G be a group with subgroups H and K. If $G = HK$, $H \cap K = \{e\}$, and $hk = kh$ for all $k \in K$ and $h \in H$, then we say that G is the **internal direct product** of H and K.

Example 7.4.7. Let $G = \mathbb{Z}_8^*$. Let $H = \langle 3 \rangle = \{1, 3\}$ and $K = \langle 5 \rangle = \{1, 5\}$. So, $H \cap K = \{1\}$. Also,

$$|HK| = \frac{2 \cdot 2}{1} = 4.$$

Since $|G| = 4$, we conclude that $HK = G$. Of course, since G is abelian, all elements of G commute with each other, especially the elements of H and K. Therefore $G = HK$ is an internal direct product.

Example 7.4.8. Let $G = D_6$. Let $H = \langle r^3 \rangle = \{1, r^3\}$ and let $K = \langle r^2, s \rangle = \{1, r^2, r^4, s, r^2s, r^4s\} \cong S_3$ (or D_3). Certainly, $H \cap K = \{1\}$. Then

$$|HK| = \frac{2 \cdot 6}{1} = 12.$$

Thus, $HK = G$. Of course, for any element $\pi \in D_6$, we have $1\pi = \pi = \pi 1$. Also, if $\pi = r^k$, then $r^3\pi = r^3r^k = r^{3+k} = r^{k+3} = r^kr^3 = \pi r^3$. If $\pi = r^k s$, then $r^3\pi = r^3r^k s = r^k r^3 s = r^k s r^{-3} = r^k s r^3 = \pi r^3$. Thus, H commutes with all elements of G, in particularly with the elements of K. Therefore, $D_6 = HK = \langle r^3 \rangle \langle r^2, s \rangle$ is an internal direct product.

Example 7.4.9. Let $G = S_3$, $H = \langle (123) \rangle$, and $K = \langle (12) \rangle$. In Example 7.4.4, we saw that $G = HK$ and $H \cap K = \{1\}$. But $(123)(12) = (13) \neq (23) = (12)(123)$. Thus, the elements of H and K do not commute with each other. Therefore, S_3 is not an internal direct product of H and K. In fact, S_3 is not an internal direct product of any of its subgroups.

Theorem 7.4.10. *Let G be the internal direct product of subgroups H and K. Then $G \cong H \times K$*

Proof. Since G is an internal direct product of H and K, we have $G = HK$. Let $\varphi : H \times K \to HK$ be defined as in Theorem 7.4.5. Thus, φ is a surjective function. Since $H \cap K = \{e\}$, we also can conclude

that φ is injective. Thus, φ is a bijection between $H \times K$ and G. Let $h, h' \in H$ and $k, k' \in K$. Then

$$\varphi(hh', kk') = (hh')(kk') = h(h'k)k' = h(kh')k' = (hk)(h'k') = \varphi(h, k)\varphi(h', k').$$

Therefore, φ is homomorphic since the elements of H and K commute with each other. Therefore, φ is an isomorphism $H \times K \to HK = G$. □

Internal direct products get their name because the formative groups come from inside the group G. On the other hand, the group $G = H \times K$, called the **external direct product**, can be created using any two groups H and K. Thus, G was formed from without.

Example 7.4.11. We have already seen that $\mathbb{Z}_6 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$. Let $H = \langle 3 \rangle$ and $K = \langle 2 \rangle$. Then $|H| = 2$, $|K| = 3$, and $|H \cap K| = 1$. Then $\mathbb{Z}_6$ is the internal direct product H and K. In essence, we were able to find isomorphic copies of the external groups $\mathbb{Z}_2$ and $\mathbb{Z}_3$ internally in $\mathbb{Z}_6$ as $\langle 3 \rangle$ and $\langle 2 \rangle$.

[i] See §9.2 Direct Products in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. For $H = \langle 2, 3\rangle \leq \mathbb{Q}^*$, prove that $H \cong \mathbb{Z} \times \mathbb{Z}$ using an internal direct product.
2. Prove that D_4 is not an internal direct product of any of its proper subgroups.
3. Let G be a group such that $H, K \leq G$ with $|G| = 20$, $|H| = 4$, and $|K| = 5$. If $hk = kh$ for all $h \in H$ and $k \in K$, prove that $G \cong H \times K$.
4. Let G be an internal direct product of subgroups H and K. Prove that $\varphi : G \to H \times K : \varphi(g) = (h, k)$ for $g = hk$, where $h \in H$ and $k \in K$. Prove that φ is an isomorphism.
5. If G is the internal direct product of $H_1, H_2, \ldots, H_n$, prove that $G \cong \prod_{i=1}^{n} H_i$.

7.5 Supplemenatry Exercises

(Go to Solutions)

1. Find five non-isomorphic groups of order 8.
2. Prove or disprove: Every abelian group of order divisible by 6 contains a subgroup of order 6.
3. Prove or disprove: Every nonabelian group of order divisible by 6 contains a subgroup of order 6.
4. Prove or disprove: If $G \times K \cong H \times K$, then $G \cong H$.
5. Prove that D_n is isomorphic to a subgroup of S_n.
6. Prove that if $G \cong H \cong H \times G$.
7. Prove that if $G \cong G'$ and $H \cong H'$, then $G \times H \cong G' \times H'$.
8. Find all of the abelian groups of order less than or equal to 40 up to isomorphism.
9. Let G be an abelian group of order n. If $d \mid n$, prove that G has a subgroup of order d.

Chapter 8

Quotient Groups and Homomorphisms

"The only normal people are the ones you don't know very well." – Alfred Adler

Lecture Videos

Normal Subgroups

Properties of Normal Subgroups

Conjugation

Conjugates in the Symmetric Group

8.1 Normal Subgroups

Definition 8.1.1. A subgroup H of a group G is called **normal** if $gH = Hg$ for all $g \in G$. We denote this as $H \trianglelefteq G$.

Let G be any group. Then $\{e\} \trianglelefteq G$ since $g\{e\} = \{g\} = \{e\}g$ for all $g \in G$. Likewise, $G \trianglelefteq G$. To see this, note that $gG = \text{im}(\lambda_g)$ where λ_g is the permutation on G induced by left multiplication on G. Since λ_g is surjective, we have that $\text{im}(\lambda_g) = G$. By similar reasoning, $Gg = \text{im}(\rho_{g^{-1}}) = G$, where $\rho_{g^{-1}}$ is right multiplication by $(g^{-1})^{-1}$. Therefore, $gG = G = Gg$.

Example 8.1.2. Let G be an abelian group. Since multiplication is commutative, $gh = hg$ for all $g \in G$ and $h \in H$ for some subgroup $H \leq G$. So, all left cosets are equal to their corresponding right cosets and every subgroup of G is normal!

One should be cautious about the converse of the above result. There do exist nonabelian group all of whose subgroups are normal. The reader will prove in the exercises that Q_8 is such a group.

Example 8.1.3. Let $G = S_3$ and $H = \langle(123)\rangle$. Then the left cosets of H are $H = \{1, (123), (132)\}$ and $(12)H = \{(12), (13), (23)\}$. Similarly, the right cosets of H are $H = \{1, (123), (132)\}$ and $H(12) = \{(12), (13), (23)\}$. Therefore, $H \trianglelefteq G$.

On the other hand, let $K = \langle(12)\rangle$. Then the left cosets of K are $K = \{1, (12)\}$, $(123)K = \{(123), (13)\}$, and $(132)K = \{(132), (23)\}$. On the other hand, the rights cosets of K are $K = \{1, (12)\}$, $K(123) = \{(123), (23)\}$, and $K(132) = \{(132), (13)\}$. Therefore, $K \not\trianglelefteq G$ since $(123)K \neq (132)K$.

Theorem 8.1.4. *Let G be a group with subgroup H. If $[G : H] = 2$ then $gH = Hg$ for all $g \in G$. In particular, $H \trianglelefteq G$.*

Proof. Let $x \in G \setminus H$. If $[G : H] = 2$, then there are two left cosets of H, namely H and $xH = G \setminus H$. Likewise, the two right cosets must be H and $Hx = G \setminus H$. Thus, $xH = G \setminus H = Hx$. Of course, if $y \in H$, then $yH = H = Hy$. Therefore, the left and right cosets agree. □

Example 8.1.5. The alternating group A_n is normal in S_n for all n.

Theorem 8.1.6. *Let G be a group and $N \leq G$. Then the following statements are equivalent.*

(i) $N \trianglelefteq G$.

(ii) *For all* $g \in G$, $gNg^{-1} \subseteq N$.

(iii) *For all* $g \in G$, $gNg^{-1} = N$.

Proof. We begin by proving $(i) \Rightarrow (ii)$. Let $x \in gNg^{-1}$. Then there exists some $n \in N$ such that $x = gng^{-1}$. Now, $gn \in gN$, so there exists some $n' \in N$ such that $gn = n'g$, by (i). Thus, $x = (gn)g^{-1} = (n'g)g^{-1} = n' \in N$. Therefore, $gNg^{-1} \subseteq N$. This proves $(i) \Rightarrow (ii)$.

We next prove $(ii) \Rightarrow (iii)$. Let $n \in N$. By (ii), we have that $g^{-1}N(g^{-1})^{-1} \subseteq N$. Thus, there exists some $n' \in N$ such that $g^{-1}ng = n'$. This gives $n = gn'g^{-1} \in gNg^{-1}$. Thus, $N \subseteq gNg^{-1}$. Since (ii) also gives $gNg^{-1} \subseteq N$, we conclude (iii).

Finally, we prove $(iii) \Rightarrow (i)$. We may assume that $gNg^{-1} = N$ for all $g \in G$. Let $x \in gN$, which means there exists some $n \in N$ such that $x = gn$. Of course, by (iii), there exists some $n' \in N$ such that $xg^{-1} = gng^{-1} = n'$. Thus, $x = n'g \in Ng$. We conclude that $gN \subseteq Ng$. A similar argument shows that $Ng \subseteq gN$, proving (i). □

Definition 8.1.7. Let $x, g \in G$. Then we call ${}^{g}x = gxg^{-1} \in G$ a **conjugate**[i] of x. Then set of all conjugates of x in G is denoted ${}^{G}x$ and is called the **conjugacy class** of x.

If y is a conjugate of x in G, we write $x \sim y$. This, of course, is an equivalence relation.

The study of conjugacy classes is an important part of group theory. For example, ${}^{g}x = x$ if and only if $gx = xg$, that is, x and g commute with each other. As such, ${}^{G}x = \{x\}$ for all $x \in G$ if and only if G is abelian. For another example, the **center** of a group, denoted $Z(G) = \{z \in G \mid zg = gz, \forall g \in G\}$, is a normal subgroup of G (to be proven by the reader) such that $Z(G) = \{z \in G \mid {}^{G}z = \{z\}\}$. Similarly, the commutator subgroup G', is another normal subgroup of G (to also be proven by the reader) generated by commutators $[g, x] = gxg^{-1}x^{-1} = {}^{g}xx^{-1}$.[ii] Both of these groups in way measure how close a group is to being abelian because they are constructed using conjugation.

Theorem 8.1.6.(ii) can be restated as the following: for all $g \in G$ and $x \in N$ we have that $gxg^{-1} \in N$. Thus, we can define normal subgroups as exactly the subgroups of G which are closed under conjugation. In other words, normal subgroups are subgroup which are unions of conjugacy classes.

Example 8.1.8. In S_3, the three conjugacy classes as

$$\{1\}, \{(123), (132)\}, \{(12), (13), (23)\}.$$

In D_4, the five conjugacy classes are

$$\{1\}, \{r^2\}, \{r, r^3\}, \{s, r^2s\}, \{rs, r^3s\}.$$

In A_4, the conjugacy classes are

$$\{1\}, \{(12)(34), (13)(24), (14)(23)\}, \{(123), (134), (142), (243)\}, \{(132), (143), (124), (234)\}.$$

Theorem 8.1.9. *Let $\sigma, \tau \in S_n$. Let $a \in \{1, 2, \ldots, n\}$. Then $\sigma\tau\sigma^{-1} : \sigma(a) \mapsto \sigma(\tau(a))$. In particular, if $\tau = (a_1 a_2 \ldots a_k)$ is a k-cycle, then*

$$\sigma\tau\sigma^{-1} = (\sigma(a_1)\sigma(a_2)\ldots\sigma(a_k)).$$

Furthermore, two permutations are conjugates if and only if they have the same cycle structure.

Proof. The first statement is immediate: $\sigma\tau\sigma^{-1}(\sigma(a)) = \sigma(\tau(a))$. Thus, if τ is a k-cycle, then $\sigma\tau\sigma^{-1}(\sigma(a_i)) = \sigma(\tau(a_i)) = \sigma(a_{i+1})$, proving the second statement. Finally, since conjugation preserves cycle length and $\sigma(\rho\tau)\sigma^{-1} = (\sigma\rho\sigma^{-1})(\sigma\tau\sigma^{-1})$, cycle structure is an invariant of conjugation. Let ρ and τ have the same cycle structure. Then there exists a bijection from the letter in the cycle decomposition of ρ to the letters in the cycle decomposition of τ that respects the cycles. This bijection is a permutation, say σ. Then $\sigma\rho\sigma^{-1} = \tau$. □

[ii]Similarly, we may denote conjugates as $x^g = g^{-1}xg$. The distinction between x^g and ${}^g x$ is known as the **right convention** and **left convention**, respectively. These different conventions give slightly different interpretations of conjugates but the exact same theory. For example, ${}^G x = x^G$, where the second set is defined analogously to ${}^G x$. The difference between left conjugation and right conjugation is analogous to the difference between left cosets and right cosets (or left translation versus right translation). There is no difference between the theory of left cosets and the theory of right cosets but some convention must be used. In this course, we use the left convention as all these notations are derived from the fact that we write functions on the left, that is, we write $f(x)$ instead of $(x)f = x^f$. Be aware though that there are many times where left and right conventions MUST interact. For example, normal subgroups are those subgroups for which LEFT cosets equal RIGHT cosets. More such interactions will be seen in the study of group actions. As such, the reader should be aware of both conventions.

[ii]Under the right convention, $[g, x] = g^{-1}x^{-1}gx$ and G' is defined analogously. Note that this is another instance where left and right interact. Let $[g, x]$ be the right commutator and $(gx)G'$ be the left coset. Then

$$(gx)G' = (gx)[x, g]G' = (gx)(x^{-1}g^{-1}xg)G' = (xg)G'.$$

Therefore, $(gx)G' = (xg)G'$, that is, the cosets "commute" even though g and x do not necessarily commute as elements.

[iii]See §10.1 Factor Groups and Normal Subgroups in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

Definition 8.1.10. Let G be a group and let $x, y \in G$. The **commutator** associated to x and y, denoted $[x, y]$, is given as

$$[x, y] := xyx^{-1}y^{-1}.$$

Note that $[x, y]$ measures the cost of commutation between x and y since $xy = [x, y]yx$. Note that $[x, y] = e$ if and only if $xy = yx$. Let $G' := \langle [x, y] \mid x, y \in G \rangle$ denote the **commutator subgroup** of G.

For Exercises 1–6, for the given G, compute the conjugacy classes of G, compute $Z(G)$, and compute G'. Recall $Z(G)$ is the center of a group G, as defined in Definition 2.4.8.

1. S_3
2. D_4
3. Q_8
4. $\mathbb{Z}_9$
5. D_5
6. A_4

For Exercises 7–11, prove the following statements about normal subgroups.
(Pick 2 from Exercises 7-9; Pick 1 from Exercise 10-11)

7. The intersection of two normal subgroups is a normal subgroup.
8. If a group has exactly one subgroup of order n, then that subgroup of order n is normal.

Definition. For a group G and $x \in G$, the set

$$C(x) = \{z \in G \mid xz = zx\},$$

is called the **centralizer** of x.[iv]

9. For all $x \in G$, $C(x) \leq G$. If $\langle x \rangle \trianglelefteq G$, then $C(x) \trianglelefteq G$.
10. For any group G, $Z(G) \trianglelefteq G$.
11. For any group G, $G' \trianglelefteq G$.

[iv] Recall that Definition 2.4.9 defines the centralizer of a subgroup $H \leq G$. This definition is a special case of that definition since $C(x) = C_G(\langle x \rangle)$.

"Whenever you're in conflict with someone, there is one factor that can make the difference between damaging your relationship and deepening it. That factor is attitude." – William James

Lecture Videos

Conjugacy Classes in the Dihedral Group

Quotient Groups

8.2 Quotient Groups

Theorem 8.2.1. *In D_n, the conjugacy classes for all rotations are of the form $\{r^k, r^{-k}\}$. For reflections, if n is odd, then there is a single conjugacy class; if n is even, then there are two conjugacy class: $\{r^k s \mid k = 2\ell\}$ and $\{r^k s \mid k = 2\ell + 1\}$.*

Proof. Consider r^k. Since all powers of r commute with each, we have $r^m r^k r^{-m} = r^k r^m r^{-m} = r^k$ for all $m \in \mathbb{Z}$. For reflections, we have

$$(r^m s)r^k(r^m s)^{-1} = r^m(sr^k s^{-1})r^{-m} = r^m(r^{n-k}ss^{-1})r^{-m} = r^m r^{n-k}r^{-m} = r^{n-k}r^m r^{-m} = r^{n-k}.$$

Thus, $(r^k)^{D_4} = \{r^k, r^{n-k}\}$.

For the conjugates of a reflection, note that

$$r^m(r^k s)r^{-m} = r^m r^k r^m s = r^{k+2m}s.$$

Similarly, $(r^m s)(r^k s)(r^m s)^{-1} = r^{n-k+2m}s$. Thus, $r^k s \sim r^\ell s$ if and only if $k \equiv \ell \pmod 2$ or $n - k \equiv \ell \pmod 2$. If n is even, then $k \equiv \ell$ if and only if $k + \ell \equiv n \equiv 0 \pmod 2$ and we get the two classes mentioned above. If n is odd, then $k + \ell \equiv n \equiv 1 \pmod 2$. Hence, $r^k s \sim r^\ell s$ if either k and ℓ have the same parity or not, that is, all reflections are conjugate to $r^k s$. □

In particular, $\langle r \rangle \trianglelefteq D_n$ since $\langle r \rangle$ is a union of conjugacy classes of the form $\{r^k, r^{-k}\}$. In other words, if $x \in D_n$ then $xr^k x^{-1} = r^k, r^{-k} \in \langle r \rangle$. Thus, the cyclic subgroup is normal.

