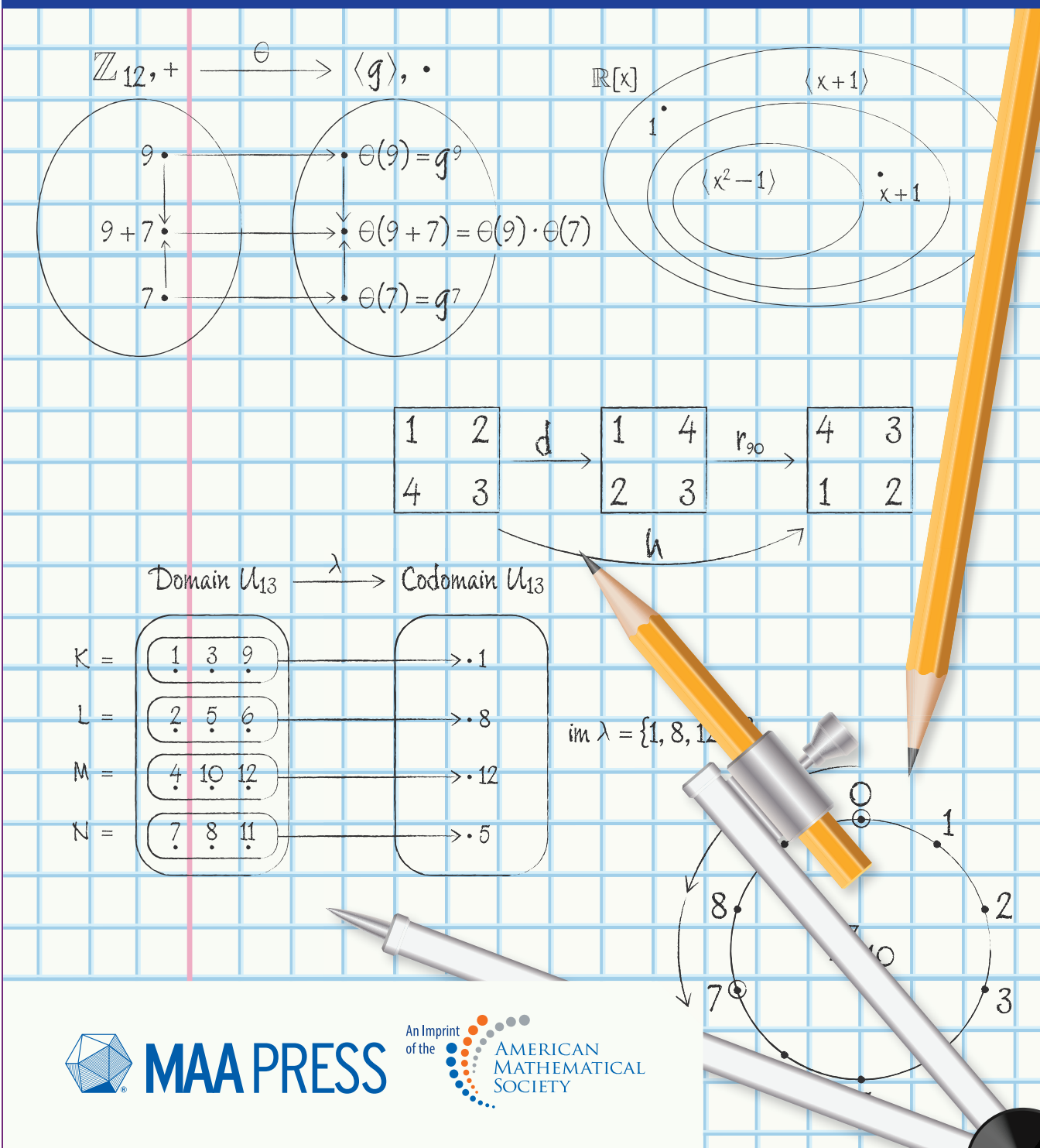
AMS / MAA | TEXTBOOKS VOL 72

A Friendly Introduction to Abstract Algebra

Ryota Matsuura



MAA PRESS
An Imprint of the American Mathematical Society

A Friendly Introduction to Abstract Algebra

AMS/MAA | TEXTBOOKS

VOL **72**

A Friendly Introduction to Abstract Algebra

Ryota Matsuura

MAA PRESS

An Imprint of the AMERICAN MATHEMATICAL SOCIETY

Providence, Rhode Island

2020 *Mathematics Subject Classification*. Primary 12-XX, 13-XX, 16-XX, 20-XX.