Lemma 8.2.2. *Let $H \leq G$. Then $HH = H$.*

Proof. Let $x \in HH$. Then there exists some elements $h, h' \in H$ such that $x = hh'$. Since H is a subgroup, H is closed under multiplication and contains $x = hh'$. So, $HH \subseteq H$. Of course, for all $h \in H$, we have $h = he \in HH$. So, $H \subseteq HH$, which proves the lemma. □

Let $N \trianglelefteq G$ and let $a, b \in G$. Then $(aN)(bN) = a(Nb)N = a(bN)N = (ab)(NN) = (ab)N$. Thus, for a normal subgroup N, the product of two cosets is again a coset. This defines a multiplication of cosets, and we can use this multiplication on the set of cosets G/N to form a group.

Theorem 8.2.3. *Let $N \trianglelefteq G$. The set G/N with coset multiplication is a group.*

Proof. The associativity of multiplication of subsets comes from the fact that element-wise, multiplication is associative. It is also simple to check that $N = eN$ is the identity element and $g^{-1}N$ is the inverse of gN. Thus, G/N is a group, called the **factor group** (or **quotient group**) of G and H. It is pronounced

$G \bmod N$. □

Note that we have defined coset multiplication above using subset multiplication in G and not on the representative of the coset. This allows us to avoid a tedious well-defined argument that typically appears in the proof of the above theorem.

When considering the factor group G/N it is important to realize that the elements of G/N are cosets of N and not elements of G.

Example 8.2.4. Note that $A_3 \trianglelefteq S_3$. The cosets of A_3 are $A_3 = \{1, (123), (132)\}$ and $(12)A_3 = \{(12), (13), (23)\}$. The factor group S_3/A_3 is then a group of order 2 whose Cayley table is provided below. It is not hard to prove that $S_3/A_3 \cong \mathbb{Z}_2$. In fact, $S_n/A_n \cong \mathbb{Z}_2$ for all n.

	A_3	$(12)A_3$
A_3	A_3	$(12)A_3$
$(12)A_3$	$(12)A_3$	A_3

Example 8.2.5. Consider the normal subgroup $3\mathbb{Z}$ in $\mathbb{Z}$. The cosets of $3\mathbb{Z}$ are

$$3\mathbb{Z}, 1+3\mathbb{Z}, 2+3\mathbb{Z}.$$

The factor group $\mathbb{Z}/3\mathbb{Z}$ is then a group of order 3 whose Cayley table is provided below. It is not hard to prove that $\mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}_3$. In fact, $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$ for all n.

	$3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$
$3\mathbb{Z}$	$3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$
$1+3\mathbb{Z}$	$1+3\mathbb{Z}$	$2+3\mathbb{Z}$	$3\mathbb{Z}$
$2+3\mathbb{Z}$	$2+3\mathbb{Z}$	$3\mathbb{Z}$	$1+3\mathbb{Z}$

Example 8.2.6. Note $R = \langle r \rangle \trianglelefteq D_n$ since R is a union of conjugacy classes of the form $\{r^k, r^{-k}\}$. In other words, if $x \in D_n$ then $xr^kx^{-1} = r^k, r^{-k} \in R$. Thus, the cyclic subgroup is normal.

The cosets of R are the set of rotations (including the identity) and the set of reflections. Then D_n/R is a group of order two. Thus, $D_n/R \cong \mathbb{Z}_2$. Since $R \cong \mathbb{Z}_n$, we often denote this cyclic subgroup of D_n using $\mathbb{Z}_n$ (or Z_n). Then we say that $D_n/Z_n \cong \mathbb{Z}_2$.

[i] See §10.1 Factor Groups and Normal Subgroups in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–5, for the given G, determine whether the subgroup H is a normal subgroup of G. If so, compute the Cayley table for the quotient group G/H.

1. $A_4 \leq S_4$
2. $\langle (123) \rangle \leq A_5$
3. $D_4 \leq S_4$
4. $\langle i \rangle \leq Q_8$
5. $5\mathbb{Z} \leq \mathbb{Z}$

For Exercises 6–11, for the given G, determine the Hasse diagram of subgroups, determine which subgroups are normal, and compute G/H for each normal subgroup H.
(Pick 2 from Exercises 9-11)

6. $\mathbb{Z}_{12}$
7. $\mathbb{Z}_3 \times \mathbb{Z}_3$
8. S_3
9. D_4
10. Q_8
11. A_4

For Exercises 12–14, prove or disprove the given statement.
(Pick 2 from Exercises 12-14)

12. If G is cyclic and $H \trianglelefteq G$, then G/H is cyclic.
13. If G is abelian and $H \trianglelefteq G$, then G/H is abelian. If G is a group, $H \trianglelefteq G$, H is cyclic, and G/H is cyclic, then G is cyclic.
14. If G is a group, $H \trianglelefteq G$, H is abelian, and G/H is abelian, then G is abelian.

"If you keep feeling a point that has been sharpened, the point cannot long preserve its sharpness."
– Lao Tzu

Lecture Videos

Group Homomorphisms

Homomorphisms of Cyclic Groups

Properties of Group Homomorphisms

8.3 Homomorphisms

Definition 8.3.1. A map $\varphi : (G, *) \to (H, \circ)$ between groups G and H is a **homomorphism** if φ *preserves* the group operations, that is, for all $a, b \in G$,

$$\varphi(a * b) = \varphi(a) \circ \varphi(b).$$

In other words, a homomorphism is a function with the homomorphic property.

Note that isomorphisms are just bijective homomorphisms. While isomorphisms preserve the group structure exactly between the two groups, a homomorphism is a function that need only preserve part of the group structure.

Example 8.3.2. Many homomorphisms naturally occur in Linear Algebra. The determinant is a homomorphism of the form $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$ since $\det(AB) = \det(A)\det(B)$. Note that both groups are multiplicative.

Another example is the trace map $\mathrm{Tr} : \mathbb{R}^{n\times n} \to \mathbb{R}$ since $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$. Note that these groups are additive. On the other hand, the determinant is NOT a homomorphism from $\mathbb{R}^{n\times n}$ to $\mathbb{R}$ since the determinant does not preserve addition, that is, $\det(A + B) \neq \det(A) + \det(B)$. It is important to note that the map must preserve the operation of the domain group. The set of $n \times n$ real matrices $\mathbb{R}^{n\times n}$ is not a multiplicative group but instead an additive group.

Example 8.3.3. Let $Z_2 = \langle -1 \rangle \leq \mathbb{C}^*$. Note that $Z_2 \cong \mathbb{Z}_2$. Define $\mathrm{sgn} : S_n \to H$ by the rule $\mathrm{sgn}(\sigma) = \begin{cases} 1 & \sigma \text{ is even} \\ -1 & \sigma \text{ is odd} \end{cases}$, called the **signum** of σ. Then sgn is a homomorphism, which is simple to check. We note in the table that sgn preserves a part of the Cayley table of S_n. Although the Cayley table is much more complicated, if we forget all but the parity of a permutation, then permutation multiplication agrees with addition modulo 2. This is the idea behind homomorphisms. A part of the group structure is preserved.

	even	odd
even	even	odd
odd	odd	even

Please note that none of the above examples are isomorphisms (except for 1×1 matrices and S_2), but the homomorphisms are surjective. A surjective homomorphism is often called an **epimorphism** and it will be proven that the codomain is isomorphic to a factor group of the domain. An injective homomorphism is often called a **monomorphism** and it will be proven that the image is isomorphic to the domain.

Example 8.3.4. Let $d \mid n$ be integers. Define a map $\varphi : \mathbb{Z}_d \to \mathbb{Z}_n$ by the rule $\varphi(m) = mn/d \in \mathbb{Z}$. Of course, $\varphi(m + kd) = (m + kd)(n/d) = mn/d + kn \equiv mn/d \pmod{n}$. So this map is well-defined. It is also a homomorphism since

$$\varphi(a + b) = (a + b)n/d = a(n/d) + b(n/d) = \varphi(a) + \varphi(b).$$

In fact, this is a monomorphism. Note that $\text{im}(\varphi) = \langle n/d \rangle \cong \mathbb{Z}_d$.

Example 8.3.5. We can generalize that previous example. Let G be ANY group and let $g \in G$. Then we define $\varphi : \mathbb{Z} \to G$ as

$$\varphi(n) = g^n.$$

This is a homomorphism by the usual exponent laws. This map is a monomorphism if and only if g is infinite order. Similarly, we can define $\mathbb{Z}_d \to G$ by the rule $\varphi(n) = g^n$. This map is well-defined if and only if $d \mid |g|$. It likewise is a homomorphism and is injective if and only if $d = |g|$.

Example 8.3.6. We can define a homomorphism $\varphi : \mathbb{R} \to S^1$ (the circle group) by the rule $\theta \mapsto e^{i\theta}$. Again, usual exponent laws guarantee that this map is a homomorphism. It is also surjective.

Proposition 8.3.7. *Let $\varphi : G \to H$ be a homomorphism of groups. Then*

(i) If e is the identity of G, then $\varphi(e)$ is the identity of H.

(ii) For any element $g \in G$, $\varphi(g^{-1}) = \varphi(g)^{-1}$.

(iii) If $K \leq G$, then $\varphi(K) \leq H$. Furthermore, if K is normal in G then $\varphi(K)$ is normal in $\text{im}(\varphi) = \varphi(G)$ (but not necessarily normal in H).

(iv) If $K \leq H$, then $\varphi^{-1}(K) \leq G$. Furthermore, if K is normal in H then $\varphi^{-1}(K)$ is normal in G.

Proof. Let e' be the identity of H. Then

$$e'\varphi(e) = \varphi(e) = \varphi(ee) = \varphi(e)\varphi(e).$$

This gives $e' = \varphi(e)$ by cancellation in H. This proves (i).

Note that $\varphi(g)\varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(e) = e'$. Therefore, $\varphi(g)^{-1} = \varphi(g^{-1})$, by the uniqueness of inverses. This proves (ii).

Suppose $K \leq G$. Note that $\varphi(K) = \{\varphi(x) \mid x \in K\}$. Since $e \in K$ we have that $e' = \varphi(e) \in \varphi(K)$. Let $a, b \in \varphi(K)$. Then there exists some $x, y \in K$ such that $a = \varphi(x)$ and $b = \varphi(y)$. Thus,

$$ab = \varphi(x)\varphi(y) = \varphi(xy) \in \varphi(K)$$

since $xy \in K$. Finally, $a^{-1} = \varphi(x)^{-1} = \varphi(x^{-1}) \in \varphi(K)$ since $x^{-1} \in K$. Therefore, $\varphi(K) \leq H$.

If $\varphi(g) \in \text{im}(\varphi)$ and $\varphi(x) \in \varphi(K)$ where $g \in G$ and $x \in K$, then $\varphi(g)\varphi(x)\varphi(g)^{-1} = \varphi(gxg^{-1}) \in \varphi(K)$ since $gxg^{-1} \in K$. Therefore, $\varphi(K) \trianglelefteq \text{im}(\varphi)$, which proves (iii).

Suppose that $K \leq H$. Note that $\varphi^{-1}(K) = \{x \in G \mid \varphi(x) \in K\}$. Of course, $\varphi(e) = e' \in K$ so $e \in \varphi^{-1}(K)$. Next, let $x, y \in \varphi^{-1}(K)$. Then $\varphi(x), \varphi(y) \in K$ and $\varphi(x)\varphi(y) \in K$ since $K \leq H$. Then $\varphi(xy) = \varphi(x)\varphi(y) \in K$. So, $xy \in \varphi^{-1}(K)$. Finally, if $x \in \varphi^{-1}(K)$ then $\varphi(x) \in K$. But $= \varphi(x^{-1}) = \varphi(x)^{-1} \in K$ since $K \leq H$. Thus, $x^{-1} \in \varphi^{-1}(K)$. Therefore, $\varphi^{-1}(K) \leq G$.

Suppose next that $K \trianglelefteq H$. Let $x \in \varphi^{-1}(K)$ and $g \in G$. Then $\varphi(gxg^{-1}) = \varphi(g)\varphi(x)\varphi(g^{-1}) = \varphi(g)\varphi(x)\varphi(g)^{-1} \in K$ since $\varphi(x) \in K$, $\varphi(g) \in H$ and K is normal. Therefore, $gxg^{-1} \in \varphi^{-1}(K)$. This shows that $\varphi^{-1}(K) \trianglelefteq G$, which proves (iv). □

[i] See §11.1 Group Homomorphisms in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–2, prove that the given function between given groups is a group homomorphism.

1. $\varphi : \mathbb{R}^n \to \mathbb{R}^m : \boldsymbol{x} \mapsto A\boldsymbol{x}$ for a fixed $m \times n$ matrix A
2. $\varphi : G \to G : g \mapsto g^n$ for some fixed $n \in \mathbb{Z}$ and G is abelian

For Exercises 3–4, find all homomorphisms between the two given groups.
(Pick 1 from Exercise 3-4)

3. $\varphi : \mathbb{Z}_{24} \to \mathbb{Z}_{18}$
4. $\varphi : \mathbb{Z} \to \mathbb{Z}_{12}$

5. If $\varphi : G \to H$ is a group homomorphism and G is abelian, then $\varphi(G)$ is also abelian.
6. If $\varphi : G \to H$ is a group homomorphism and G is cyclic, then $\varphi(G)$ is also cyclic.

"Belief in oneself is incredibly infectious. It generates momentum, the collective force of which far outweighs any kernel of self-doubt that may creep in." – Aimee Mullins

Lecture Videos

Kernels

The First Isomorphism Theorem

8.4 Kernels

Definition 8.4.1. Let $\varphi : G \to H$ be a group homomorphism. Then the **kernel** of φ, denoted $\ker \varphi$, is the set $\ker \varphi = \{g \in G \mid \varphi(g) = e\}$.

In other words, $\ker \varphi = \varphi^{-1}(\{e\})$, the pre-image of the trivial subgroup of H. Since $\{e\} \trianglelefteq H$, we can conclude that $\ker \varphi \trianglelefteq G$. This offers a very convenient way to construct normal subgroups in a group G.

Example 8.4.2. We saw that the determinant is a homomorphism of the form $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^*$. The kernel of this homomorphism is the set of matrices with determinant equal to 1, the identity of $\mathbb{R}^*$. This is exactly the special linear group $\mathrm{SL}_n(\mathbb{R})$, which is necessarily a normal subgroup of $\mathrm{GL}_n(\mathbb{R})$.

Similarly, the kernel of the trace map $\mathrm{Tr} : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ is the set of **traceless matrices**, that is, $\ker \mathrm{Tr} = \{A \in M_{n\times n}(\mathbb{R}) \mid \mathrm{Tr}(A) = 0\}$. This forms a subgroup (or subspace) of dimension of $n^2 - 1$ by the Rank-Nullity Theorem.

Example 8.4.3. The kernel of the signum map $\mathrm{sgn} : S_n \to \{-1, 1\}$ is A_n, the alternating group, since even permutations maps to 1, the identity of $\{1, -1\}$. We have already established that $A_n \trianglelefteq S_n$.

Example 8.4.4. Let $\varphi : \mathbb{R} \to S^1$ be the homomorphism given by $\theta \mapsto e^{i\theta}$. Then the kernel of φ is given by the normal subgroup $\langle 2\pi \rangle \leq \mathbb{R}$. Of course, $\langle 2\pi \rangle \cong \mathbb{Z}$.

Example 8.4.5. Let $\varphi : \mathbb{Z} \to G$ as $\varphi(m) = g^m$, for some fixed $g \in G$. We have seen that this is a homomorphism. If the order of g is finite, say $|g| = n$, then any multiple of n will map to the identity since $\varphi(kn) = g^{kn} = (g^n)^k = e^k = e$. In fact, by Lagrange's Theorem, multiples of n are the only kernel elements of φ, that is, $\ker \varphi = n\mathbb{Z}$.

If $|g|$ is infinite, then $\ker \varphi = \{e\}$. Likewise, φ is a one-to-one map. This is not a coincidence as is illustrated in the below theorem whose proof is left to the reader.

Theorem 8.4.6. *Let $\varphi : G \to H$ be a group homomorphism. Then φ is injective if and only if $\ker \varphi = \{e\}$.*

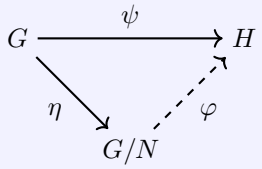
Definition 8.4.7. Let N be a normal subgroup of a group G. Then there is a homomorphism $\eta : G \to G/N$ given by the rule $g \mapsto gN$ for all $g \in G$. This is a homomorphism, since

$$\eta(gh) = (gh)N = (gN)(hN) = \eta(g)\eta(h),$$

called the **natural** (or **canonical**) **map**.

The natural map gets its name because it is really the only natural way to send an element of G into G/N without knowing anymore about the group structure of G. Note also that the kernel of the natural map is N. Thus, every kernel of a homomorphism is a normal subgroup but also every normal subgroup is the kernel of a homomorphism. In this way, normal subgroups and kernels of homomorphism represent the same collection of subsets of a group.

Theorem 8.4.8 (First Isomorphism Theorem)**.** *If $\psi : G \to H$ is a group homomorphism with $\ker \psi = N$, then there exists a unique homomorphism $\varphi : G/N \to H$ such that $\psi = \varphi \circ \eta$, where $\eta : G \to G/N$ is the natural map. In particular, $\varphi : G/N \to \mathrm{im}(\psi)$ is an isomorphism, that is, $G/N \cong \mathrm{im}(\psi)$.*



Proof. The construction of the map $\varphi : G/N \to H$ is also natural in the regard that there is really only one choice to define φ, that is, we define $\varphi(gN) = \psi(g)$. We first check that this is well-defined. Let $h \in gN$. Then there exists some $n \in N$ such that $h = gn$. Then

$$\varphi(hN) = \psi(h) = \psi(gn) = \psi(g)\psi(n) = \psi(g)e = \psi(g) = \varphi(gN).$$

Therefore, the map φ is well-defined. It is also a homomorphism since if $gN, hN \in G/N$ then

$$\varphi(gN \cdot hN) = \varphi(ghN) = \psi(gh) = \psi(g)\psi(h) = \varphi(gN)\varphi(hN).$$

It also has the property that

$$\psi(g) = \varphi(gN) = \varphi(\eta(g)) = \varphi \circ \eta(g)$$

for all $g \in G$. Thus, $\psi = \varphi \circ \eta$.

Suppose that $\chi : G/N \to H$ has the property that $\psi = \chi \circ \eta$. Then $\varphi(gN) = \psi(g) = \chi \circ \eta(g) = \chi(gN)$ for all $gN \in G/N$. Thus, $\varphi = \chi$, which shows that φ is the unique such map.

Finally, we show that φ is injective. Suppose that $\varphi(gN) = e$. But this implies that $\psi(g) = e$, that is, $g \in \ker \psi$. Since $N = \ker \psi$, we conclude that $g \in N$ and $gN = N$, the identity of G/N. Thus, $\ker \varphi = \{N\}$, which shows that φ is one-to-one. Since every function maps its domain onto its image, we conclude that $\varphi : G/N \to \mathrm{im}(\psi)$ is an isomorphism. □

The above diagram is referred to as a **commutative diagram** because all paths from G to H are equal. The dashed line implies that φ is uniquely constructed from ψ.

The First Isomorphism Theorem shows us that images of homomorphisms correspond to factor groups. It also provides a tool to show us when a $G/N \cong H$, exactly when an epimorphism $G \to H$ exists.

Example 8.4.9. Let $\mathbb{Z} \to \mathbb{Z}_n$ be given by the map $m \mapsto m \pmod n$. As seen before, this is a surjective homomorphism and the kernel of this map is $n\mathbb{Z}$. Therefore, $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$.

Example 8.4.10. Define $\psi : \mathbb{R} \to S^1$ using the map $\varphi(\theta) = e^{2\pi i\theta}$. Then this map is a surjective homomorphism with $\ker \psi = \mathbb{Z}$. Therefore, $\mathbb{R}/\mathbb{Z} \cong S^1$. Similarly, $\mathbb{Q}/\mathbb{Z}$ is isomorphic to the group of all roots of unity, which is a subgroup of S^1.

[i]See §11.1 Group Homomorphisms and §11.2 The Isomorphism Theorems in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–6, determine the kernel of the given homomorphism.

1. $\varphi : \mathbb{Z} \to \mathbb{Z} : n \mapsto 7n$
2. $\varphi : \mathbb{Z}_{18} \to \mathbb{Z}_{24} : n \mapsto 4n$
3. $\varphi : \mathbb{R}^* \to \mathrm{GL}_2(\mathbb{R}) : \boldsymbol{x} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$
4. $\varphi : \mathbb{R} \to \mathrm{GL}_2(\mathbb{R}) : \boldsymbol{x} \mapsto \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$
5. $\varphi : M_2(\mathbb{R}) \to \mathbb{R} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a + d$
6. $\varphi : M_2(\mathbb{R}) \to \mathbb{R} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto b$

7. Prove Theorem 8.4.6.
8. Given a group homomorphism $\varphi : G \to H$, define a relation $\sim$ on G such that $x \sim y$ if and only if $\varphi(x) = \varphi(y)$. Then $\sim$ is an equivalence relation on G and the equivalence classes are the cosets of $\ker \varphi$.
9. Prove or disprove: $\mathbb{Q}/\mathbb{Z} \cong \mathbb{Q}$
10. Let $\varphi : G_1 \to G_2$ be a surjective group homomorphism. Let $H_1 \trianglelefteq G_1$ and $H_2 = \varphi(H_1)$. Prove or disprove: $G_1/H_1 \cong G_2/H_2$.

"By three methods we may learn wisdom: First, by reflection, which is noblest; Second, by imitation, which is easiest; and third by experience, which is the bitterest." – Confucius

Lecture Videos

The Product of Subgroups is a Subgroup

The Second Isomorphism Theorem

The Third Isomorphism Theorem

The Correspondence Theorem

8.5 The Isomorphism Theorems

Proposition 8.5.1. *Let G be a group with subgroup H and normal subgroup N. Then $HN \leq G$. If $H \trianglelefteq G$, then $HN \trianglelefteq G$.*

Proof. Since $e \in H, N$, then $e = ee \in HN$. Let $x, y \in HN$. Then there exist $h, h' \in H$ and $n, n' \in N$ such that $x = hn$ and $y = h'n'$. Also, since N is normal, we have $h'N = Nh'$, that is, there exists some $n'' \in N$ such that $nh' = h'n''$. Then

$$xy = (hn)(h'n') = h(nh')n' = h(h'n'')n' = (hh')(n''n') \in HN,$$

since $hh' \in H$ and $n''n' \in N$. So, HN is closed under multiplication. Finally, since $h^{-1}N = Nh^{-1}$ there exists some $n''' \in N$ such that $n^{-1}h^{-1} = h^{-1}n'''$. Then

$$x^{-1} = (hn)^{-1} = n^{-1}h^{-1} = h^{-1}n''' \in HN,$$

which shows that HN contains inverse. Therefore, $HN \leq G$.

Suppose $H \trianglelefteq G$. Let $g \in G$. Then

$$gxg^{-1} = g(hn)g^{-1} = gh(g^{-1}g)ng^{-1} = (ghg^{-1})(gng^{-1}) \in HN,$$

since $ghg^{-1} \in H$ and $gng^{-1} \in N$. Therefore, HN is closed under conjugation. □

Theorem 8.5.2 (Second Isomorphism Theorem)**.** *Let G be a group with subgroup H and normal subgroup N. Then*

$$HN/N \cong H/(H \cap N).$$

Note that since $N \trianglelefteq G$, N contains all G-conjugates of elements from N. In particular, N contains all H-conjugates of elements from N. So, $N \trianglelefteq HN$. So, HN/N is well-defined. Similarly, we have shown previously that $H \cap N \trianglelefteq H$. So, $H/(H \cap N)$ is also well-defined.

Proof. We will prove the Second Isomorphism Theorem using the First Isomorphism Theorem. Let $K = H \cap N$, and define the map $\varphi : HN \to H/K$ by the rule $\varphi(hn) = hK$. Suppose that $h, h' \in H$ and $n, n' \in N$ such that $hn = h'n'$. Then we have that $(h')^{-1}h = n'(n^{-1})$, and if we let $x = (h')^{-1}h = n'(n^{-1})$, then $x \in H \cap N$. Thus,

$$\varphi(h'n') = h'K = (h'x)K = hK = \varphi(hn).$$

This shows that the map is well-defined.

Next we see that φ is homomorphic, namely

$$\varphi((hn)(h'n')) = \varphi(hh'n''n') = (hh')K = (hK)(h'K) = \varphi(hn)\varphi(h'n').$$

The map φ is clearly surjective by construction ($he \mapsto hK$). Finally, if $hn \mapsto K$, then $h \in K = H \cap N$, that is, $h \in N$. This implies that $\ker\varphi = N$ since also $n = en \mapsto eK = K$. Therefore, the First Isomorphism Theorem implies that $HN/N \cong H/K$. □

Example 8.5.3. Let $G = S_4$, $N = V_4$, and $H = \langle (123), (12) \rangle = S_3$. Then $|HN| = \dfrac{|H||N|}{|H \cap N|} = \dfrac{6 \cdot 4}{1} = 24$, since $H \cap N = \{1\}$. But the only subgroup of S_4 of order 24 is S_4 itself. Therefore, $\dfrac{S_4}{V_4} = \dfrac{HN}{N} \cong \dfrac{H}{H \cap N} = \dfrac{S_3}{\{1\}} \cong S_3$, by the Second Isomorphism Theorem.

Theorem 8.5.4 (Third Isomorphism Theorem)**.** *Let G be a group and H and N are normal subgroups of G with $N \leq H$. Then*

$$\frac{G/N}{H/N} \cong G/H.$$

Note that since $H \trianglelefteq G$, H contains all G-conjugates of elements from H, that is, $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$. Then $(gN)(hN)(gN)^{-1} = (ghg^{-1})N \in H/N$ since $ghg^{-1} \in H$. Thus, H/N contains all G/N-conjugates of elements from H/N. So, $H/N \trianglelefteq G/N$. So, $\dfrac{G/N}{H/N}$ is well-defined.

Proof. We will also prove the Third Isomorphism Theorem using again the First Isomorphism Theorem. Define a map $\varphi : G/N \to G/H$ by the rule $gN \mapsto gH$. Suppose that $x \in gN$, that is, there exists some $n \in N$ such that $x = gn$. Since $N \leq H$, we know that $n \in H$. Then

$$\varphi(xN) = xH = (gn)H = gH = \varphi(gN).$$

Therefore, φ is well-defined. It is also homomorphic since if $gN, xN \in G/N$ then

$$\varphi(gN \cdot xN) = \varphi((gx)N) = (gx)H = (gH)(xH) = \varphi(gN)\varphi(xH).$$

We can also see by construction that φ is surjective ($gN \mapsto gH$). Thus, it suffices to check that $\ker\varphi = H/N$. This is immediate since $\varphi(gN) = gH = H$ if and only if $g \in H$. Therefore, by the First Isomorphism Theorem, $\dfrac{G/N}{H/N} \cong G/H$. □

Example 8.5.5. The groups $\dfrac{\mathbb{Z}/mn\mathbb{Z}}{m\mathbb{Z}/mn\mathbb{Z}} \cong \mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}_m$ by the Third Isomorphism Theorem.

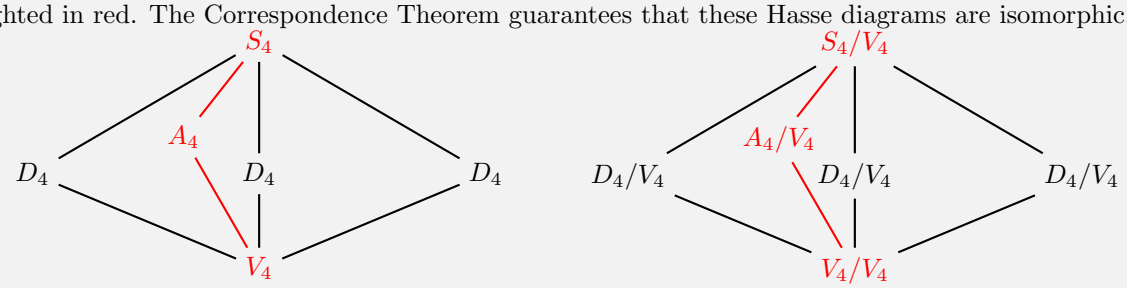
Example 8.5.6. Let $G = S_4$, $H = A_4$, and $N = V_4$. Then $G/H = S_4/A_4 \cong \mathbb{Z}_2$. Conversely, $G/N = S_4/V_4 \cong S_3$, as seen before, and $H/N = A_4/V_4 \cong A_3$, the only subgroup of order 3 in S_3. Then $\dfrac{G/N}{H/N} = \dfrac{S_4/V_4}{A_4/V_4} \cong \dfrac{S_3}{A_3} \cong \mathbb{Z}_2 \cong S_4/A_4$, which was to be expected by the Third Isomorphism Theorem.

The final theorem of this section, sometimes called the "Fourth Isomorphism Theorem," is called the Correspondence Theorem. It essentially says that homomorphisms preserve Hasse diagrams.

Theorem 8.5.7 (Correspondence Theorem)**.** *Let N be a normal subgroup of a group G. Then there is a one-to-one correspondence between the set of subgroups of G containing N and the set of subgroups of G/N, namely $H \mapsto H/N$ is a bijective map. Furthermore, this correspondence restricts to a one-to-one correspondence between the set of normal subgroups of G containing N and the set of normal subgroups of G/N.*

Proof. Let $\eta : G \to G/N$ be the natural map. Let $H \leq G$ such that $N \leq H$. Then $H/N \leq G/N$ since η is a homomorphism and, in fact, every subgroup of G/N is of this form since η is surjective. Therefore, the natural map establishes a correspondence between the subgroups of G containing N and the subgroups of G/N, namely $\eta : H \mapsto H/N$. Furthermore, $H = \eta^{-1}(H/N)$, the pre-image of H/N with respect to η. Since this correspondence has an inverse, it must be bijective. Therefore, η is a one-to-one correspondence. Finally, if $H/N \trianglelefteq G/N$, then we know that $H = \eta^{-1}(H/N) \trianglelefteq G$. Conversely, if $H \trianglelefteq G$, then $H/N = \eta(H) \trianglelefteq G/N$ since η is surjective. Thus, η restricts to a one-to-one correspondence on normal subgroups. □

Example 8.5.8. Note that $V_4 \trianglelefteq S_4$. There are six subgroups of S_4 which contain V_4, namely S_4, A_4, V_4 and three subgroups isomorphic to D_4: $\langle (1234), (12)(34) \rangle$, $\langle (1342), (13)(42) \rangle$, and $\langle (1423), (14)(23) \rangle$. The first three groups are normal while the three dihedral groups are not normal. The Hasse diagrams for S_4 and $S_4/V_4 \cong S_3$ are illustrated below, where normal subgroups are highlighted in red. The Correspondence Theorem guarantees that these Hasse diagrams are isomorphic.



[i] See §11.2 The Isomorphism Theorems in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Let $G = \mathbb{Z}_{24}$, $H = 4\mathbb{Z}_{24}$, and $N = 6\mathbb{Z}_{24}$. List all elements of $H + N$, $H \cap N$, $(H + N)/N$, and $H/(H \cap N)$. Give the correspondence between $(H + N)/N$ and $H/(H \cap N)$, as given in the 2nd Isomorphism Theorem.

2. Show that a group homomorphism whose domain is a cyclic group is completely determined by the image on a generator of the domain.

3. Let G be a group and $N \trianglelefteq G$. If $H \leq G/N$, prove that $\eta^{-1}(H) \leq G$ of order $|H||N|$, where $\eta : G \to G/N$ is the natural map.

4. Let G_1 and G_2 be groups such that $N_1 \trianglelefteq G_1$ and $N_2 \trianglelefteq G_2$. Let $\varphi : G_1 \to G_2$ be a group homomorphism. If $\varphi(N_1) \leq N_2$, Prove that there exists a homomorphism $\overline{\varphi} : G_1/N_1 \to G_2/N_2$ such that $\eta \circ \varphi = \overline{\varphi} \circ \eta$, where η is the natural map.

5. If $H, K \trianglelefteq G$ and $H \cap K = \{e\}$, prove that G is isomorphic to a subgroup of $G/H \times G/K$.

8.6 Supplemenatry Exercises

(Go to Solutions)

1. Let $G = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \middle| \, a, b, c \in \mathbb{R} \right\} \leq \mathrm{GL}_2(\mathbb{R})$. Let $H = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \middle| \, b \in \mathbb{R} \right\} \subseteq G$. Show that H is an abelian subgroup of G, $H \trianglelefteq G$, and $G/H \cong \mathbb{R}^* \times \mathbb{R}^*$.
2. If G is a group and $H \leq G$ such that $[G : H] = 2$, then $H \trianglelefteq G$.
3. Prove or disprove: If G is a group, $H \trianglelefteq G$, H is abelian, and G/H is abelian, then G is abelian.
4. Prove that $\det : \mathrm{GL}_2(\mathbb{R}) \to \mathbb{R}^*$ is a group homomorphism. What is its kernel?

Chapter 9

Representation Theory

"Your expectations opens or closes the doors of your supply, If you expect grand things, and work honestly for them, they will come to you, your supply will correspond with your expectation." – Orison Swett Marden

Lecture Videos

The Classical Matrix Groups | The Orthogonal Group | The Unitary Group

9.1 The Classical Matrix Groups

In this section, we will present the classical matrix groups. We have seen some already, namely the *general linear groups* $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{GL}_n(\mathbb{C})$, the groups of nonsingular $n \times n$ matrices with real or complex scalar with matrix multiplication. We have seen already that since det is a homomorphism from the general linear group, that is,

$$\det(AB) = \det(A)\det(B),$$

its kernel is a normal subgroup, namely the *special linear groups* $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{SL}_n(\mathbb{C})$.

Recall from linear algebra the *transpose* operator $^\top : A = (a_{ij}) \mapsto A^\top = (a_{ji})$. The following properties are also standard from linear algebra:

(i) $(A+B)^\top = A^\top + B^\top$

(ii) $(rA)^\top = rA^\top$

(iii) $(A^\top)^\top = A$

(iv) $(AB)^\top = B^\top A^\top$

(v) $\det(A^\top) = \det(A)$

Definition 9.1.1. Define the set $O(n)$ (or $O_n(\mathbb{R})$), called the **orthogonal group**, as the subset of *orthogonal matrices*, those invertible real matrices whose inverse is equal to its transpose. In other words,

$$O(n) = \{Q \in \mathrm{GL}_n(\mathbb{R}) \mid Q^\top = Q^{-1}\}.$$

It is an exercise of the reader to show that $O(n) \leq \mathrm{GL}_n(\mathbb{R})$.

Note that if $Q^{-1} = Q^\top$, then it must be that $QQ^T = QQ^{-1} = I_n$. Likewise, $Q^\top Q = I_n$.

Theorem 9.1.2. *A square matrix U is orthogonal if and only if its column vectors form an orthonormal set.*

Proof. Let U be an $n \times n$ orthogonal matrix and let $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ be the column vectors of U. Then

$$U^\top U = (u_{ij})^\top (u_{ij}) = (\boldsymbol{u}_i^\top \boldsymbol{u}_j) = (\boldsymbol{u}_i \cdot \boldsymbol{u}_j).$$

Therefore, $U^\top U = I_n$ if and only if $\boldsymbol{u}_i \cdot \boldsymbol{u}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$ if and only if the columns of U are an orthonormal set. □

Since $U^\top = U^{-1}$ and $UU^\top = UU^{-1} = I$, it also follows that the row vectors of an orthogonal matrix U must also form an orthonormal set.

Geometrically speaking, the orthogonal group $O(n)$ consists of all rotation and reflection matrices of $\mathbb{R}^n$.

Example 9.1.3. The matrix

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthogonal matrix. Note that

$$U^\top U = \begin{bmatrix} 3/\sqrt{11} & 1/\sqrt{11} & 1/\sqrt{11} \\ -1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{66} & -4/\sqrt{66} & 7/\sqrt{66} \end{bmatrix} \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

$$= \begin{bmatrix} (9+1+1)/11 & (-3+2+1)/\sqrt{66} & (-3-4+7)/\sqrt{726} \\ (-3+2+1)/\sqrt{66} & (1+4+1)/6 & (1-8+7)/\sqrt{396} \\ (-3-4+7)/\sqrt{726} & (1-8+7)/\sqrt{396} & (1+16+49)/66 \end{bmatrix} = I_3$$

Theorem 9.1.4. *Let U be an orthogonal matrix, and let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Then*

$$(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y},$$

that is, the matrix transformation $\boldsymbol{x} \mapsto \boldsymbol{U}\boldsymbol{x}$ preserves inner products.

Proof.

$$(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = (U\boldsymbol{x})^\top (U\boldsymbol{y}) = (\boldsymbol{x}^\top U^\top)(U\boldsymbol{y}) = \boldsymbol{x}^\top (U^\top U)\boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}. \qquad \square$$

Since multiplication by an orthogonal matrix U preserves dot product, as a consequence, it also preserves lengths, distances, angles, and orthogonality of vectors. For example, $\|U\boldsymbol{x}\| = \|\boldsymbol{x}\|$ for any vector $\boldsymbol{x}$ and any orthogonal matrix U. In particular, orthogonal matrices are exactly those linear transformations of $\mathbb{R}^n$ that preserve lengths, called a **rigid motion**.