For additional information and updates on this book, visit
www.ams.org/bookpages/text-72

Library of Congress Cataloging-in-Publication Data

Names: Matsuura, Ryota, 1974- author.
Title: A friendly introduction to abstract algebra / Ryota Matsuura.
Description: Providence, Rhode Island : American Mathematical Society, [2022] | Series: AMS/MAA textbooks, 2577-1205 ; Volume 72 | Includes index.
Identifiers: LCCN 2021062966 | ISBN 9781470468811 (paperback) | ISBN 9781470470371 (ebook)
Subjects: LCSH: Algebra, Abstract. | AMS: Field theory and polynomials. | Commutative algebra. | Associative rings and algebras. | Group theory and generalizations.
Classification: LCC QA162 .M38 2022 | DDC 512/.02–dc23/eng20220228
LC record available at https://lccn.loc.gov/2021062966

---

10 9 8 7 6 5 4 3 2 1 27 26 25 24 23 22

Contents

Preface

"The only way to learn mathematics is to do mathematics."
Quote attributed to Paul Halmos

For the student

Embracing Halmos's quote above, this textbook provides you (the student) with opportunities to *do* mathematics: to perform experiments and grapple with problems; to formulate, test, and revise conjectures; to develop theories that bring coherence to observed results; and to express understanding using precise mathematical language. In essence, you get to experience mathematics as a mathematician does.

As its title suggests, this textbook is a friendly introduction to abstract algebra, a study of algebraic structures and relationships. I hope you'll not only read the textbook, but *actively engage* with it. As you work through the examples or exercises, you should seek patterns, make bold conjectures, and try proving them. Create your own examples to further your understanding. Ask your own questions. Have fun!

For the instructor

This textbook is intended to be used in a first-semester course in abstract algebra. Students should have completed Calculus I and II, not because they need the calculus content, but to acquire enough mathematical experience and maturity to handle the abstraction and proof writing that are part of the textbook. Familiarity with matrices is helpful, but not required. (Chapter 7 contains a brief introduction to matrices and covers matrix concepts that are needed in the textbook.) Prior proof-writing experience is *not* expected. In fact, an underlying goal of the textbook is to guide students toward writing clear and precise mathematical proofs. To support this goal, the textbook includes "Proof know-how," frequent, context-specific, short tips on proof writing.

Paul Halmos once said, "A good stack of examples, as large as possible, is indispensable for a thorough understanding of any concept." Following this advice, almost all concepts in this textbook are introduced through concrete examples. Ideas are foreshadowed, revisited, and developed over time. For instance, in an exercise about modular arithmetic (in Chapter 4), students compute the order of units modulo p and make conjectures about these orders. Eight chapters later (in Chapter 12), the notion of the order of a group element is formally defined. By then, students have seen enough examples so that the concept feels familiar.

Abstract algebra often acts as a "gateway" to upper-level mathematics courses and to a successful completion of a mathematics major. But it can seem impenetrable due to its (seemingly) theoretic nature. By taking a more concrete approach to the subject and by allowing students to develop their own understanding of the material, this textbook makes abstract algebra more accessible to more students.

Below, we highlight the pedagogical features of this textbook.

"Under the hood" perspective. This textbook provides students access to the "under the hood" work that mathematicians do. Rather than starting with a general theorem or definition (i.e., a finished product), the textbook lets students in on how that finished product is developed. In Chapter 13, for instance, we consider the cyclic group $\langle g \rangle$ generated by a group element g of order 12. We compute the product $g^9 \cdot g^7 = g^{9+7} = g^4$ in $\langle g \rangle$ and notice how this is just like the sum $9 + 7 = 4$ in $\mathbb{Z}_{12}$. Indeed, multiplication in $\langle g \rangle$ *feels like* addition in $\mathbb{Z}_{12}$. In Chapter 16, we build on this observation to motivate the definition of a *group isomorphism*. By providing access to the mathematical thinking that goes into the finished product, this textbook helps students make sense of the concepts in abstract algebra.

"How did you come up with that?" This is a question that I often get from abstract algebra students, especially when it comes to proof writing. Students can typically follow a proof that is presented to them, but they struggle with deriving the key steps on their own. This textbook addresses this issue through in-depth analyses of the proofs. In the proof of Theorem 19.14, for instance, there is a tricky step of coming up with an element $h = a^{-1}g$. Here's an excerpt from the "Proof know-how" following the proof:

> Coming up with the element $h = a^{-1}g$ employed the familiar "working backwards" technique. Our goal was to show that $g = ah$ for some $h \in H$, so we solved this equation for h by left-multiplying each side by a^{-1}, which yielded $h = a^{-1}g$. As before, this process of solving for h is scratch work and does *not* belong in the proof. Instead, the focus of the argument is showing that $g = ah$ for $h = a^{-1}g$.

Students, even mathematics majors, often have a (false) impression of the subject, that mathematicians produce new ideas out of thin air by writing down a theorem and effortlessly proving it. This textbook teaches students not only the content, but also the skills and know-how to do mathematics and create new ideas on their own.

Experience before formality. Providing students with *concrete experiences* is at the heart of this textbook, accomplished through example-driven exposition where theoretical concepts are introduced through examples. This approach makes the content more accessible to more students.