Let $SO(n)$ (or $SO_n(\mathbb{R})$), called the **special orthogonal group**, as the subset of orthogonal matrices with determinant equal to 1, that is, $SO(n) = O(n) \cap \mathrm{SL}_n(\mathbb{R})$. This implies that $SO(n) \trianglelefteq O(n)$. It can be shown that $\det(Q) = \pm 1$ for all $Q \in O(n)$. A consequence of this fact is that $[O(n) : SO(n)] = 2$, which also shows that it is normal in the orthogonal groups.

For complex matrices, we define the notion of the "orthogonal" group differently. In particular, the dot product $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^\top \boldsymbol{v}$ does not quite work the correct way anymore. For example, $\begin{bmatrix} 1 \\ i \end{bmatrix} \cdot \begin{bmatrix} 1 \\ i \end{bmatrix} = 1 - 1 = 0$ but $\begin{bmatrix} 1 \\ i \end{bmatrix} \neq \mathbf{0}$. To repair this problem, the dot product on complex spaces is replaced with the *Hermitian product.*

Definition 9.1.5. Let A be an $m \times n$ complex matrix. Then we define $A^* = (\overline{A})^\top$, which is called the **conjugate transpose**. This replaces the role of transposes in complex space.

Example 9.1.6. Let $A = \begin{bmatrix} 1+i & -i & 0 \\ 2 & 3-2i & i \end{bmatrix}$ and $B = \begin{bmatrix} 1 & i & 1+i \\ -i & -5 & 2-i \\ 1-i & 2+i & 3 \end{bmatrix}$.

Note that

$$A^* = \begin{bmatrix} 1-i & 2 \\ i & 3+2i \\ 0 & -i \end{bmatrix}, \qquad B^* = \begin{bmatrix} 1 & i & 1+i \\ -i & -5 & 2-i \\ 1-i & 2+i & 3 \end{bmatrix}.$$

Definition 9.1.7. We define the **Hermitian product**[i] $\cdot : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ of two vectors, $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ by the rule

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^* \boldsymbol{v} = \begin{bmatrix} \overline{u_1} & \overline{u_2} & \dots & \overline{u_n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \overline{u_1} v_1 + \overline{u_2} v_2 + \ldots + \overline{u_n} v_n.$$

The concepts of vector norms (lengths), vector distance, unit vectors, normalizations, angles, etc. that were defined via the dot product can naturally be extended to complex vector space via the Hermitian product.

Example 9.1.8. Let $\boldsymbol{u} = (1+i,\ i,\ 3-i)$ and $\boldsymbol{v} = (1+i,\ 2,\ 4i)$. Find $\boldsymbol{u} \cdot \boldsymbol{v}$, $\boldsymbol{v} \cdot \boldsymbol{u}$, and $\|\boldsymbol{u}\|$.

$\boldsymbol{u} \cdot \boldsymbol{v} = (\overline{1+i})(1+i) + \overline{i}(2) + (\overline{3-i})(4i) = (1-i)(1+i) - 2i + (3+i)(4i) = 2 - 2i + 12i - 4 = \boxed{-2+10i}$

$\boldsymbol{v} \cdot \boldsymbol{u} = (\overline{1+i})(1+i) + \overline{2}(i) + (\overline{4i})(3-i) = (1-i)(1+i) + 2i - (4i)(3-i) = 2 + 2i - 12i - 4 = \boxed{-2-10i}$

$$\|\boldsymbol{u}\| = \sqrt{(\overline{1+i})(1+i) + (\overline{i})i + (\overline{3-i})(3-i)} = \sqrt{(1-i)(1+i) + (-i)i + (3+i)(3-i)}$$
$$= \sqrt{2+1+10} = \boxed{\sqrt{13}}$$

Theorem 9.1.9. *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{C}^n$ and let $z \in \mathbb{C}$. Then*

(i) $\boldsymbol{u} \cdot \boldsymbol{v} = \overline{\boldsymbol{v} \cdot \boldsymbol{u}}$;

(ii) $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$;

(iii) $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$;

(iv) $(z\boldsymbol{u}) \cdot \boldsymbol{v} = \overline{z}(\boldsymbol{u} \cdot \boldsymbol{v})$;

(v) $\boldsymbol{u} \cdot (z\boldsymbol{v}) = z(\boldsymbol{u} \cdot \boldsymbol{v})$;

(vi) $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$, *and* $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ *if and only if* $\boldsymbol{u} = \boldsymbol{0}$.

Definition 9.1.10. Define the set $U(n)$ (or $U_n(\mathbb{C})$), called the **unitary group**, as the subset of *unitary matrices*, those invertible complex matrices whose inverse is equal to its conjugate transpose. In other words,

$$U(n) = \{U \in \mathrm{GL}_n(\mathbb{C}) \mid Q^* = Q^{-1}\}.$$

Note that $U \in U(n)$ if and only if $U^*U = UU^* = I_n$.

The **special unitary group** $SU(n)$ is the subgroup of $U(n)$ whose determinants are 1.

Example 9.1.11. Let $U = \begin{bmatrix} \frac{1}{2}(1+i) & \frac{1}{2}(1+i) \\ \frac{1}{2}(1-i) & \frac{1}{2}(-1+i) \end{bmatrix}$.

Note that

$$UU^* = \begin{bmatrix} \frac{1}{2}(1+i) & \frac{1}{2}(1+i) \\ \frac{1}{2}(1-i) & \frac{1}{2}(-1+i) \end{bmatrix} \begin{bmatrix} \frac{1}{2}(1-i) & \frac{1}{2}(1+i) \\ \frac{1}{2}(1-i) & \frac{1}{2}(-1-i) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}+\frac{1}{2} & \frac{i}{2}-\frac{i}{2} \\ -\frac{i}{2}+\frac{i}{2} & \frac{1}{2}+\frac{1}{2} \end{bmatrix} = I_2$$

Therefore, U is unitary.

Analogously, unitary matrices are exactly the complex matrices that preserve Hermitian products, and thus also preserve lengths and distances. In other words, $U(n)$ is the group of rigid motions of $\mathbb{C}^n$.

[i] Many textbooks alternatively define the Hermitian product as $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^\top \overline{\boldsymbol{v}}$. Although this does not at all change the theory and applications of the Hermitian product, it does change intermediate calculations.

[ii] See §12.1 Matrix Groups in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–4, determine whether the matrix A belongs to $O(n)$ or $SO(n)$.

1. $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$

2. $\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$

3. $\begin{pmatrix} \frac{4}{5} & 0 & \frac{3}{5} \\ -\frac{3}{5} & 0 & \frac{4}{5} \\ 0 & -1 & 0 \end{pmatrix}$

4. $\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$

5. Prove that $O(n)$ is a group.

6. Prove that $SU(n) \trianglelefteq U(n)$.

7. Prove or disprove: there exists an infinite abelian subgroup of $O(n)$.

8. Let $\boldsymbol{x} \in \mathbb{R}^n$ such that $\|\boldsymbol{x}\| = 1$, that is, $\boldsymbol{x} \in S^{n-1}$, where S^{n-1} is the unit $n-1$ sphere. If $U \in O(n)$, prove that $U\boldsymbol{x} \in S^{n-1}$.

'Never mistake motion for action." – Ernest Hemingway

Lecture Videos

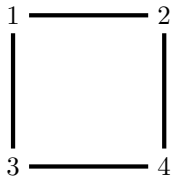
Symmetries

Symmetries of Euclidean Space $\mathbb{R}^n$

9.2 Isometries

Recall that a **symmetry** $\sigma : X \to X$ is a permutation which preserves the underlying structure of the object X. If X is a set, then the the symmetry group S_X is the collection of all permutations from X into itself. If X has no other structure, this captures all the symmetries of the object X. For example, if $X = \{1, 2, 3, 4\}$, then the symmetry group of X is just S_4.

But maybe X does have some other structure, for example say that $X = \{1, 2, 3, 4\}$ but these four points are the vertices of a square. Then we will not consider all permutations of X but only those permutations which also preserve the adjacency relations of vertices, that is, those permutations that preserve the shape of the square. Of course, only the permutations in D_4 preserve the square-ness of X.



For another example, consider $X = \{1, 2, 3, 4\}$ as a group with the following Cayley table. Of course, $X \cong \mathbb{Z}_4$. In this situation, the symmetries of X as a group are all permutations $\varphi : X \to X$ such that φ preserves the group operation, that is, $\varphi(ab) = \varphi(a)\varphi(b)$. Notice that this implies that φ is an isomorphism from X to itself. Such an isomorphism is called an **automorphism**. Viewing X as a group, the symmetry group of X becomes $\langle(24)\rangle = \{1, (24)\}$.

	1	2	3	4
1	1	2	3	4
2	2	3	4	1
3	3	4	1	2
4	4	1	2	3

Thus, when we view X as just a set, its symmetry group is S_4. When we view X as a square, its symmetry group is D_4, a subgroup of S_4. When we view X as a cyclic group, its symmetry group is $\mathbb{Z}_2$, another subgroup of S_4. In this manner, when we consider the symmetry group of an object we need to focus not just on the elements which create the object but also the relations between the elements that capture the aspects we are focusing on.

The Euclidean space $\mathbb{R}^n$ is a beautifully complicated mathematical object, with a deep interplay between both algebra and topology. Algebraically, $\mathbb{R}^n$ is a vector space with vector addition and scalar multiplication. When given the dot product, $\mathbb{R}^n$ becomes an inner product, a stronger algebraic object. But this inner product leads toward the geometry of $\mathbb{R}^n$. With the dot product, we can discuss lengths, distances, and angles. This makes $\mathbb{R}^n$ into a metric space, which forms a topology on $\mathbb{R}^n$ by which we can start discussing convergent sequences, Cauchy sequences, completeness, compactness, limits, derivatives, etc. Based upon the focus we have on $\mathbb{R}^n$, we will discuss the symmetries of this great object.

Example 9.2.1. Viewing $\mathbb{R}^n$ as a vector space, the symmetries of $\mathbb{R}^n$ will be those permutations of $\mathbb{R}^n$ which preserve the vector space structure, namely vector addition and scalar multiplication. Maps that preserve vector addition and scalar multiplications are just *linear transformations* from linear algebra and they are the analog of group homomorphisms for vector space. It is a common

fact from linear algebra that any linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ can be represented as an $m \times n$ uniquely using the standard bases. If this linear transformation is bijective, since it is a permutation, then the linear transformation must both be one-to-one and onto. This implies that the standard matrix is nonsingular. Therefore, $\mathrm{GL}_n(\mathbb{R})$ is symmetry group of $\mathbb{R}^n$ with respect to its vector space structure.

Example 9.2.2. Viewing $\mathbb{R}^n$ as an inner product space, the symmetries of $\mathbb{R}^n$ will be those permutations of $\mathbb{R}^n$ which preserve the inner product structure, namely vector addition, scalar multiplication and the inner product. As mentioned before, this symmetry must be an element of $\mathrm{GL}_n(\mathbb{R})$. As we discussed earlier, a matrix A preserves all dot products if and only if A is orthogonal. Therefore, $O(n)$ is symmetry group of $\mathbb{R}^n$ with respect to its inner product space structure. Geometrically speaking, the $O(n)$ consists of all *rotations* about the origin and *reflections* through the origin or across lines through the origin.

Example 9.2.3. Because orthogonal matrices preserve dot products, lengths and distances are also preserved. When viewing $\mathbb{R}^n$ as a metric space, our symmetries are all those permutations of $\mathbb{R}^n$ that preserve distance, which are called **isometries**. Of course, all orthogonal matrices are isometries of $\mathbb{R}^n$, but orthogonal matrices preserve the whole algebraic structure of $\mathbb{R}^n$. A general isometry only needs to preserve the geometric (and topological) structure of $\mathbb{R}^n$.

Rotations about any point and *reflections* across any line are examples of isometries. These are not linear transformations, and hence not orthogonal matrices, if the point of rotation is not the origin or if the line of reflection does not pass through the origin. Another isometry is *translation* which maps the point $\boldsymbol{x}$ to $\boldsymbol{x} + \boldsymbol{v}$ for some fixed vector $\boldsymbol{v}$. Translations are linear transformations if and only if $\boldsymbol{v} = \boldsymbol{0}$, in which case the translation is just the identity map. These three types of isometries generate the isometry group of $\mathbb{R}^n$.

Define the set $E(n)$ as the Cartesian product of $\mathbb{R}^n$ and $O(n)$, namely $E(n) = \mathbb{R}^n \times O(n)$. We will turn $E(n)$ into a group but not by using the direct product multiplication. Instead, if $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ and $A, B \in O(n)$, then we define multiplication by the rule:

$$(\boldsymbol{u}, A) \circ (\boldsymbol{v}, B) = (\boldsymbol{u} + A\boldsymbol{v}, AB) \in E(n).$$

The multiplication forms a group structure on $E(n)$, called the **Euclidean group**. It is left as an exercise to the reader to prove that $E(n)$ is a group.

The Euclidean group $E(n)$ forms the group of distance-preserving-symmetries of $\mathbb{R}^n$ in the following way: let $\boldsymbol{x}, \boldsymbol{v} \in \mathbb{R}^n$ and $A \in O(n)$. Then we define $(\boldsymbol{v}, A) : \mathbb{R}^n \to \mathbb{R}^n$ by the rule $(\boldsymbol{v}, A)\boldsymbol{x} = \boldsymbol{v} + A\boldsymbol{x} \in \mathbb{R}^n$. In particular, $(\boldsymbol{v}, I)$ is simple translation by the vector $\boldsymbol{v}$ and $(\boldsymbol{0}, A)$ is simple multiplication by the orthogonal matrix A.

In $\mathbb{R}^2$, $E(2)$ consists of all rotations around a point, reflections across a line, translations, and glide reflections, which is a composite of a translations and a reflection, namely $(\boldsymbol{v}, A)$ where $\boldsymbol{v} \in \mathbb{R}^n$ and A is a reflection matrix. All other elements of $E(2)$ can be expressed as one of these four types. In $\mathbb{R}^3$, the Euclidean group $E(3)$ consists of seven types: translations, rotations around a line, reflections through a point, reflections across a plane, glide reflections, screw displacements (a composite of a translation and a rotation), and improper rotations (rotation around an axis and reflection through the orthogonal complement).

Example 9.2.4. Finally, the topology of $\mathbb{R}^n$ is a weaker structure than the geometry of $\mathbb{R}^n$, that is, a permutation can preserve closeness between two points without necessarily preserving distance. These describe continuous permutations. Define $\mathrm{Aff}(n)$, called the **Affine group**, to be the Cartesian product $\mathrm{Aff}(n) = \mathbb{R}^n \times \mathrm{GL}_n(\mathbb{R})$[i] and multiplication as the rule:

$$(\boldsymbol{u}, A) \circ (\boldsymbol{v}, B) = (\boldsymbol{u} + A\boldsymbol{v}, AB) \in \mathrm{Aff}(n).$$

This forms a group by the same reasoning as the Euclidean group. In fact, $E(n) \leq \mathrm{Aff}(n)$. In addition to the types of transformations mentioned above in $E(n)$, $\mathrm{Aff}(n)$ also contains stretches/-compressions/dilations and shears, a linear map that displaces each point in fixed direction, by an amount proportional to its signed distance from a hyperplane that is parallel to that direction.

Example 9.2.5. Is $\mathrm{SL}_n(\mathbb{R})$ a symmetry group for $\mathbb{R}^n$? Yes, in a way. If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. If A is 3×3, the volume of the parallelepiped determined by the columns of A is $|\det A|$. The higher dimensional analogues also hold. It is left as an exericse that $\mathrm{SL}_n(\mathbb{R})$ is the symmetry group of $\mathbb{R}^n$ which preserves areas, but does not necessarily preserve distance.

[i] Both $E(n)$ and $\mathrm{Aff}(n)$ give examples of **semi-direct product**, which is an alternative multiplication placed upon a Cartesian product.

[ii] See §12.2 Symmetry in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

1. Prove that $E(n)$ is a group for all $n \geq 1$.
2. Prove or disprove: there exists some $(A, \boldsymbol{x}) \in E(2)$ where $\boldsymbol{x} \neq \mathbf{0}$ and $|(A, \boldsymbol{x})| = \infty$.

Definition 9.2.6. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$. Then we say a linear combination $c_1\boldsymbol{x}_1 + \ldots + c_k\boldsymbol{x}_k$ is a **convex combination** if $c_i \geq 0$ for $1 \leq i \leq k$ and $c_1 + \ldots + c_k = 1$.

Let $X \in \mathbb{R}^n$. Then the **convex span** (or **convex hull**), denoted $\text{ConSpan}(X)$, is the set of all convex combinations involving only finitely many vectors in X. Note that this is the smallest convex subset of $\mathbb{R}^n$ containing X.

We say that $X \subseteq \mathbb{R}^n$ is a **hyper-parallelepiped**[iii] if $X = \text{ConSpan}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ for some vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^n$. If A is the $n \times n$ matrix whose columns are $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, then the **hyper-volume** of the hyper-parallelepiped X is $|\det(A)|$.

3. Prove that $\text{SL}_n(\mathbb{R})$ preserves the hyper-volume of each hyper-parallelepiped in $\mathbb{R}^n$.

Definition 9.2.7. Let G be a group with normal subgroup H and (not necessarily normal) subgroup K. If $G = HK$ and $H \cap K = \{e\}$, then we say that G is the **internal semidirect product** of H and K, denoted $G = H \rtimes K$.

4. Prove that $S_3 = A_3 \rtimes \langle(12)\rangle$.
5. Prove that $\text{Aff}(n) = R^n \rtimes \text{GL}_n(\mathbb{R})$.
6. Prove that Q_8 is not a semidirect product of any of its subgroups.

[iii] In dimensions $n = 1, 2, 3$, a hyper-parallelepiped is instead called a **segment**, **parallelogram**, and **parallelepiped**, and its hyper-volume is instead called **length**, **area**, and **volume**, respectively.

9.3 Supplemenatry Exercises

(Go to Solutions)

1. Prove Theorem 9.1.9.

2. For all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$, prove
$$\boldsymbol{u} \cdot \boldsymbol{v} = \frac{1}{2}\left(\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}\|^2 - \|\boldsymbol{v}\|^2\right).$$

3. Suppose that $G \leq E(2) = \mathbb{R}^2 \rtimes O(2)$ and let $T = G \cap \mathbb{R}^2$, where T is the translation subgroup of G. Prove that $G/T \cong G \cap O(2)$.

Chapter 10

Rings

"A bell's not a bell 'til you ring it, A song's not a song 'til you sing it,
Love in your heart wasn't put there to stay, Love isn't love 'til you give it away!" – Oscar Hammerstein II

Lecture Videos

Rings | Subrings | The Dominance of Zero in a Ring | Matrix Rings | Polynomial Rings

10.1 Rings

Up until now, we have focused on algebraic sets with a single binary operation. Often in algebra, it is useful to consider a set with multiple operations that interact with each other nicely, for example vector spaces in linear algebra have two operations: vector addition and scalar multiplication in which scalar multiplication distributes over vector addition. Inner product spaces have a third operation, the inner product, that also behaves nicely with the two other operations of the vector space. Many other operations from linear algebra could be listed, like the cross product and outer product, which behave nicely with the other algebraic operations often in the form of some kind of distributive property or homogeneity property. We have also seen in set theory that binary operations of intersection, union, and set difference are very compatible with each other and other important set operations.

The study of algebraic rings is in this vein, that is, a ring is, loosely speaking, an algebraic set with two binary operations, called addition and multiplication, that are compatible (distributive) with each. The precise definition is below.

Definition 10.1.1. We say that $(R, +, \cdot)$ is a **ring** if the following seven axioms are satisfied by the binary operations:

(i) (**additive associativity**) For all $r, s, t \in R$, it holds that

$$r + (s + t) = (r + s) + t.$$

(ii) (**additive commutativity**) For all $r, s \in R$, it holds that

$$r + s = s + r.$$

(iii) (**additive identity**) There exists an element $0 \in R$ such that for all $r \in R$ we have

$$r + 0 = 0 + r = r.$$

(iv) (**additive inverses**) For all $r \in R$ there is an element $(-r) \in R$ such that

$$r + (-r) = (-r) + r = 0.$$

(v) (**multiplicative associativity**) For all $r, s, t \in R$, it holds that

$$r \cdot (s \cdot t) = (r \cdot s) \cdot t.$$

(vi) (**left distributivity**) For all $r, s, t \in R$, it holds that

$$r \cdot (s + t) = (r \cdot s) + (r \cdot t).$$

(vii) (**right distributivity**) For all $r, s, t \in R$, it holds that

$$(r + s) \cdot t = (r \cdot t) + (s \cdot t).$$

In particular, under addition, $(R, +)$ is an abelian group. Under multiplication, $(R, \cdot)$ is only required to be associate, which is known as a *semigroup*. When the binary operations $+$ and $\cdot$ are clear from context, we will say that R is a ring instead of $(R, +, \cdot)$.

Furthermore, we say R is a **ring with unity** if R is a ring which satisfies another additional axiom:

(viii) (**multiplicative identity**) There exists an element $1 \in R$ such that $1 \neq 0$ and for all $r \in R$ we have
$$r \cdot 1 = 1 \cdot r = r.$$

Additionally, we say R is a **commutative ring** if R is a ring which satisfies another additional axiom:

(ix) (**multiplicative commutativity**) For all $r, s \in R$, it holds that
$$r \cdot s = s \cdot r.$$

Example 10.1.2. The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all commutative rings with unity under their usual operations of addition and multiplication. The set $\mathbb{N}$ with its usual addition and multiplication is NOT a ring since $(\mathbb{N}, +)$ is not an abelian group.
Let $n \in \mathbb{Z}$. Then as usual, we have $n\mathbb{Z} = \{nx \mid x \in \mathbb{Z}\}$. Then $n\mathbb{Z}$ is a commutative ring but it does not have unity if $n \neq \pm 1$.

Example 10.1.3. Let $n \in \mathbb{Z}$. Then $\mathbb{Z}_n$, the set of congruence classes modulo n, is a ring under its usual addition and multiplication. In fact, $\mathbb{Z}_n$ is a commutative ring with unity for all n. Also, $\mathbb{Z}_n$ gives a family of finite rings.

Definition 10.1.4. Let R be a ring and let $S \subseteq R$. We say that S is a subring of R, denoted $S \leq R$, if S is an additive subgroup of R and is closed under multiplication. If R is a ring with unity, we say that S is a subring with unity if S is a subring of R and contains the unity of R.

Note that it is possible that S can be a subring of R and S can be a ring with unity without being a subring with unity, that is, S can be closed under the multiplication of R and can have a multiplicative identity which is NOT the multiplicative identity of R. The reader will explore this idea in the homework.

Example 10.1.5. The chain $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$ is a sequence of subrings. Likewise, $n\mathbb{Z} \leq \mathbb{Z}$ for each $n \in \mathbb{Z}$.

Proposition 10.1.6. *Let R be a ring with $r, s \in R$. Then*

(i) $r \cdot 0 = 0 \cdot r = 0$;

(ii) $r(-s) = (-r)s = -(rs)$;

(iii) $(-r)(-s) = rs$.

Proof. Note that $0r = (0+0)r = 0r + 0r$, by the distributive law. Then we add $-0r$ to both sides of this equation:

$$\begin{aligned} 0r + (-0r) &= (0r + 0r) + (-0r) = 0r + (0r + (-0r)) \\ 0 &= 0r + 0 = 0r \end{aligned}$$

A similar calculation gives $r0 = 0$. This proves (i).

Note that $rs + (-r)s = (r + (-r))s = 0s = 0$, by (i) and the distributive law. Thus, $(-r)b$ acts like the additive inverse of rs. Since $(R, +)$ is a group, additive inverses are unique. Therefore, $(-r)s = -rs$. Similarly, $r(-s) = -rs$, which proves (ii).

Finally, $(-r)(-s) = -r(-s) = -(-rs)$ by applying (ii) twice. In a group, the inverse of the inverse is the original element, that is, $-(-rs) = rs$, which proves (iii). □

Example 10.1.7. The sets of $n \times n$ matrices $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ are rings under the usual matrix addition and matrix multiplication. These are rings with unity since the identity matrix I_n acts as the multiplicative identity of matrix multiplication. On the other hand, these rings are noncommutative since matrix multiplication is noncommutative. When we say that a ring is noncommutative or commutative, we are always describing the multiplication of the ring because the addition is always commutative in a ring.

In fact, if R is any ring then we can form a new **matrix ring** $M_n(R)$ which is the set of $n \times n$ matrices with entries coming from R. Addition of matrices is defined analogously where terms are added component-wise, a binary operation which is well-defined since we can add elements in a ring R. Multiplication of matrices is also defined analogously where rows are multiplied by columns (using dot products), a binary operations which is well-defined since we can multiply and add elements in a ring R. The matrix ring $M_n(R)$ will be noncommutative if and only if $n > 1$, even if R is a commutative ring. Furthermore, if R is a ring with unity, then $M_n(R)$ will be a ring with unity, namely the identity matrix I_n.

Example 10.1.8. Let $\mathbb{R}[x]$ denote the set of all polynomials with real coefficients. This forms a ring where addition is defined as the usual polynomial addition ("combine like-terms") and multiplication is defined as the usual polynomial multiplication ("extended FOIL method"). Then this is a commutative ring with unity. In fact, if R is any ring then we can form a new **polynomial ring** $R[x]$ which is the set of all polynomials with coefficients coming from R. Since we can add and multiply elements from R, polynomial addition and multiplication are binary operations on R-polynomials, which makes $R[x]$ into a ring. The ring $R[x]$ is commutative if and only if R is commutative. Also, $R[x]$ has unity if and only if R has unity. In this case, the unity of $R[x]$ is the unity of R viewed as a constant polynomial.

[i] See §16.1 Rings in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–8, determine if the given set with its usual notion of addition and multiplication is a ring.

1. $7\mathbb{Z}$
2. $\mathbb{Z}_{18}$
3. $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$
4. $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\}$
5. $\mathbb{Z}[\sqrt{3}] = \{a + b\sqrt{3} \mid a, b \in \mathbb{Z}\}$
6. $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$
7. $R = \{a + b\sqrt[3]{3} \mid a, b \in \mathbb{Q}\}$
8. $\mathbb{Q}(\sqrt[3]{3}) = \{a + b\sqrt[3]{3} + c\sqrt[3]{9} \mid a, b, c \in \mathbb{Q}\}$

For Exercises 9–14, evaluate the given polynomial.

9. $(x^3 + 3x - 4) + (4x^2 - x + 9) \in \mathbb{Z}_{12}[x]$
10. $(x^3 + 3x - 4)(4x^2 - x + 9) \in \mathbb{Z}_{12}[x]$
11. $(x^3 + 3x - 4)^2 \in \mathbb{Z}_{12}[x]$
12. $(7x^3 + 3x^2 - x) + (6x^2 - 8x + 4) \in \mathbb{Z}_9[x]$
13. $(3x^2 + 2x - 4) + (4x^2 + 2) \in \mathbb{Z}_5[x]$
14. $(3x^2 + 2x - 4)(4x^2 + 2) \in \mathbb{Z}_5[x]$

For Exercises 15–22, prove the given statement about rings.
(Pick 4 from Exercises 15-22)

15. For a ring R with unity 1, $(-1)r = -r$ for all $r \in R$.
16. If R is a commutative ring with unity, then $R[x]$ is a commutative ring with unity.
17. Let R be a ring and $S \subseteq R$. Then $S \leq R$ if and only if
    (i) $S \neq \emptyset$,
    (ii) $rs \in S$, $\forall r, s \in S$,
    (iii) $r - s \in S$, $\forall r, s \in S$.

> **Definition 10.1.9.** A ring R is a **Boolean ring** if for every $r \in R$, $r^2 = r$.

18. Every Boolean ring is commutative.
19. Let R be a ring such that $r^3 = r$ for all $r \in R$. Then R must be commutative.

> **Definition 10.1.10.** Let R be a ring. Then the **center** of R, denoted $Z(R)$, is
>
> $$Z(R) := \{z \in R \mid zr = rz, \forall r \in R\}.$$

20. For any ring R, $Z(R) \leq R$ is a commutative subring. If R is a ring with unity, then $Z(R)$ is a subring with unity.
21. Let p be a prime number. Then
    $$\mathbb{Z}_{(p)} := \left\{ \frac{a}{b} \;\middle|\; a, b \in \mathbb{Z}, \gcd(b, p) = 1 \right\}$$
    is a ring.

> **Proposition 10.1.11.** *Let R and S be rings. Then $R \times S$ can be made into a ring, called the **direct product** of R and S, using component–wise addition and multiplication, that is,*
>
> $$(r, s) + (r', s') = (r + r', s + s'), \qquad (r, s)(r', s') = (rr', ss').$$

22. Proposition 10.1.11.

"Where there is unity there is always victory." – Publilius Syrus

Lecture Videos

Group Rings | Fields | Division Rings | Units of Rings

10.2 Fields

We can generalize the construction above to an arbitrary group. Let R be a ring and G be a multiplicative group. Then we can form the **group ring** $R[G] = \left\{\sum_{g\in G} r_g g \mid r_g \in R\right\}$ as the set of all finite linear combinations of elements of the group G with coefficients coming from R. Addition is defined by "combining like-terms," that is,

$$\sum_{g\in G} r_g g + \sum_{g\in G} s_g g = \sum_{g\in G} (r_g + s_g) g.$$

Multiplication is defined by the "extended FOIL method," that is, the group multiplication is extended to the set of all linear combinations so that the distributive laws hold:

$$\left(\sum_{g\in G} r_g g\right)\left(\sum_{g\in G} s_g g\right) = \sum_{g\in G}\left(\sum_{hk=g} r_h s_k\right) g.$$

This will form a ring with unity (the identity of the group G) and is commutative if and only if R is commutative and G is abelain.

Example 10.2.1. Consider the group ring $\mathbb{R}[\mathbb{Z}]$, where we view $\mathbb{Z} = \langle x\rangle = \{\ldots, x^{-2}, x^{-1}, 1, x, x^2, x^3, \ldots\}$ multiplicatively. Some elements of this group ring would be

$$3x^2 + 2x - 1 + x^{-1} + 3x^{-2} - 5x^{-3}, x^3 - x + 2, 3 - x^{-1} - x^{-2} \in \mathbb{R}[\mathbb{Z}].$$

Some examples of addition and multiplication would be

$$\begin{aligned}
(3x^2 + 2x - 1 + x^{-1} + 3x^{-2} - 5x^{-3}) + (x^3 - x + 2) &= x^3 + 3x^2 + x + 1 + x^{-1} + 3x^{-2} - 5x^{-3};\\
(3x^2 + 2x - 1 + x^{-1} + 3x^{-2} - 5x^{-3}) + (3 - x^{-1} - x^{-2}) &= 3x^2 + 2x + 2 + 2x^{-2} - 5x^{-3};\\
(x^3 - x + 2) + (3 - x^{-1} - x^{-2}) &= x^3 - x + 5 - x^{-1} - x^{-2};
\end{aligned}$$

$$\begin{aligned}
(3x^2 + 2x - 1 + x^{-1} + 3x^{-2} - 5x^{-3})(x^3 - x + 2) &= 3x^5 + 2x^4 - x^3 + x^2 + 3x - 5\\
&\quad - 3x^3 - 2x^2 + x - 1 - 3x^{-1} + 5x^{-2}\\
&\quad + 6x^2 + 4x - 2 + 2x^{-1} + 6x^{-2} - 10x^{-3}\\
&= 3x^5 + 2x^4 - 4x^3 + 5x^2 + 8x - 8 - x^{-1} + 11x^{-2} - 10x^{-3};
\end{aligned}$$

$$\begin{aligned}
(3x^2 + 2x - 1 + x^{-1} + 3x^{-2} - 5x^{-3})(3 - x^{-1} - x^{-2}) &= 9x^2 + 6x - 3 + 3x^{-1} + 9x^{-2} - 15x^{-3}\\
&\quad - 3x - 2 + x^{-1} - x^{-2} - 3x^{-3} + 5x^{-4}\\
&\quad - 3 - 2x^{-1} + x^{-2} - x^{-3} - 3x^{-4} + 5x^{-5}\\
&= 9x^2 + 3x - 8 + 2x^{-1} + 9x^{-2} - 19x^{-3} + 2x^{-4} + 5x^{-5};
\end{aligned}$$

$$\begin{aligned}
(x^3 - x + 2)(3 - x^{-1} - x^{-2}) &= 3x^3 - x^2 - x\\
&\quad - 3x + 1 + x^{-1}\\
&\quad + 6 - x^{-1} - 2x^{-2}\\
&= 3x^3 - x^2 - 4x + 7 - 2x^{-2}
\end{aligned}$$

It is important to realize that although we require additive inverses in a ring, we are not requiring that a ring have inverses for multiplication. But there do exist ring which have multiplicative inverses.

Definition 10.2.2. We say that R is a **field** if R is a commutative ring with unity which satisfies another additional axiom:

(x) (**multiplicative inverses**) For all $r \in R \setminus \{0\}$ there exists an element $r^{-1} \in R \setminus \{0\}$, it holds that

$$r^{-1} \cdot r = r \cdot r^{-1} = 1.$$

In every ring, addition, subtraction, and multiplication are always possible. Fields are the commutative rings where division is possible.

Example 10.2.3. The rings $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all examples of infinite fields. The ring of integers $\mathbb{Z}$ is a not a field since ± 1 are the only nonzero integers with multiplicative inverses in $\mathbb{Z}$.

Consider the finite ring $\mathbb{Z}_n$. We have seen previously that every element of $\mathbb{Z}_n^*$, the set of integers coprime to n, has a multiplicative inverse. In particular, $\mathbb{Z}_n^* = \mathbb{Z}_n \setminus \{0\}$ if and only if n is a prime number. Therefore, $\mathbb{Z}_n$ is a field if and only if n is prime. Finite fields, like $\mathbb{Z}_p$, are very important in algebra and its applications. For example, the finite field $\mathbb{Z}_2$ was very important to algebraic coding theory.

As we have seen in group theory, commutativity is not required for inverses.

Definition 10.2.4. A ring R is called a **division ring** (or **skew-field**) if R is a ring with unity that satisfies the above multiplicative inverse axiom above.

Essentially, a division ring is a not-necessarily-commutative field.

Example 10.2.5. Consider the quaternion group $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\} \leq \mathrm{GL}_2(\mathbb{C})$. Of course, $\mathrm{GL}_2(\mathbb{C}) \subseteq M_{2\times 2}(\mathbb{C})$. Viewing $M_{2\times 2}(\mathbb{C})$ as a 8-dimensional real vector space, let $\mathbb{H}^{\mathrm{i}} = \mathrm{Span}_{\mathbb{R}}(Q_8) = \mathrm{Span}_{\mathbb{R}}(1, i, j, k) = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}$, which is called the (ring of) quaternions. Since Q_8 forms a group, the set $\mathbb{H}$ forms a ring with unity. It is noncommutative, since multiplication in Q_8 is nonabelian. It is a division ring, which we show now.

In a natural way, we can view $\mathbb{H}$ as an extension of $\mathbb{C}$. As such, we extend the notion of complex conjugation to quaternion conjugation as the rule:

$$\overline{a + bi + cj + dk} = a - bi - cj - dk.$$

Then quaternions are added and subtracted using "like-terms." Multiplication comes from the "extended FOIL method," using the usual multiplication of quaternions. Finally, division is computed using conjugates:

$$(a + bi + cj + dk)^{-1} = \frac{1}{a + bi + cj + dk} = \frac{1}{a + bi + cj + dk}\left(\frac{a - bi - cj - dk}{a - bi - cj - dk}\right) = \frac{a - bi - cj - dk}{a^2 + b^2 + c^2 + d^2}$$

Therefore, $\mathbb{H}$ is a noncommutative division ring.

Of course, many rings like $\mathbb{Z}$, $\mathbb{R}[x]$, $\mathbb{R}[G]$ are not fields or skew-fields, that is, not all nonzero elements are invertible. Let R be a ring with unity and let $r \in R$. We say that r is a **unit** in R if there exists a multiplicative inverse $r^{-1} \in R$. Let R^* (or $U(R)$) denote the set of units in R. This forms a multiplicative group. A ring is a skew-field if and only if $R^* = R \setminus \{0\}$.

Example 10.2.6. The unit of $\mathbb{Z}$ are ± 1, that is, $\mathbb{Z}^* = \{1, -1\}$. For $\mathbb{Z}_n$, the units are exactly those integers coprime to n, an observation we have used all semester long.

In the ring $M_{n\times n}(R)$, the units are the general linear group, that is, $M_{n\times n}(R)^* = \mathrm{GL}_n(R)$.