A key feature of this textbook is the set of exercises at the end of each chapter, in which students work on examples that lead to or reveal certain patterns. After working on such exercises, students are often asked to make a conjecture and/or prove a generalization. I tell my students that the role of the proof is *not* to convince, but to understand and explain. In other words, they shouldn't try to prove any theorem that they don't already believe is true. And this belief typically comes from concrete experiences that lead to the statement of the theorem. This is another instance of how, through this textbook, students experience mathematics as a mathematician does. Again, such an experience has the effect of making the subject more accessible to more students.

Accessible but still rigorous content. In this textbook, certain choices were made with the aim of making the content as accessible as possible, while still maintaining the mathematical rigor. For instance, we define a polynomial $f(x)$ to be *factorable* when $f(x) = p(x) \cdot q(x)$ with $\deg p(x), \deg q(x) < \deg f(x)$. (In other words, $f(x)$ is a product of "smaller" polynomials.) Otherwise, we say that $f(x)$ is *unfactorable*. This treatment is a bit unorthodox in a couple of ways. First, we use the terms *factorable* and *unfactorable*, rather than the more commonly used *reducible* and *irreducible*. Second, a more typical approach is to define *irreducible* polynomials as satisfying the following property: If $f(x) = p(x) \cdot q(x)$, then $\deg p(x) = 0$ or $\deg q(x) = 0$; and otherwise, $f(x)$ is said to be *reducible*. In this textbook, this property of irreducible (or unfactorable) polynomials is proved in Theorem 30.8.

These decisions about polynomials were made because the approach we take more closely resembles students' prior experiences with polynomials. In fact, many of the topics that we cover (e.g., polynomials, integers, the commutative law, putting on socks and shoes, to name a few) are already familiar to students. Whenever possible, we build on students' existing knowledge to make the content more accessible.

Note about rings

In this textbook, a ring will contain the multiplicative identity element *by definition*. (For instance, we do *not* consider the set $2\mathbb{Z}$ of even integers to be a ring.) We do so for two reasons. First, we wanted our definition of a ring to closely mimic what we observe in the ring of integers $\mathbb{Z}$. Second, every relevant example of a ring that we examine contains the multiplicative identity.

Road map

There are 37 chapters in the textbook. Depending on your students' background and/or the structure of your algebra course, here are some suggested road maps:

- Chapters 1 through 25 cover group theory. If students have had prior experience with proof writing, you may choose to omit Chapters 1 and 2. Similarly, students may have already studied the notion of divisibility (of integers), particularly if they have taken an introductory number theory course. In such a case, Chapter 3 may be skipped as well. I recommend *not* skipping Chapter 4 on modular arithmetic, however, since many group-theoretic concepts are introduced in that chapter.
- Chapter 7 contains a brief introduction to matrices and covers matrix concepts that are needed in the textbook. If your students have had a linear algebra course, then Chapter 7 may be omitted or assigned to the students to read on their own as a refresher.
- Chapters 26 through 37 cover ring theory, with an emphasis on polynomial rings. It is possible to end the course with Chapter 35, where we complete the proof of Theorem 35.1 (i.e., $F[x]/\langle g(x)\rangle$ is a field if and only if $g(x)$ is unfactorable).
- Chapters 36 and 37 guide the students to prove Theorem 35.1 again using the notion of maximal ideals. It is a lovely proof, so I recommend its inclusion in your course, if time permits.

Acknowledgments

This textbook is based on the materials that I developed for Math 252, a first-semester abstract algebra course at St. Olaf College. I want to thank all of my Math 252 students (past, present, and future) for their hard work, enthusiasm, and patience, as I put them through the experience of learning mathematics by *doing* mathematics.

My mathematical mentors, David Rohrlich, Glenn Stevens, and Al Cuoco, taught me how to think mathematically and what it means to experience mathematics as a mathematician does. I hope their impact can be seen on every page of this textbook.

I am grateful to Steve Kennedy from the Mathematical Association of America for seeing the potential in the first draft of this textbook and for his continual guidance and encouragement. I also want to thank the MAA editors Will Green, for his thorough review of an early manuscript, and Suzanne Larson, for her support throughout the writing and production stages. Moreover, I appreciate the meticulous work by Arlene O'Sean and Chris Thivierge from the American Mathematical Society to make this textbook a reality.

My kids, Elizabeth and Anita, lent their names to the recurring characters in this textbook. They take full credit for the insightful comments that the characters make but bear no responsibility for their errors.

As always, I am most indebted to my wife, Sarah, the real algebraist of the family. When I showed her my Math 252 course notes, she immediately said, "Send it to Steve Kennedy." She knew that I could write an algebra textbook long before I believed it myself.