In the ring $R[x]$, the units are just the units of R, that is, $R[x]^* = R^*$. In the group ring $R[G]$, the calculation of units is much more difficult. We can see that $R^*, G \leq R[G]^*$, called the *trivial units*. In many group rings there exists nontrivial units, for example in $\mathbb{Z}[S_3]$ we have the elements $\mu = 1 + (123) - (132) + (13) - (23)$ and $\mu^{-1} = 1 - (123) + (132) - (13) + (23)$. Note that

$$\begin{aligned}
\mu\mu^{-1} &= (1 + (123) - (132) + (13) - (23))(1 - (123) - (13) + (23) + (132)) \\
&= 1 - (123) + (132) - (13) + (23) \\
&\quad +1 + (123) - (132) + (12) - (23) \\
&\quad -1 + (132) - (12) + (13) + (23) \\
&\quad -1 + (123) - (12) + (13) - (23) \\
&\quad +1 - (123) - (132) + (12) - (13) \\
&= 1
\end{aligned}$$

Therefore μ is a unit in $\mathbb{Z}[S_3]$. The study of units in group rings is a very active research area in modern algebra.

[i] This is not to be confused with the group ring $\mathbb{R}[Q_8]$ which is an 8-dimensional real vector space. The quaternion ringa topic studied at length in representation theory.

[ii] See §16.1 Rings and §16.2 Integral Domains and Fields in Judson's Abstract Algebra: Theory and Applications for additional reading.

Exercises

(Go to Solutions)

For Exercises 1–6, list all units in the given ring.

1. $\mathbb{Z}_{10}$
2. $\mathbb{Z}_{12}$
3. $\mathbb{Z}_7$
4. $M_2(\mathbb{Z})$
5. $M_2(\mathbb{Z}_2)$
6. $\mathbb{Z}[x]$

For Exercises 7–10, prove the given statement about units.

7. Let $R = \left\{ \begin{pmatrix} r & s \\ 0 & 0 \end{pmatrix} \middle| \, r, s \in \mathbb{R} \right\} \leq M_2(\mathbb{R})$. Then R is a ring without unity but has a nonzero subring S which has unity.
8. There exists a ring R with unity 1_R and a subring $S \leq R$ which also has unity 1_S but $1_R \neq 1_S$.
9. If R is a ring with unity such that $0 = 1$, then $R = \{0\}$.
10.
11. For any ring R, $R[x]$ is NOT a field.

"Must is a hard nut to crack, but it has a sweet kernel." – Charles Spurgeon

Lecture Videos

Ring Homomorphisms

Examples of Ring Homomorphisms

Kernels of Ring Homomorphisms

10.3 Ring Homomorphisms

he analogous notion of homomorphism extends to rings.

Definition 10.3.1. Let R and S be rings. Then a map $\varphi : R \to S$ is a **(ring) homomorphism** if φ *preserves* the ring operations, that is, for all $r, s \in R$,

$$\varphi(r + s) = \varphi(r) + \varphi(s),$$

$$\varphi(rs) = \varphi(r)\varphi(s).$$

In other words, a homomorphism is a function with the homomorphic property. The definitions of monomorphism, epimorphism, isomorphism, automorphism, and endomorphism are extended to rings analogously.

We define the **kernel** of a ring homomorphism, denoted $\ker \varphi$, as the set of all elements of R which map to the zero element of S, that is,

$$\ker \varphi = \{r \in R \mid \varphi(r) = 0\}.$$

We note that as a homomorphism preserves the additive structure of R, in addition to the multiplicative structure, every ring homomorphism on $(R, +, \cdot)$ is necessarily a group homomorphism on $(R, +)$. As such, many of the following properties of ring homomorphisms are carried over directly from group theory. The ones which do not can be proven analogously and are left as an exercise to the reader.

Proposition 10.3.2. *Let $\varphi : R \to S$ be a ring homomorphism. Then*

(i) $\ker \varphi \leq R$, *that is,* $\ker \varphi$ *is a subring of* R,

(ii) $\operatorname{im} \varphi \leq S$, *that is,* $\operatorname{im} \varphi$ *is a subring of* S,

(iii) $\varphi(0_R) = 0_S$,

(iv) $\varphi(1_R) = 1_{\operatorname{im} \varphi}$[i],

(v) if $R' \leq R$, *then* $\varphi(R') \leq S$,

(vi) if $S' \leq S$, *then* $\varphi^{-1}(S) \leq R$,

(vii) if R *is commutative,* $\operatorname{im} \varphi$ *is commutative.*

Example 10.3.3. The map $\varphi : \mathbb{Z} \to \mathbb{Z}_n$ given by $m \mapsto m \pmod n$ is a ring homomorphism. More generally, if S is a ring with unity 1_S, then the map $\varphi : \mathbb{Z} \to S$ given by the rule $m \mapsto m1_S$ is a ring homomorphism since

$\varphi(m+n) = (m+n)1_S = m1_S + n1_S = \varphi(m) + \varphi(n)$ and $\varphi(mn) = (mn)1_S = (m1_S)(n1_S) = \varphi(m)\varphi(n)$.

Note that the kernel of φ will be $n\mathbb{Z}$ where $\operatorname{Char} S = n$. In particular, if $\operatorname{Char} S = 0$, then no finite sum of 1_S will ever equal 0. As such, the kernel of φ would be trivial, which implies that $\ker\varphi = \{0\} = 0\mathbb{Z}$. This justifies why we call such rings characteristic 0.

Example 10.3.4. Let X be a set and R a ring. Then recall that R^X is the set of all functions of the form $f : X \to R$. Then R^X can be made into a ring by defining addition and multiplication of functions via addition and multiplication of their images (the same method used in Calculus). Let $x \in X$. Then define the map

$$\varphi_x : R^X \to R : f \mapsto f(x),$$

called the **evaluation map**. The definitions of function addition and multiplication are defined exactly to guarantee that this map is a ring homomorphism.

Example 10.3.5. Let R be a ring (or even an abelian group would suffice!). Recall that an endomorphism on R is a homomorphism of the form $\varphi : R \to R$. Let $\operatorname{End}(R)$ be the set of all endomorphisms on R. This forms a ring with unity (even if R is not), called the **endomorphism ring** where addition is defined as function addition and multiplication is defined as function composition (the details are left to the reader to check that these satisfy the ring axioms). The zero of this ring is the zero map: $r \mapsto 0$ for all $r \in R$. The unity of this ring is the identity map. Typically this ring is noncommutative even if R is commutative. The group of units of the endomorphism ring $\operatorname{End}(R)$ is the automorphism group $\operatorname{Aut}(R)$.

We have seen when studying the Isomorphism Theorems that kernels of group homomorphisms are equivalent to normal subgroups and that normal subgroups are the only type of subgroup which can be modded out by to form a quotient group with well-defined multiplication. What is the ring analog? As $(R, +)$ is abelian, every subring I of R is necessarily a normal additive group. Thus, if R/I is a well-defined ring, then it is the multiplication which must be well-defined.

Let R/I denote the set of all additive cosets of the subring I in the ring R, for example, if $r \in R$, then $r + I \in R/I$. Let $r + I, s + I \in R/I$. Note that as a product of sets (the Frobenius product)

$$(r + I)(s + I) = rs + rI + Is + II \subseteq rs + rI + Is + I,$$

where the last equality holds since $I \leq R$ and is closed under multiplication. Because I is also closed under addition, we know that $I + I = I$. What we want is that $(r + I)(s + I) \subseteq rs + I$.[ii] This would require that $rI + Is + I \subseteq I$ for all $r, s \in I$. In particular, if $s = 0$, then $rI + I \subseteq I$, and if $r = 0$, then $Is + I \subseteq I$. These conditions imply that $rI \subseteq I$ and $Is \subseteq I$ for all $r, s \in I$. This is a stronger type of multiplicative closure, that is, I is closed not just under multiplication in I but if one factor is from I and the other is from R then I is still closed.

Definition 10.3.6. Let R be a ring and let I be a subring of R. We say that I is an **ideal**[iii] of R if $rI, Is \subseteq I$ for all $r, s \in I$. If I is an ideal, then we denote this as $I \trianglelefteq R$.[iv]

Ideals are exactly the subrings of a ring that guarantee that the set of cosets will have a well-defined addition and multiplication, thus forming a well-defined notion of quotient ring.

Theorem 10.3.7. *Let $\varphi : R \to S$ be a ring homomorphism. Then* $\ker\varphi \trianglelefteq R$.

Proof. As we already know that $\ker\varphi \leq R$, it suffices to show that $\ker\varphi$ satisifies the ideal closure

property. Let $r \in R$ and $x \in \ker \varphi$. Then

$$\varphi(rx) = \varphi(r)\varphi(x) = \varphi(r)0 = 0.$$

Thus, $rx \in \ker \varphi$. In particular, $r \ker \varphi \subseteq \ker \varphi$. Similarly, we see that $\ker \varphi s \subseteq \ker \varphi$ for all $s \in R$. Therefore, $\ker \varphi$ is an ideal. □

If $I \trianglelefteq R$, then the natural quotient map $\eta : R \to R/I$ is a ring homomorphism whose kernel is clearly I. This shows that the notion of kernels and ideals are equivalent in ring theory.

[i]If R is a ring with unity, then $\operatorname{im}\varphi$ is a ring with unity but it is not necessarily a subring with unity in S, as we have seen this potential problem before. In other words, $\varphi(1)$ acts like unity for $\varphi(R)$ although it might not be the unity of S. Ways to guarantee that $\varphi(1_R) = 1_S$: first, if φ is a epimorphism, that is, if φ is onto then $\operatorname{im}\varphi = S$ and $1_{\operatorname{im}\varphi} = 1_S$; second, as $\varphi(1)$ is necessarily an idempotent element of S ($\varphi(1)^2 = \varphi(1)$), it suffices to forbid S from having idempotents other than 0 and 1, a condition true in any domain; third, we could require that $\varphi(1_R) = 1_S$ as part of the definition of a homomorphism: if $\varphi : R \to S$ be a homomorphism between rings with unity, we say that φ is **unital** if $\varphi(1_R) = 1_S$. Many authors define all ring homomorphisms to be unital and never use the adjective as it is understood to be always that case. These are the same authors that require all rings to have unity. Typically, these authors will use the term *non-unital* to describe what we are general rings and homomorphism.

[ii]In the case of group cosets, one could define coset multiplication in terms of equivalence relations or in terms of set operations, which we will call the *modular* and *Frobenius* versions, respectively, for clarity's sake. We previously defined coset multiplication in the Frobenius-sense but the textbook defines it using the modular approach.

Suppose that G is a group and let $C, D \subseteq G$. Recall that the (**Frobenius**) **product** of subsets is

$$CD := \{cd \mid c \in C, d \in D\} \subseteq G.$$

Hence, we can create a product on the subsets of a group by defining multiplication between sets element-wise. Recall that a coset is defined using this Frobenius product, ie,

$$gN = \{g\}N = \{gn \mid n \in N\}.$$

Thus, it seems natural to define the product of cosets to be their Frobenius product, as subsets of G. If N is a normal subgroup of G, then

$$(gN)(hN) = g(Nh)N = g(hN)N = (gh)(NN) = (gh)N,$$

where each equality holds as a (Frobenius) set equality. With this perspective there is no need to "define" a product of cosets in any other way because there already exists a product of subsets that restricts exactly to the product of cosets we need. Conversely, the modulo product of cosets, $(gN)\sharp(hN)$, defines this as the unique coset containing gh, namely, $(gN)\sharp(hN) := (gh)N$.

While no one introduces the alternative notation of $\sharp$ for coset product. Instead, the two meanings of coset multiplication is overloaded, but this is no concern for groups, because these two different definitions of coset multiplication produce the exact same binary operation on the set of cosets (not just isomorphic, they are literally the same operation on the set of cosets). Therefore, for group cosets, $(gN)(hN) = (gN)\sharp(hN) = (gh)N$. Hence, we do not have to define a new operation when introducing coset multiplication on groups because the product is already well-defined and is associative. One only needs to show that the product of two cosets is a coset (which we have already shown). That is, the Frobenius product makes the power set of a group into a semigroup and the restriction to cosets of a subgroup forms a subsemigroup of the power set if and only if this subgroup is normal. This subsemigroup of the power set is in fact a group, which is identical to the quotient group created using the modular approach.

It is natural to suppose this same pattern occurs for ring cosets. Let R be a ring and I be an ideal. It is true that

$$(a + I) + (b + I) = \{(a + i) + (b + j) \mid i, j \in I\} = \{(a + b) + I \mid i \in I\} = (a + b) + I,$$

where this is the sum of two subsets of R in the Frobenius-sense. This follows from above since $(R, +)$ is an abelian group and I is an additive subgroup of $(R, +)$. Unfortunately, the assumption that I is an ideal is insufficient to prove that $(a+I)(b+I) = ab+I$ in the Frobenius-sense. Even in a commutative ring, this is false. Let $R = \mathbb{Z}_2[x]/(x^3 - x^2)$ and let $I = (x) = \{0, x, x^2, x^2 + x\}$. Note that

$$I^2 = \{0, x, x^2, x^2 + x\}\{0, x, x^2, x^2 + x\} = \{0, x^2\} \neq I.$$

This defect is a consequence of $(R, *)$ only being a semigroup (well, commuative monoid in this case). Without cancellation, we cannot guarantee the Frobenius product of two cosets is itself a coset.

It is easy to show from the properties of an ideal that

$$(a + I)(b + I) \subseteq ab + I$$

always, but, as we saw above, equally can fail. So while the Frobenius product of subsets of a ring still applies in this context, the Frobenius product of two cosets is not necessarily a coset of the ideal. But, and this is an important but, because the system of cosets of a fixed ideal of a ring forms a partition of the ring, if $(a + I)(b + I) \subseteq ab + I$, then $ab + I$ is the only coset that contains $(a + I)(b + I)$. Therefore, we can still define the modular product of cosets and the above set containment is sufficient to prove this modular coset multiplication is well-defined.

[iii]The object defined above is often called a **two-sided ideal**. This is becuase in for commutative rings the second requirement $Is \subseteq I$ is implied by the first but it is not so for general noncommutative rings. That is, there exists subrings I of R such that $rI \subseteq I$ for all $r \in R$ but not necessarily $Is \subseteq R$ for all $s \in R$. Also, there exists subrings I of R such that $Is \subseteq I$ for all $s \in R$ but not necessarily $rI \subseteq R$ for all $r \in R$. These are called **left-** and **right-ideals**, respectively. These one-sided ideals are extremely important in more advanced ring theory, such as module theory or representation theory, but we will not need to consider them further in this course.

[iii]Unfortunately, this notation is not always consistent. Many ring theorists simply denote I is an ideal of R as $I \subseteq R$. As

Exercises

(Go to Solutions)

1. What is the same and what is different about homomorphisms for groups and rings?

rarely anybody is discussing a strict subset of R without some other algebraic notion, this rarely leads to confusion, but the author prefers the less ambiguous notion mentioned above, especially as it emphasizes the parallel with normal subgroups in group theory.

[iv] See §16.3 Ring Homomorphisms and Ideals in Judson's Abstract Algebra: Theory and Applications for additional reading.

10.4 Supplemenatry Exercises

(Go to Solutions)

Definition 10.4.1. Let R be a commutative ring. An element $r \in R$ is **nilpotent** if $r^n = 0$ for some positive integer n.

1. Show that the set of nilpotent elements of a commutative ring R forms a subring.
2. List all polynomials in $\mathbb{Z}_2[x]$ of degree less than or equal to 3.
3. Find a unit $f(x) \in \mathbb{Z}_4[x]$ such that $\deg(f) > 1$.
4. Let R be ring with unity. For $u \in R^*$, define $\iota_u : R \to R : r \mapsto uru^{-1}$. Let $\text{Inn}(R) = \{\iota_u \mid u \in R^*\}$. Show that $\text{Inn}(R)$ is a group.

Appendix A

Complex Numbers

Lecture Videos


Addition/Subtraction of Complex Numbers


Multiplication of Complex Numbers


Division of Complex Numbers


The Complex Plane


Polar Form of Complex Numbers


Computing the Polar Form of a Complex Number


Euler's Identity

A.1 Algebra of Complex Numbers

There exists no real number r such that $r^2 = -1$, that is, $\sqrt{-1}$ is not a real number. This doesn't mean that $\sqrt{-1}$ doesn't exists or shouldn't exist. It just means that we shouldn't call it a "real" number. It is not my desire to now go into a long philosophical or historical discussion on what a "number" is. So, needless to say $\sqrt{-1}$ is an *imaginary number*.

Definition A.1.1. Let $i = \sqrt{-1}$. Then we say z is **complex number** if it is of the form

$$z = a + bi$$

where $i = \sqrt{-1}$ and a and b are real numbers. The real number a is called the **real part** of z; the real number b is called the **imaginary part** of z.

When adding or subtracting complex numbers, follow the simple rule of combining like-terms, that is, combine the real parts together and combine the imaginary parts together.

Example A.1.2.

(a) $(3+5i)+(-2+3i) = (3-2)+(5+3)i = \boxed{1+8i}$

(b) $(6+4i)-(3+6i) = (6-3)+(4-6)i = \boxed{3-2i}$

Multiplying complex numbers boils down to the FOIL method. When doing arithmetic, such as multipli-

cation, on complex numbers, all complex numbers should be simplified to the **standard form** $a+bi$. With this said, it is useful to note that $i^2 = -1$.

Example A.1.3.

(a)
$$\begin{aligned}(5+3i)(2+7i) &= 10+35i+6i+21i^2\\ &= (10-21)+(35+6)i\\ &= \boxed{-11+41i}\end{aligned}$$

(b)
$$\begin{aligned}(2+i)^3 &= (2+i)(2+i)(2+i) = (4+2i+2i+i^2)(2+i)\\ &= (4+4i-1)(2+i) = (3+4i)(2+i)\\ &= 6+3i+8i-4 = \boxed{2+11i}\end{aligned}$$

Before introducing complex division, we introduce the conjugate of a complex number.

Definition A.1.4. Let $z = a+bi$ be a complex number. Then the **complex conjugate** of z, denoted by $\overline{z}$, is $\overline{z} = a - bi$.

Example A.1.5.

(a) $\overline{2+3i} = \boxed{2-3i}$

(b) $\overline{-6-2i} = \boxed{-6+2i}$

(c) $\overline{(2-3i)+(5+2i)} = \overline{7-i} = \boxed{7+i}$

(d) $(2+3i)+\overline{(2+3i)} = (2+3i)+(2-3i) = \boxed{4}$

(e)
$$\begin{aligned}(3+4i)\overline{(3+4i)} &= (3+4i)(3-4i)\\ &= 9-12i+12i-(-16)\\ &= 9+16 = \boxed{25}.\end{aligned}$$

Theorem A.1.6 (Properties of Complex Conjugates)**.** *Let $z, w \in \mathbb{C}$.*

(a) $\overline{z+w} = \overline{z}+\overline{w}$

(b) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$

(c) $\overline{\overline{z}} = z$

(d) $\overline{z} = z$ *if and only if* z *is a real number*

(e) $z+\overline{z}$ *is equal to twice the real part of* z.