Unit I: Preliminaries

This book is about algebraic structures called *groups* and *rings*. We'll also be *proving* various facts about groups and rings, which is a mathematician's way of understanding and explaining what they observe and why those observations are true. The first three chapters provide preliminary knowledge needed for our exploration throughout the book. Chapter 1 introduces the notion of proofs and some foundational proof-writing techniques. Chapter 2 deals with *sets* and *subsets*, which are central to our study, since groups and rings are sets that are equipped with an operation or two. (We'll say much more about this later.) Chapter 3 is about *divisibility* (e.g., 4 is a divisor of 24), which will be a useful and recurring tool as we study the properties of groups and rings.

Below is a taste of what you'll be able to accomplish in this unit:

- Prove that n is odd if and only if n^2 is odd. (Here, n is an integer.)
- Learn how to show that two sets are equal, i.e., they contain the same elements.
- Understand why consecutive integers such as 451 and 452 are relatively prime, i.e., their greatest common divisor is 1. (**Note:** It's because $451 \cdot (-1) + 452 \cdot 1 = 1$.)

1

Introduction to Proofs

In mathematics, a *proof* is a series of logical arguments that explains why something is true. The goal of this opening chapter is to begin to understand how to write mathematical proofs. In particular, we will introduce various proof-writing techniques which will be used throughout the textbook. Note that many of the proofs in the chapter involve the set of *integers*, which includes the counting numbers (1, 2, 3, 4, ...), their negatives (-1, -2, -3, -4, ...), and zero (0).

1.1 Proving an implication

Consider the following statement: *If n is an odd integer, then n^2 is odd.* To better understand this statement, we look at a few examples:

- If $n = 7$ is odd, then $n^2 = 49$ is also odd.
- If $n = 213$ is odd, then $n^2 = 45{,}369$ is also odd.
- If $n = -1{,}081$ is odd, then $n^2 = 1{,}168{,}561$ is also odd.

The importance of these examples cannot be overstated. Concrete examples help us make sense of an abstract statement. Sometimes, they provide insight into *why* the statement is true. But, as we will see shortly, writing a proof is different from creating examples.

Before proceeding, let's be more precise about what it means for an integer to be *odd* (and also *even*). Notice that 7 is odd, because $7 = 2 \cdot 3 + 1$; i.e., when we put 7 into groups of two, there is a remainder of 1. However, 10 is even, because $10 = 2 \cdot 5$; i.e., 10 can be put into groups of two without a remainder.

Definition 1.1 (Odd and Even)**.** Let n be an integer. Then:

- n is *odd* when $n = 2k + 1$ for some integer k.
- n is *even* when $n = 2k$ for some integer k.

Example 1.2. We categorize the following integers as odd or even:

- $213 = 2 \cdot 106 + 1$ so that 213 is odd.
- $-1{,}081 = 2 \cdot (-541) + 1$ so that $-1{,}081$ is odd.
- $-314 = 2 \cdot (-157)$ so that -314 is even.
- $0 = 2 \cdot 0$ so that 0 is even.

Now, back to our statement: *If n is an odd integer, then n^2 is odd.* This is an example of an *implication*, i.e., an "if ..., then ..." statement. The if-part is called the *hypothesis* ("n is an odd integer") and the then-part is called the *conclusion* ("n^2 is odd"). To *prove* an implication, we take the following steps:

Proof know-how. To *prove* an implication:

(1) *Assume* that the hypothesis is true.

(2) *Show* that the conclusion is true.

Equipped with this know-how, let's brainstorm how to write this proof. Our initial step is to assume that the hypothesis is true, which means that the first sentence of the proof should be: Assume n is an odd integer. Our eventual goal is to show that the conclusion is true, which means that the last sentence of the proof should be: Thus, n^2 is odd. What about the second sentence of the proof? We know that n is odd, so Definition 1.1 suggests that we write: Then $n = 2k + 1$ for some integer k.

We now have $n = 2k + 1$, but how do we proceed from here? Knowing this next step is often the most difficult part of proof writing. A useful tip is to *look ahead* at our goal, which, in this proof, is to show that n^2 odd. Hence, we compute $n^2 = (2k+1)^2 = 4k^2+4k+1$ and show that the expression $4k^2+4k+1$ can be written as $2\cdot(\text{integer})+1$. Then, we can conclude that n^2 is odd.

Without further ado, here are the first theorem and proof of this book!

Theorem 1.3. *If n is an odd integer, then n^2 is odd.*

PROOF. Assume n is an odd integer. Then $n = 2k + 1$ for some integer k. Therefore,

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

Hence, $n^2 = 2 \cdot (2k^2 + 2k) + 1$, where $2k^2 + 2k$ is an integer. Thus, n^2 is odd. ■

Proof know-how. In the proof above, we start with an arbitrary odd integer n, rather than a specific one like 7, 213, or $-1{,}081$. Here, "arbitrary" means we are not making any assumptions about n, other than that it is an odd integer. Then we proved that n^2 is odd. Because the theorem is true for an *arbitrary* odd integer, we may conclude that it is true for *all* odd integers.

1.2 Proof by cases

Consider the statement: *If n is an integer, then $n^2 + n$ is even.* As before, we begin by creating some concrete examples. Since the only assumption about n is that it is an

integer, we consider the cases where (1) n is odd and (2) n is even:

- If $n = 7$ (i.e., n is odd), then $n^2 + n = 56$ is even.
- If $n = 213$ (i.e., n is odd), then $n^2 + n = 45{,}582$ is even.
- If $n = 10$ (i.e., n is even), then $n^2 + n = 110$ is even.
- If $n = -314$ (i.e., n is even), then $n^2 + n = 98{,}282$ is even.

These examples suggest a proof technique called *proof by cases*. In this method, we split the given scenario into multiple cases and then prove the statement for each case. It is important that the cases considered cover all the possibilities. For instance, if n is an integer, then the cases (1) n is odd and (2) n is even would suffice, since every integer is either odd or even.

Theorem 1.4. *If n is an integer, then $n^2 + n$ is even.*

PROOF. Assume n is an integer. We consider the two cases: (1) n is odd and (2) n is even.

Case (1). Suppose n is odd, so that $n = 2k + 1$ for some integer k. Then

$$n^2 + n = (2k + 1)^2 + (2k + 1) = 4k^2 + 6k + 2 = 2 \cdot (2k^2 + 3k + 1),$$

where $2k^2 + 3k + 1$ is an integer. Thus, $n^2 + n$ is even.

Case (2). Suppose n is even, so that $n = 2k$ for some integer k. Then

$$n^2 + n = (2k)^2 + 2k = 4k^2 + 2k = 2 \cdot (2k^2 + k),$$

where $2k^2 + k$ is an integer. Thus, $n^2 + n$ is even. ■

Remark. Notice how Theorem 1.4 is a statement about *all* integers. We prove it by showing that it is true for an *arbitrary* integer n. In fact, the first sentence of the proof, "Assume n is an integer," may be considered as a shorthand for "Assume n is an arbitrary integer."

Next, consider the statement: *If n is an integer, then $n^2 + 1$ is not a multiple of 3.* As we examine some concrete examples, observe that when an integer n is divided by 3, the possible remainders are 0, 1, or 2. (**Note:** For a more in-depth discussion of remainders, see Theorem 12.16.)

- If $n = 12 = 3 \cdot 4 + 0$ (i.e., remainder = 0), then $n^2 + 1 = 145$ is *not* a multiple of 3.
- If $n = 13 = 3 \cdot 4 + 1$ (i.e., remainder = 1), then $n^2 + 1 = 170$ is *not* a multiple of 3.
- If $n = 14 = 3 \cdot 4 + 2$ (i.e., remainder = 2), then $n^2 + 1 = 197$ is *not* a multiple of 3.

In the above calculation, 145 is *not* a multiple of 3, because $145 = 3 \cdot 48 + 1$. In other words, when we divide 145 by 3, there is a non-zero remainder. Similar arguments can be made to explain why neither 170 nor 197 is a multiple of 3. These three examples suggest the three cases that we consider in the proof below.

Theorem 1.5. *If n is an integer, then $n^2 + 1$ is not a multiple of 3.*

PROOF. Assume n is an integer. When n is divided by 3, the possible remainders are 0, 1, or 2.

Case (1). The remainder is 0, so that $n = 3 \cdot q + 0$ for some integer q. Then

$$n^2 + 1 = (3q)^2 + 1 = 9q^2 + 1 = 3 \cdot (3q^2) + 1.$$

Hence, when we divide $n^2 + 1$ by 3, there is a remainder of 1. Thus, $n^2 + 1$ is not a multiple of 3.

Case (2). The remainder is 1, so that $n = 3 \cdot q + 1$ for some integer q. Then

$$n^2 + 1 = (3q + 1)^2 + 1 = 9q^2 + 6q + 2 = 3 \cdot (3q^2 + 2q) + 2.$$

Hence, when we divide $n^2 + 1$ by 3, there is a remainder of 2. Thus, $n^2 + 1$ is not a multiple of 3.

Case (3). The remainder is 2. (This case is left for you as an exercise at the end of the chapter.) ■

1.3 Contrapositive

Consider the statement: *If n^2 is odd, then n is odd.* (Here, n is an integer.) To prove it, we might start by assuming the hypothesis, i.e., n^2 is odd. Then $n^2 = 2k + 1$ for some integer k. We wish to show that n is odd, but we're stuck, since solving $n^2 = 2k + 1$ for n requires us to take the square root of $2k + 1$. Yikes!

We will introduce a new proof technique to handle a statement like the following: *If n^2 is odd, then n is odd.* But first, consider these four implications:

(a) If I live in Tokyo, then I live in Japan.