Theorem A.1.7. *Let $z = a+bi$. Then*

$$z\overline{z} = a^2+b^2.$$

In particular, $z\overline{z}$ is a nonnegative real number.

Proof.
$$\begin{aligned}(a+bi)\overline{(a+bi)} &= (a+bi)(a-bi)\\ &= a^2-abi+abi-b^2i^2 = a^2-b^2i^2\\ &= a^2+b^2 \quad \square\end{aligned}$$

Complex division is accomplishing by multiplying the quotient $\dfrac{a+bi}{c+di}$ by $\dfrac{c-di}{c-di}$, where $\overline{c+di} = c-di$ is called the **complex conjugate** of $c+di$.

Example A.1.8. Write $\dfrac{1}{3+4i}$ in the form $a+bi$, that is, compute the reciprocal of $3+4i$.

$$\frac{1}{3+4i} = \frac{1}{3+4i}\left(\frac{\overline{3+4i}}{\overline{3+4i}}\right) = \frac{3-4i}{3^2+4^2} = \frac{3-4i}{25} = \boxed{\frac{3}{25}-\frac{4}{25}i}$$

This method of multiplying by the conjugate shows that you can divide by any nonzero complex number.

Example A.1.9.

(a)
$$\begin{aligned}\frac{1+4i}{5-12i} &= \frac{1+4i}{5-12i}\left(\frac{5+12i}{5+12i}\right) = \frac{(1+4i)(5+12i)}{25+144}\\ &= \frac{5+12i+20i-48}{169} = \boxed{-\frac{43}{169}+\frac{32}{169}i}\end{aligned}$$

(b)
$$\begin{aligned}\frac{2-3i}{4-3i} &= \frac{2-3i}{4-3i}\left(\frac{4+3i}{4+3i}\right) = \frac{(2-3i)(4+3i)}{16+9}\\ &= \frac{8+6i-12i+9}{25} = \boxed{\frac{17}{25}-\frac{6}{25}i}\end{aligned}$$

(c)
$$\begin{aligned}\frac{2-3i}{5+2i} &= \frac{2-3i}{5+2i}\left(\frac{5-2i}{5-2i}\right) = \frac{(2-3i)(5-2i)}{25+4}\\ &= \frac{10-4i-15i-6}{29} = \boxed{\frac{4}{29}-\frac{19}{29}i}\end{aligned}$$

Definition A.1.10. If N is a positive real number, we define the **principal square root** of $-N$, denoted $\sqrt{-N}$, as
$$\sqrt{-N} = i\sqrt{N} = \sqrt{N}i.$$

Note that $(i\sqrt{N})^2 = i^2\sqrt{N}^2 = -N$. So, $i\sqrt{N}$ is a square root of $-N$. The other square root is $-i\sqrt{N}$.

Example A.1.11.

(a) $\sqrt{-1} = \sqrt{1}i = \boxed{i}$

(b) $\sqrt{-4} = \sqrt{4}i = \boxed{2i}$

(c) $\sqrt{-8} = i\sqrt{8} = \boxed{2\sqrt{2}i}$

Example A.1.12. Solve the equation $x^2 = -9$.

By the square root method, the solutions should be $x = \pm\sqrt{-9}$. Considering the above, we have $x = \pm\sqrt{9}i = \pm 3i = 3i, -3i$. Note that $(3i)^2 = 9i^2 = -9$ and $(-3i)^2 = 9i^2 = -9$. Thus, $\boxed{x = 3i, -3i}$ are the solutions of this quadratic equation.

When the discriminant of a quadratic equation is negative, the equation will have two distinct, non-real, complex solutions.

Example A.1.13. Solve $x^2 - 4x + 8 = 0$.

A quick check will show that no magic pair exists for 8 and -4. So, we rely on the Quadratic Formula. Thus,

$$\begin{aligned} x &= \frac{-(-4) \pm \sqrt{(-4)^2 - 4(1)(8)}}{2(1)} = \frac{4 \pm \sqrt{16 - 32}}{2} \\ &= \frac{4 \pm \sqrt{-16}}{2} = \frac{4 \pm \sqrt{16}i}{2} \\ &= \frac{4 \pm 4i}{2} = 2 \pm 2i = \boxed{2 + 2i, 2 - 2i} \end{aligned}$$

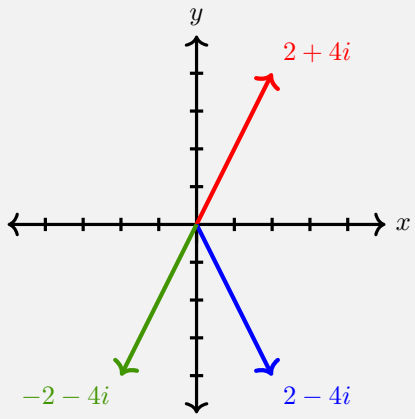
Note that $(2 + 2i)^2 - 4(2 + 2i) + 8 = (4 + 2i + 2i - 4) + (-8 - -8i) + 8 = 0$. So, $2 + 2i$ is a solution. Similarly, $(2 - 2i)^2 - 4(2 - 2i) + 8 = 0$.

A.2 Polar Form of Complex Numbers

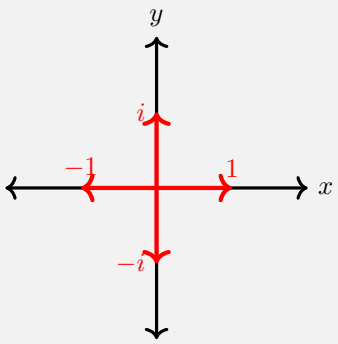
Definition A.2.1. The graph of a complex number $z = x + yi$ is the vector $\boldsymbol{v} = \begin{bmatrix} x \\ y \end{bmatrix}$.

When graphing complex numbers, the x-axis becomes the **real axis** and the y-axis becomes the **imaginary axis**. The resulting plane is called the **complex plane**.

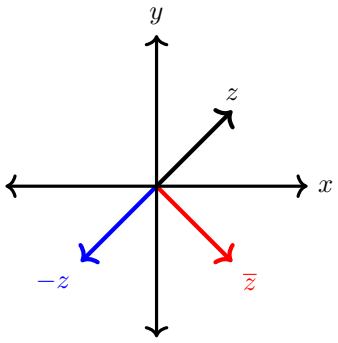
Example A.2.2. Graph each complex number: $2 + 4i$, $-2 - 4i$, and $2 - 4i$.



Example A.2.3. Graph each complex number: 1, -1, i, and $-i$.



If $z = x + yi$ is a complex number, then its inverse $-z = -x - yi$ is graphed by the vector pointing in the opposite direction and same length. Its conjugate $\overline{z} = x - yi$ is graphed by the vector formed by reflecting z across the real axis.



Definition A.2.4. The **absolute value** or **modulus** of the complex number $z = x + yi$ is the distance from the origin to the point (x, y) in the complex plane. In other words, it is the magnitude of the vector $\boldsymbol{z} = \langle x, y \rangle$. If this distance is denoted by r, then

$$r = |z| = |x + yi| = \sqrt{x^2 + y^2}.$$

Example A.2.5. Find the modulus of each complex number: $5i$, 7, and $3 + 4i$.

$$\begin{aligned} |5i| &= \sqrt{0^2 + 5^2} = \sqrt{25} = \boxed{5} \\ |7| &= \sqrt{7^2 + 0^2} = \sqrt{49} = \boxed{7} \\ |3 + 4i| &= \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = \boxed{5} \end{aligned}$$

Definition A.2.6. The **argument** of the complex number $z = x + yi$, denoted $\arg(z)$, is the smallest positive angle θ from the positive real axis to the graph of z. In other words, it is the direction of the vector $\boldsymbol{z} = \begin{bmatrix} x \\ y \end{bmatrix}$.

Note that $\arg(z) = \tan^{-1}\left(\frac{y}{x}\right) = \sin^{-1}\left(\frac{y}{|z|}\right) = \cos^{-1}\left(\frac{x}{|z|}\right)$.

Using basic trigonometry, we know that the components of $\boldsymbol{z} = \begin{bmatrix} x \\ y \end{bmatrix}$ are $x = r\cos\theta$ and $y = r\sin\theta$ where r and θ are the magnitude and direction of $\boldsymbol{z}$, respectively. Therefore, for the complex number $z = x + yi$, if $r = |z|$ and $\theta = \arg(z)$, then

$$\begin{aligned} z &= x + yi = (r\cos\theta) + (r\sin\theta)i \\ &= r\cos\theta + ri\sin\theta = r(\cos\theta + i\sin\theta)^{\text{i}} \end{aligned}$$

Definition A.2.7. Let $z = x + yi$, let $r = |z|$, and $\theta = \arg(z)$. Then

$$z = r(\cos\theta + i\sin\theta)$$

is the **polar form** of the complex number.

Theorem A.2.8 (Euler's Formula)**.** *Let $z = x + yi$, let $r = |z|$, and $\theta = \arg(z)$. Then*

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta}.$$

In particular,

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Proof. It suffices to prove the second equation. To do so, remember the Taylor series of e^x, $\sin x$, and $\cos x$, which can be naturally extended to any complex number:

$$\begin{aligned} e^z &= \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots \\ \sin(z) &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots \\ \cos(z) &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots \end{aligned}$$

The power series for e^z then allows us to compute $e^{i\theta}$.

$$\begin{aligned} e^{i\theta} &= \sum_{n=0}^{\infty} \frac{i\theta^n}{n!} = 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \frac{(i\theta)^6}{6!} + \frac{(i\theta)^7}{7!} + \dots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} - \frac{\theta^6}{6!} - \frac{i\theta^7}{7!} + \dots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots\right) + \left(i\theta - \frac{i\theta^3}{3!} + \frac{i\theta^5}{5!} - \frac{i\theta^7}{7!} + \dots\right) \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \dots\right) \\ &= \left(\sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}\right) + i\left(\sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}\right) = \cos(\theta) + i\sin(\theta) \end{aligned}$$

□

When $\theta = \pi$, we get

$$e^{\pi i} = \cos\pi + i\sin\pi = -1$$

Example A.2.9. Write $z = -1 + i$ in polar form.

Since $x = -1$ and $y = 1$, z will be in $Q2$. Also, we have $r = |z| = \sqrt{(-1)^2 + 1^2} = \sqrt{2}$. Likewise, $\theta = \arg(z) = \tan^{-1}\left(\frac{1}{-1}\right) = \tan^{-1}(-1) = \frac{3\pi}{4} (= 135^\circ)$. Therefore,

$$z = \boxed{\sqrt{2}\left(\cos\left(\frac{3\pi}{4}\right) + i\sin\left(\frac{3\pi}{4}\right)\right) = \sqrt{2}e^{3\pi i/4}}.$$

Example A.2.10. Write $z = 2e^{\pi i/3}$ in standard form.

$$z = 2e^{\pi i/3} = 2\left(\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)\right) = 2\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = \boxed{1 + i\sqrt{3}}.$$

[i] Many trigonometry textbooks use the notation $\operatorname{cis}\theta = \cos\theta + i\sin\theta$, which is not commonly used in advanced mathematics. For this reason, it would be better for you to not use cis. No cissing in class! Instead, use $e^{i\theta}$ to abbreviate $\cos\theta + i\sin\theta$.

Appendix B

MAGMA

In abstract algebra, like other mathematical disciplines, it can be very useful to have a "calculator" to assist with computations. Two computational algebra softwares appropriate for calculations in this book are MAGMA and GAP. Because MAGMA has a free, cloud based compiler, found at : http://magma.maths.usyd.edu.au/calc/, we will utilize MAGMA code from time-to-time in this book. A detailed explanation for MAGMA can be found in its handbook: http://magma.maths.usyd.edu.au/magma/handbook/, but we can provide a few tips here.

In MAGMA, all commands must end with a semicolon `;` (line breaks are ignored completely. The special character `//` starts a comment in MAGMA. Any character in a line after `//` is ignored by the compiler and is used simply for notes for the programmer. Variables are declared with `:=`, for example, `ID := 1234;` MAGMA does not require you to cast your variable type. It is *dynamic* in that the compilers learns what quantity based upon context. If $\circ$ is an operation, then `var` $\circ$`= num` replaces the current values of the variable `var` with `var` $\circ$ `num`. For example, if `var := 3;`, then `var += 2;` changes the value of `var` from `3` to `5`.

An array is started and ended with square brackets, `[` and `]`, and individual elements are separated by commas, or example, `[3,1,4,1,5,9]`. Arrays are lists, meaning they are ordered and repetition is allowed. The array $[1, 2, \ldots, n]$ can quickly be produced as `[1..n]`. The special character # gives the number of elements in an array. For example, If `array` is an array, then `#array` returns the number of elements in `array`. Referencing the ith element of `array` is done as `array[i]`. If the ith element of `array` is itself an array, then the jth element of `array[i]` is received as `array[i,j]`. If `array1` and `array2` are two arrays, then `array1 cat array2` is the concatenation of the two arrays, putting `array1` first, followed immediately by `array2`. For example, `[1,2,3] cat [4,5];` return `[1,2,3,4,5]`. For an associative operation $\circ$, e.g. $+$, $*$, `cat`, then `&+` $\circ$ computes the `#array`-fold $\circ$-product `array`. For example, `&+array` computes the sum of `array`.

Example. One can print to screen simply by stating the variable or quantity you want. For example,

```
p := 37316150861565913091;
q := 62369111002657007863; // This is a comment.
n := p*q; n;
```

returns 2327375155276899373645006956575561634533, the product of two large primes.

Example. To compute $\gcd(a, b)$ for $a, b \in \mathbb{Z}$ in MAGMA, type `Gcd(a,b);`. This will execute an algorithm similar to the Euclidean algorithm given in Example 1.6.4. For example, `Gcd(6,561);` returns `3`.

Example. To compute $a^{-1} \pmod{n}$ for $a, b \in \mathbb{Z}$ and $\gcd(a, n)$ in MAGMA, type `InverseMod(a,n);`. This will execute an algorithm similar to the Euclidean algorithm given in Example 1.6.4. For example, `InverseMod(629,3432);` returns `1997`.

Example. To compute $a^m \pmod{n}$ for $a, m, n \in \mathbb{Z}$ in MAGMA, type `Modexp(a,m,n);`. This will execute an algorithm similar to the Repeated Squares algorithm given in Example 3.3.6. For example, `Modexp(2,560,561);` returns `1`.

Alternatively, one could use `(a^m) mod n;`, but, unfortunately, this command will compute a^m first in $\mathbb{Z}$ then reduce modulo n which can lead to runtime errors if a^m is too large. The command `Modexp` instead runs the repeated squares algorithm which is much faster and not prone to error.

Example. To factor an integer $n \in \mathbb{Z}$ in MAGMA, type `Factorization(n);`. This will produce a factorization of n as a list of pairs `<p,k>`, where p is a prime factor of n and k is the largest power of p which divides n. Hence, $p^k \mid n$ but $p^{k+1} \nmid n$. For example, `Factorization(67989839287140);` returns `[ <2, 2>, <3, 4>, <5, 1>, <19, 1>, <109, 1>, <2287, 1>, <8861, 1> ]`, since

$$67989839287140 = 2^2 \cdot 3^4 \cdot 5 \cdot 19 \cdot 109 \cdot 2287 \cdot 8861.$$

To access specific pairs within the factorization, say `fact := Factorization(67989839287140);`, call an entry of the array by concatenating `[i]` at the end of the array. For example, `fact[2];` return `<3,4>`. If you want a specific prime or exponent in this factorization, use a double index `[i,j]`. For example, `fact[2,1];` return `3`. Note that we can compute the product of any n by `&*[fact[i,1]^fact[i,2] : i in [1..#fact]];`, where `fact := Factorization(n);`.

Like in Section 6.2, imagine that Eve find Bob's RSA public key (n, e). In order to decrypt messages sent to Bob from Alice, Eve needs to compute Bob's private key d, which is the solution to the congruence $de \equiv 1 \pmod{\phi(n)}$. She knows that $n = pq$ is a semiprime, that is, p and q are primes. As n is not very large, she is able to factor n, revealing p and q. With this factorization, Eve can easily compute $\phi(n) = \phi(p)\phi(q) = (p-1)(q-1)$. The following code finds Bob's private key:

```
n := 3551; e := 629;

fact := Factorization(n);
p:=fact[1,1]; q:=fact[2,1];

phi := (p-1)*(q-1);
d := InverseMod(e,phi); d;
```

which returns `1997`. Hence, the protection RSA comes down to how efficient `Factorization` runs. The general strategy to avoid factorization of n is to use BIG primes p and q, so that even the best factoring algorithms cannot find it in a short amount of time.

Example. Continuing the previous example, how does one find BIG primes? Given any integer n, MAGMA can find the next prime after n with the command `NextPrime(n);` (there is a similar function `PreviousPrime`.[i] For example, the command `NextPrime(67989839287140);` returns `67989839287147` (if every you are not sure if a number n is prime or not, run `IsPrime(n);`, which returns `True` if n is prime and `False` otherwise).

Another useful MAGMA function is `Random(a,b);` which computes a random integer between $a, b \in \mathbb{Z}$. For example, `Random(629,1997);` may return `1626` or `1373` or `1200`, since each run will produce a random value (if you are using random numbers, make sure to record

them before your next compile). Combining this with `NextPrime`, we can create a random prime generator, where `NextPrime(Random(a,b));` returns a random prime between a and b. For example, `NextPrime(Random(10^10,10^20));` may return `37316150861565913091` or `62369111002657007863`, which are two primes between 10 and 20 digits long. Such a strategy could be useful in generating keys for RSA.

[i]Because of primality tests, some of which are hinted toward in the Exercises of Section 5.2, we can find primes easy enough, but factoring is still hard.