(b) If I live in Japan, then I live in Tokyo.

(c) If I don't live in Tokyo, then I don't live in Japan.

(d) If I don't live in Japan, then I don't live in Tokyo.

Statement (a) is true, because Tokyo is a city inside Japan. For the same reason, statement (d) is also true; i.e., if I live outside of Japan, then I can't possibly live in Tokyo. However, statements (b) and (c) are false, because I could be living in Osaka, for example.

Note how (d) is obtained from (a) by swapping the hypothesis and conclusion and negating both of them; and (a) is obtained from (d) in the same way. Thus, (a) and (d) are said to be *contrapositives* of each other. Similarly, (b) and (c) are contrapositive pairs. The key fact about contrapositives is that they are equivalent; i.e., proving one ensures that the other must be true also.

Here, to *negate* a statement means to write down its opposite. Thus, when we negate "I live in Tokyo," we obtain its *negation*: "I don't live in Tokyo." Observe that if a statement is true, then its negation is false; and if a statement is false, then its negation is true.

Example 1.6. When $n = 7$, then the statement "n is odd" is true, and its negation "n is *not* odd" is false. Moreover, when $n = 6$, the statement "n^2 is odd" is false, and its negation "n^2 is *not* odd" is true.

Back to our statement: *If n^2 is odd, then n is odd.* Since it's difficult to prove this implication directly, we can prove its contrapositive: *If n is* not *odd, then n^2 is* not *odd*; or more succinctly:

Theorem 1.7. *If n is even, then n^2 is even.*

PROOF. Given in the exercises. ■

Proof know-how. Proving "If p, then q" is exactly the same as proving the contrapositive: "If not q, then not p." (Choose the easier one!) Here, p and q are statements such as "n is odd."

Example 1.8. Let $ABCD$ be a quadrilateral, and consider the statement: *If $ABCD$ is a rectangle, then $ABCD$ is a square.* This statement is false, since there exist rectangles that are *not* squares. The contrapositive is: *If $ABCD$ is not a square, then $ABCD$ is not a rectangle.* This is also false for the same reason; i.e., even if $ABCD$ is not a square, it can still be a rectangle. As expected, the statement and contrapositive have the same truth value (they are both false), as they are equivalent.

1.4 Proof by contradiction

Proof by contradiction is another technique for proving an implication, i.e., an "if ..., then ..." statement. Here are the steps of this proof method:

(1) Assume that the hypothesis is true (as usual).

(2) Also assume that the conclusion is false, or equivalently, that the negation of the conclusion is true.

(3) Obtain a *contradiction*, i.e., an absurd outcome. This would indicate that the conclusion couldn't have been false, and so it must be true.

Consider the statement: *If n^2 is even, then n is even.* Let's prove this using proof by contradiction. To start, we assume that the hypothesis is true, which means that the first sentence of the proof should be: Assume n^2 is even. Next, we must assume that the negation of the conclusion is true: Assume n is *not* even, i.e., n is odd. To complete the proof, we must obtain a contradiction. Knowing which contradiction to derive is typically the most challenging aspect of proof by contradiction. Here, we will show that n^2 is odd (because n is odd), which contradicts our assumption that n^2 is even.

Theorem 1.9. *Let n be an integer. If n^2 is even, then n is even.*

PROOF. Assume n^2 is even. Also assume *for contradiction* that n is odd. Since n is odd, Theorem 1.3 implies that n^2 is odd. But this contradicts the fact that n^2 is even. Hence, n cannot be odd. Thus, n is even. ■

Proof know-how. In a proof by contradiction, we make two assumptions: (1) The hypothesis is true and (2) the negation of the conclusion is true. The phrase "for contradiction" (as seen in the above proof) is often used with the negation of the conclusion to differentiate between these two assumptions.

Here is another example. Note that a *rational number* is a fraction of the form $\frac{m}{n}$ where m and n are integers (and n is non-zero, since we cannot divide by zero).

Theorem 1.10. *If r is a rational number, then $r^2 \neq 2$.*

PROOF. Assume r is a rational number. Also assume *for contradiction* that $r^2 = 2$. Let m and n be integers such that $r = \frac{m}{n}$ is a reduced fraction; i.e., m and n do not share a common factor greater than 1. Then $r^2 = 2$ implies $\frac{m^2}{n^2} = 2$, so that $m^2 = 2n^2$. Thus m^2 is even. Then m is even by Theorem 1.9, so that $m = 2k$ for some integer k. But $2n^2 = m^2 = (2k)^2 = 4k^2$, so that $2n^2 = 4k^2$. Dividing both sides of $2n^2 = 4k^2$ by 2 yields $n^2 = 2k^2$. Thus, n^2 is even and so n is even. Now, m and n are both even, which implies that they have a common factor of 2. However, this contradicts the fact that m and n do not share a common factor greater than 1. Hence, r^2 cannot equal 2. Thus, $r^2 \neq 2$. ■

Remark. Proof by contradiction is different from proving the contrapositive of a given statement. The contrapositive of the above theorem is "If $r^2 = 2$, then r is not a rational number." This is not what we proved in the sample proof above.