Appendix C

Solutions to Select Exercises

Chapter 1 : Abstract Prealgebra

1.1 Sets

1. $\{2\}$
2. $\{5\}$

1.2 Functions

1. $\{(a,1),(a,2),(a,3),(b,1),(b,2),(b,3),(c,1),(c,2),(c,3)\}$
4. $\emptyset$
5. Not function
6. Function
7. Not function
8. Function
9. One-to-one, not onto
10. Not one-to-one, not onto
11. Not one-to-one, not onto
12. Not one-to-one, onto

1.3 Inverse Functions and Permutations

7. $\operatorname{dom} f = \operatorname{im} f = (-\infty, 1) \cup (1, \infty)$, $f^{-1}(x) = \dfrac{x+1}{x-1}$, $(f \circ f^{-1})(x) = (f^{-1} \circ f)(x) = x$

1.4 Equivalence Classes

2. Not symmetric
3. Not transitive
4. Not reflexive
5. Equivalence relation
6. Not symmetric
7. Equivalence relation
8. Equivalence relation

1.5 Induction and the Well–Ordering Principle

10. *For the inductive step where $|X| = n+1$, let $x \in X$ be some distinguished element. If $A \subseteq X$, then $A \smallsetminus \{x\} \subseteq X \smallsetminus \{x\}$. How many options are there for $A \smallsetminus \{x\}$ then? Either $x \in A$ or $x \notin A$, which gives two options independent of the previous number.*

1.6 Divisibility of Integers

1. $1 = 14(14) - 5(39)$
2. $3 = -17(165) + 12(234)$
3. $1 = 3709(1739) - 650(9923)$

4. $1 = -105(471) + 88(562)$ 5. $1 = -1050(19945) + 881(23771)$ 6. $1 = 1463(-4357) + 1698(3754)$

10. *Use the Well-Ordering Principle combined with the Division Algorithm.*

14. *$ar + bs = 1$ implies that $arc + bsc = c$.*

1.7 Supplementary Exercises

7. *Use the Fundamental Theorem of Arithmetic.*

10. *Every prime has the form $2, 3, 6n + 1$, or $6n + 5$.*

12. *Prove by contradiction. If not, $\sqrt{2} = \frac{p}{q}$ for some $p, q \in \mathbb{Z}$ and $\gcd(p, q) = 1$. Show that $p^2 = 2q^2$. Apply Euclid's Lemma.*

Chapter 2 : Group Theory

2.1 Groups

1. $x \equiv 3$ 2. $x \equiv 7$ 3. $x \equiv 18$ 4. $x \equiv 2$ 5. $x \equiv 5$ 6. $\emptyset$

12. *Show that the operations do not depend on the choice of the representative from the equivalence classes modulo n.*

2.2 Cayley Tables

1. Not group 2. Group 3. Group 4. Not Group

9. *There is a nonabelian group containing six elements.*

10. *The are five different groups of order 8.*

2.3 Properties of Groups

6. *Since $abab = (ab)^2 = e = a^2b^2 = aabb$, use Cancellation.*

8. $a^4b = a^3(ab)$

2.4 Subgroups

3. *The identity of H is $1 = 1 + 0\sqrt{2}$. Show that $(a + b\sqrt{2})(c + d\sqrt{2}) = (ad + bc) + (ad + bc)\sqrt{2}$ and $(a + b\sqrt{2})^{-1} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}$.*

8. *Look at S_3.*

2.5 Hasse Diagrams

5. *S_3 has five subgroups: $H_1 = \{id\}, H_2 = \{id, \rho_1, \rho_2\}, H_3 = \{id, \sigma_1\}, H_4 = \{id, \sigma_2\}, S_3$.*

2.6 Supplementary Exercises

2. *Pick any two matrices and you probably have what you need.*

4. *The group will have to be nonabelian and $n \geq 2$. So maybe try D_4.*

5. $(ghg^{-1})^n = (ghg^{-1})(ghg^{-1}) \cdots (ghg^{-1}) = gh(g^{-1}g)h(g^{-1}g) \cdots (g^{-1}g)hg^{-1}$

6. *Prove that the number of elements whose order is not 2 must be odd*

Chapter 3 : Cyclic Groups

3.1 Cyclic Groups

1. 12 2. ∞ 3. ∞ 4. 4 5. 10 6. 157

3.2 Roots of Unity

1. $\sqrt(3)+i$
2. $\frac{5\sqrt{2}}{2}+\frac{5\sqrt{2}}{2}i$
3. -3
4. $\frac{\sqrt{2}}{4}-\frac{\sqrt{2}}{4}i$
5. $\sqrt{2}e^{7\pi i/4}$
6. $5e^{\pi i}$
7. $2\sqrt{2}e^{\pi i/4}$
8. $2e^{\pi i/6}$
9. $3e^{3\pi i/2}$
10. $4e^{\pi i/3}$
11. 15
12. $-3+3i$
13. $8-i$
14. $43-18i$
15. 82
16. i
17. 2
18. ∞
19. ∞
20. 4
21. 10
22. 157
23. $\frac{1}{2}-\frac{i}{2}$
24. $8i$
25. $-16\sqrt{3}+16i$
26. -1
27. $-\frac{1}{4}$
28. -4096
29. $\frac{1}{256}+\frac{i}{256}$

3.3 Orders of Group Elements

1. $7\mathbb{Z}=\{\ldots,-7,0,7,14,\ldots\}$
2. $\langle 15\rangle=\{0,3,6,9,12,15,18,21\}=3\mathbb{Z}_{24}$
3. $\langle 8\rangle=\{0,4,8\}=4\mathbb{Z}_{12}$
4. $\langle 8\rangle=\{0,4,8,12,16,20,24,28,32,36,40,44,48,52,56\}=4\mathbb{Z}_{60}$
5. $\langle 8\rangle=\{0,1,2,3,4,5,6,7,8,9,10,11,12\}=\mathbb{Z}_{13}$
6. $\langle 8\rangle=\{0,8,16,24,32,40\}=8\mathbb{Z}_{48}$
7. $\langle 3\rangle=\{1,3,7,9\}$
8. $\langle 5\rangle=\{1,5,7,11,13,17\}=\mathbb{Z}_{18}^*$
9. $\langle 7\rangle=\left\{\ldots,\frac{1}{49},\frac{1}{7},1,7,49,343\ldots\right\}$
10. $\langle i\rangle=\langle i\rangle=\{1,-1,i,-i\}$
11. $\langle 2i\rangle=\left\{\ldots,-\frac{1}{4},-\frac{i}{2},1,2i,-4,-8i,16\ldots\right\}$
12. $\langle e^{\pi i/4}\rangle=\left\{\pm1,\pm i,\pm\frac{\sqrt{2}}{2}\pm\frac{\sqrt{2}}{2}i\right\}$
13. $\langle e^{\pi i/6}\rangle=\left\{\pm1,\pm i,\pm\frac{1}{2}\pm\frac{\sqrt{3}}{2}i,\pm\frac{\sqrt{3}}{2}\pm\frac{1}{2}i\right\}$

23. 292 24. 2876 25. 1523 26. 206

3.4 Supplementary Exercises

1. False 3. False 5. True

6. $\{0\}$, $\{1,-1\}$, $\{1,-1\}$ 7. 1,2,3,4,6,8,12,24 8. 240

Chapter 4 : Permutation Groups

4.1 The Symmetric Group

1. (12453) 2. (14)(35) 3. (13)(25) 4. (1325) 5. (135)(24) 6. (253)

7. (14)(23) 8. (12)(56) 9. (1324) 10. (13254) 11. (134)(25) 12. (14)(235)

13. $(143)(25)$ 14. 1 15. 4 16. 2 17. (12) 18. (17352)

19. $(1532)(476)$ 20. *Show first that* $\sigma\tau\sigma^{-1}(\sigma(a_i)) = \sigma(a_{i+1})$.

4.2 The Alternating Group

1. $(14)(43)(35)(56)$
even
2. $(15)(56)(23)(34)$
even
3. $(16)(14)(12) = (12)(24)(46) = (16)(12)(14)(12)(14)$
odd

4. $(12)(17)(16)(14) = (14)(46)(67)(72) = (14)(15)(12)(17)(13)(12)(14)(12)(13)(16)(14)(15)$
even

5. $(14)(42)(26)(63)(37)$ odd

6. $A_5 : 1, (12)(34), (123), (12345)$; $A_6 : 1, (12)(34), (12)(3456), (123), (123)(456), (12345)$

9. No, 26 does not divide $\frac{8!}{2}$. 10. $(1,2,3,4)(5,6,7,8,9,10)$ 16. $(abc) = (ac)(ab)$

4.3 The Dihedral Group

1. $r = (12345), s = (25)(34)$

12. D_4 13. $D_2 \cong V_4$ 14. Z_4 15. Z_2 16. $Z_1 \cong 1$

17. D_4 18. D_4 19. D_4 20. $D_1 \cong Z_2$

4.4 Supplementary Exercises

1. $(a_1a_2 \dots a_n)^{-1} = (a_n \dots a_2a_1)$

2. $\{(13), (13)(24), (132), (134), (1324), (1342)\}$,
not subgroup

3. $\{1, (13), (14), (34), (134), (143)\}$,
subgroup

4. $\{1, (13), (134)\}$, not subgroup

8. $(123)(12) \neq (12)(123)$

16. Consider how rigid motions in $\mathbb{R}^3$ affect the four diagonals of the cube.

Chapter 5 : Cosets

5.1 Cosets

1. $\langle 8\rangle, 1 + \langle 8\rangle, 2 + \langle 8\rangle, 3 + \langle 8\rangle, 4 + \langle 8\rangle, 5 + \langle 8\rangle, 6 + \langle 8\rangle, 7 + \langle 8\rangle$

2. $\langle 3\rangle, 5\langle 3\rangle$ 3. $3\mathbb{Z}, 1 + 3\mathbb{Z}, 2 + 3\mathbb{Z}$ 4. $A_4, (12)A_4$

5. $A_n, (12)A_n$ 6. $D_4, (12)D_4, (14)D_4$ 7. all circles centered at the origin

8. $\langle(123)\rangle, (12)\langle(123)\rangle, (14)\langle(123)\rangle, (24)\langle(123)\rangle, (34)\langle(123)\rangle, (12)(34)\langle(123)\rangle, (13)(24)\langle(123)\rangle, (14)(23)\langle(123)\rangle$

9. *It is not well-defined.* 10. *If* $gh \in gH$, *then* $gh = (ghg^{-1})g$ *and* $ghg^{-1} \in H$.

11. *Show that* $g(H \cap K) = gH \cap gK$.

5.2 Lagrange's Theorem

1. 1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60
2. $|g|, |h| \mid |G|$
3. *Use Fermat's Little Theorem.*
5. 86, certainly composite
6. 1 maybe prime
7. 1, maybe prime
8. 1, maybe prime
9. 4, certainly composite
10. 1, maybe prime
11. 279, certainly composite
12. 1 maybe prime
13. 1, maybe prime
14. 375, certainly composite
15. 9, certainly composite
16. 1, maybe prime

5.3 Supplementary Exercises

2. $4^{\phi(15)} \equiv 4^{(5-1)(3-1)} \equiv 4^{4(2)} \equiv 4^8 \equiv 1 \pmod{15}$
6. *Note that if* $\gcd(a, b) = 1$, *then* $\phi(ab) = \phi(a)\phi(b)$.
9. *True*
10. *False*
11. *False*
12. *True*

Chapter 6 : Applied Algebra

6.1 Symmetric Key Cryptography

1. `"LAORYHAPDWK"`
2. `"WILLxTHISxBExONxTHExFINAL"`

6.2 Public Key Cryptography

For Exercises 1–4, encrypt the plaintext m using the RSA public key (n, e).

1. 2791
2. 769
3. 89518
4. 35362
5. 31
6. 2014
7. 14
8. 21712
9. 1997
10. 1103
11. 27331
12. 24649
13. 71
14. 2161
15. 37838567
16. 1566305611

6.3 Algebraic Coding Theory

1. not injective: $E(0) = E(8)$
2. 0.0022001875402126105875749513212 1
3. 0.99952075554642107658751002546 8
4. 0.18403174412961163394735042003026
5. 0.26424108696981268747849302071902
6. yes
7. no
8. $7d_1 + 9d_2 + \ldots + 7d_{11} \pmod{10}$
9. $(3,1,3,1,3,1,3,1,3,1,3,1) \cdot (0,5,0,0,0,0,0,3,0,4,2,6) = 24 \equiv 4 \pmod{10}$
10. $(3,1,3,1,3,1,3,1,3,1,3,1) \cdot (5,0,0,0,0,0,3,0,0,4,2,6) = 40 \equiv 0 \pmod{10}$
12. yes
13. no
14. no
15. yes
16. yes

6.4 The Hamming Metric

1. 3 2. 4 3. 3 4. 6 5. 2 6. 3

7. 2 8. 3 9. $d_{\min} = 2$, $k = 1$, $\ell = 0$

10. $d_{\min} = 1$, $k = 0$, $\ell = 0$ 11. $d_{\min} = 1$, $k = 0$, $\ell = 0$ 12. $d_{\min} = 2$, $k = 1$, $\ell = 0$

13. No identity

6.5 Linear Codes

1. (5,2)-code, nullity$(H) = 2$;
nul(H) = Span$\{10110, 00101\}$
$= \{00000, 10110, 00101, 10011\}$,
$G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

2. (6,2)-code, nullity$(H) = 2$;
nul(H) = Span$\{101\text{-}101, 010\text{-}111\}$
$= \{000\text{-}000, 101\text{-}101, 010\text{-}111, 111\text{-}010\}$,
$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$

3. (5,3)-code, nullity$(H) = 3$;
nul(H) = Span$\{10011, 01011, 00100\}$
$= \{00000, 10011, 01011, 00100, 11000, 10111, 01111, 11100\}$
$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$

4. (7,4)-code, nullity$(H) = 4$;
nul(H) = Span$\{111\text{-}0000, 100\text{-}1100, 010\text{-}1010, 110\text{-}1001\}$
$= \{000\text{-}0000, 111\text{-}0000, 100\text{-}1100, 010\text{-}1010, 110\text{-}1001, 011\text{-}1100, 101\text{-}1010, 001\text{-}1001,$
$110\text{-}0110, 010\text{-}0101, 100\text{-}0011, 001\text{-}0110, 101\text{-}0101, 011\text{-}0011, 000\text{-}1111, 111\text{-}1111\}$
$G = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

7. If there are no odd codewords, then you are done. If there is, modify the proof of Theorem 4.2.7.

6.6 Decoding

1. yes, $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $k = 2$, $\ell = 1$

2. yes, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, $k = 2$ $\ell = 1$

3. yes, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$, $k = 1$, $\ell = 0$

4. yes, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, $k = 1$, $\ell = 0$

5. 001 6. 101 7. 111 8. 011

9. 001, 01101 10. 110, 2 errors 11. 110, 2 errors 12. 110, 2 errors

6.7 Supplementary Exercises

1. $L \mapsto E$, $J \mapsto T$, $C \mapsto A$

3. $k = 6$, $\ell = 3$

7. For 20 information positions, at least 6 check bits are needed to ensure an error-correcting code.

Chapter 7 : Isomorphims

7.1 Isomorphisms

6. $a+bi \mapsto \begin{pmatrix} a & b \\ -a & b \end{pmatrix}$

9. $1 \mapsto n$

10. $1 \mapsto \zeta_n$

14. Assume $\mathbb{Q}$ is cyclic and find a fraction not generated by this generator

7.2 Cayley's Theorem

1. Show that $\varphi(e_G)h = h\varphi(e_G) = h$ for all $h \in H$.
2. If $G = \langle g \rangle$, then $H = \langle \varphi(g) \rangle$.
3. If $K \leq G$, then $\varphi(K) \leq H$.
4. If $K \leq G$, then $\varphi(K) \leq H$.
5. If $\sigma \in S_n$ is odd, then $\sigma \mapsto \sigma(n+1, n+2)$.

7.3 The Chinese Remainder Theorem

1. 12
2. 30
3. 5
4. 30
5. 6 abelian groups
6. 10 abelian groups
7. No
8. Yes
9. 17 (mod 55)
10. 80 (mod 850)
11. 214 (mod 2772)
12. 408 (mod 5720)

7.4 Products of Subsets

1. H is an internal direct product of $\langle 2 \rangle$ and $\langle 3 \rangle$.
2. No rotation commutes with any reflections except r^2.
3. Show that $G = HK$ and $H \cap K = 1$. 5
4. If $g = hk = h'k'$, then $h^{-1}h' = k(k')^{-1} \in H \cap K$.
5. Use induction.

7.5 Supplementary Exercises

1. There are 3 abelian and 2 nonabelian groups of order 8.
2. True
3. False

Chapter 8 : Quotient Groups and Homomorphisms

8.1 Normal Subgroups

1. $\{ \{1\}, \{(123), (132)\}, \{(12), (13), (23)\}$
   $Z(S_3) = 1$, $S_3' = S_3$
2. $\{ \{1\}, \{r^2\}, \{r, r^2\}, \{s, r^2 s\}, \{rs, r^3 s\}$
   $Z(D_4) = \{1, r^2\}$, $D_4' = \{1, r^2\}$
3. $\{ \{1\}, \{-1\}, \{i, -i\}, \{j, -j\}, \{k, -k\}$
   $Z(Q_8) = \{1, -1\}$, $Q_8' = \{1, -1\}$
4. $\{ \{0\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\},$
   $Z(\mathbb{Z}_9) = \mathbb{Z}_9$, $\mathbb{Z}_9' = 1$

5. $\{\{1\},\{r,r^2,r^3,r^4\},\{s,rs,r^2s,r^3s,r^4s\}$
$Z(D_5)=1$, $D_5'=\{1,r,r^2,r^3,r^4\}$

6. $\{\{1\},\{(12)(34),(13)(24),(14)(23)\},$
$\{(123),(142),(134),(243)\},\{(132),(124),(143),(234)\}$
$Z(A_4)=1$, $A_4'=\{1,(12)(34),(13)(24),(14)(23)\}$

8. $|gHg^{-1}|=|H|$

9. If $\langle g\rangle \trianglelefteq G$ and $x\in C(g)$, then $(yxy^{-1})g=g(yxy^{-1})$.

11. $g[x,y]g^{-1}=g(xyx^{-1}y^{-1})g^{-1}=(gxg^{-1}(gyg^{-1})(gx^{-1}g^{-1})(gy^{-1}g^{-1})=[gxg^{-1},gyg^{-1}]$

8.2 Quotient Groups

1. $G/H\cong\mathbb{Z}_2$

2. Not normal

3. Not normal

4. $G/H\cong\mathbb{Z}_2$

5. $G/H\cong\mathbb{Z}_5$

12. If $G=\langle g\rangle$, then $G/H=\langle gH\rangle$.

8.3 Homomorphisms

3. Each homomorphism φ is completely determined by where 1 maps to. There are 18 options for $\varphi(1)$. Do all of them work?

5. $\varphi(x)\varphi(y)=\varphi(xy)=\varphi(yx)=\varphi(y)\varphi(x)$

8.4 Kernels

1. $\ker\varphi=\{0\}$

2. $\ker\varphi=6\mathbb{Z}_{18}$

3. $\ker\varphi=\{1\}$

4. $\ker\varphi=\{0\}$

5. $\ker\varphi=\left\{\begin{pmatrix}a & b\\ c & -a\end{pmatrix}\,\middle|\, a,b,c\in\mathbb{R}\right\}$

6. $\ker\varphi=\left\{\begin{pmatrix}a & 0\\ c & d\end{pmatrix}\,\middle|\, a,c,d\in\mathbb{R}\right\}$

10. Disprove