1.5 If and only if

In this chapter, we studied these two implications:

- If n is odd, then n^2 is odd.
- If n^2 is odd, then n is odd.

Each is obtained from the other by swapping the if-part and the then-part. Thus, each is called the **converse** of the other. (**Careful:** They're *not* contrapositives of each other. Why not?) As a shorthand, we can combine the two and write: n is odd **if and only if** n^2 is odd. So, when you're asked to prove an "if and only if" statement, you'll have to prove two implications.

Example 1.11. Here is another pair of statements that are converses of each other:

- If I live in Tokyo, then I live in Japan. (True)
- If I live in Japan, then I live in Tokyo. (False)

As you can see, an implication can be true even though its converse is false.

Example 1.12. Similar to Example 1.11, below is a pair of statements that are converses of each other where one is true and the other is false.

- If n is positive, then n^2 is positive. (True)
- If n^2 is positive, then n is positive. (False)

The first implication is true, but the second one is false. With $n = -3$, we see that even though $n^2 = 9$ is positive, $n = -3$ is *not* positive.

1.6 Counterexample

Consider the statement: *If n is prime, then $2^n - 1$ is prime.* As usual, let's create some concrete examples by letting n take on the values of the first few prime numbers.

- If $n = 2$, then $2^2 - 1 = 3$ is prime.
- If $n = 3$, then $2^3 - 1 = 7$ is prime.
- If $n = 5$, then $2^5 - 1 = 31$ is prime.
- If $n = 7$, then $2^7 - 1 = 127$ is prime.

So far, so good. But does this mean that the statement is true? Not necessarily. In order for the statement to be true, the expression $2^n - 1$ must be prime for *every* prime n. In other words, if we can find just one counterexample, i.e., an example that invalidates the statement, then we can conclude that the statement is false. Here is a counterexample: $n = 11$ is prime (which satisfies the hypothesis), but $2^{11} - 1 = 2{,}047 = 23 \cdot 89$ is *not* prime (which fails the conclusion). Thus, the implication is false.

Proof know-how. To show that an implication is false, we only need to find one counterexample. Thus, disproving an implication (when it's false) tends to be easier than proving one (when it's true).

Example 1.13. Consider the statement: *If n is an odd prime, then $n + 2$ is prime.* Many values of n serve as valid examples of this statement, as shown below:

- If $n = 3$, then $n + 2 = 5$ is prime.
- If $n = 11$, then $n + 2 = 13$ is prime.
- If $n = 101$, then $n + 2 = 103$ is prime.

However, the statement is false because we can find a counterexample: $n = 7$ is an odd prime (which satisfies the hypothesis), but $n + 2 = 9$ is *not* prime (which fails the conclusion).

Exercises

1. **Prove:** Let m and n be integers. If m is odd and n is even, then $m + n + 3$ is even.
2. **Prove:** Let m and n be integers. If m and n are both odd, then mn is odd.

   (This exercise is referenced in Chapter 2, Exercise #17.)
3. Consider the statement: *If n is an odd integer, then n is a prime number.* Which value of n is a counterexample showing that the statement is false? Explain.
   - $n = 13$.
   - $n = 15$.
4. Consider the statement: *If n is an odd integer, then $n = 4k + 1$ for some integer k.* Give a counterexample to show that this is false. Explain why your counterexample invalidates the statement.

5. Consider the statement: *If $n \neq 4k + 1$ for any integer k, then n is an even integer.* Give a counterexample to show that this is false. Explain why your counterexample invalidates the statement.

6. Consider the statement: *If n is not a multiple of 4, then n^2 is not a multiple of 4.* Give a counterexample to show that this is false. Explain why your counterexample invalidates the statement.

7. Consider the statement: *If n^2 is a multiple of 4, then n is a multiple of 4.* Give a counterexample to show that this is false. Explain why your counterexample invalidates the statement.

8. Consider the statement: $n^2 + n + 11$ *is a prime for all positive integers n.*

   (a) Rewrite the statement in the form of an implication (i.e., "If ..., then ...").

   (b) Show that the statement is false by exhibiting a counterexample.

9. **Prove:** If n is an integer, then $5n^2 - n + 2$ is even.

10. Complete Case (3) in the proof of Theorem 1.5.

11. Consider the statement: *If n is even, then $n + 2$ is even.*

    (a) Write the converse of this statement. Is the converse true? Why or why not?

    (b) Write the contrapositive of this statement. Is the contrapositive true? Why or why not?

12. Consider the statement: *If n is a multiple of 6, then n is a multiple of 3.*

    (a) Write the converse of this statement. Is the converse true? Why or why not?

    (b) Write the contrapositive of this statement. Is the contrapositive true? Why or why not?

13. Consider the statement: *If m and n are both even, then $m + n$ is even.*

    (a) Write the converse of this statement. Is the converse true? Why or why not?

    (b) Write the contrapositive of this statement. Is the contrapositive true? Why or why not?

14. (a) Come up with a true implication whose converse is false.

    (b) Write the contrapositive of your implication in part (a). Is the contrapositive true? Explain.

15. Consider the statement: *If m and n are integers, then $4m + 6n \neq 1$.*

    (a) What is the hypothesis of this implication?

    (b) Write down the negation of its conclusion.

    (c) Suppose $m = 17$ and $n = -10$. Is $4m + 6n$ even or odd?

    (d) Choose your own integers m and n. Is $4m + 6n$ even or odd?

16. **Prove:** If m and n are integers, then $4m + 6n \neq 1$. (See Exercise #15.)

17. Consider the statement: *If x is rational and y is irrational, then $x + y$ is irrational.*

    Note: The definition of a rational number was given prior to Theorem 1.10. Moreover, an *irrational number* is a (real) number that is not rational.

    (a) First, create an example that illustrates this statement.
    (b) What is the hypothesis of this implication?
    (c) Write down the negation of its conclusion.
    (d) Suppose $x = \frac{3}{5}$ and $x + y = \frac{6}{7}$. Find the value of y. Is y rational or irrational?
    (e) Repeat part (d), but with $x = \frac{1}{7}$ and $x + y = \frac{3}{10}$.

18. **Prove:** If x is rational and y is irrational, then $x + y$ is irrational. (See Exercise #17.)

19. **Prove:** If n is an even integer, then n^2 is even.

    Recall: This is the contrapositive of "If n^2 is odd, then n is odd." When proving an implication, it is often easier to prove the contrapositive instead.

20. **Prove:** If n is an odd integer, then $n^2 = 8k + 1$ for some integer k.

    Hint: Theorem 1.4 may be useful. If you use the theorem in your proof, you should cite it like this: "By Theorem 1.4, [blah blah blah]."

21. **Prove:** If x and y are positive numbers, then $\sqrt{x+y} \neq \sqrt{x} + \sqrt{y}$.

22. **Prove:** If $n + 1$ pigeons are placed in n holes, then there is a hole containing at least two pigeons.

    Note: This is often called the "Pigeonhole principle".

23. **(Challenge) Prove:** Let n be a positive integer. If $2^n - 1$ is prime, then n is prime.

2

Sets and Subsets

Sets and subsets play a critical role in the study of abstract algebra. A set is a collection of things (called *elements*) such as numbers or people. And a subset is a set whose elements are all contained in another set. For example, the set of all people in Tokyo is a subset of the set of all people in Japan.

In this chapter, we will learn how to work with sets in a mathematical setting. For instance, we'll see how to describe sets using mathematical symbols. We will examine the set of integers, arguably the most important example in this textbook. We will also introduce the notion of *closure* and how to show that two sets are equal. Both concepts will be revisited often when we study *groups* and *rings* later in the textbook.

2.1 What is a set?

In mathematics, a *set* is a collection of objects (such as numbers, people, etc.) that are different from each other. For example,

$$\{\text{Sarah, Elizabeth, Anita, Ryota}\}$$

is the set of people in my family. The members of a set are called *elements*. Notice how a set is defined by listing its elements in between curly brackets $\{\cdots\}$, and the elements are separated by commas. In a set, the order in which the elements are listed does *not* matter. Thus, the set {Anita, Elizabeth, Ryota, Sarah} is the same as the one written above, since they contain the same elements.

Example 2.1. Consider the sets $S = \{2, 4, 6, 8\}$ and $T = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Note how every element of S is also contained in T. Thus, we say that S is a subset of T. However, T is *not* a subset of S. For instance, 3 is an element of T but not of S.

Definition 2.2 (Subset)**.** Let S and T be sets. We say S is a *subset* of T, denoted $S \subseteq T$, if every element of S is also contained in T.

Example 2.3. Any set S is a subset of itself, i.e., $S \subseteq S$, because every element of S is also contained in S.

Example 2.4. The *empty set* is a set containing *no* element. It is often denoted $\emptyset$ or $\{\}$. The empty set is a subset of any set S; i.e., $\emptyset \subseteq S$.

Example 2.5. A set need not contain the same "type" of elements. For instance, here is a set that contains both numbers and people:

$$\{\text{Sarah, } 1, \text{ } 2, \text{ Elizabeth, } 3, \text{ } 4, \text{ Anita, } 5, \text{ } 6, \text{ Ryota, } 7, \text{ } 8\}.$$

That being said, almost all examples of sets in this textbook will contain the same "type" of elements.

Example 2.6. Recall that the elements of a set must be different from each other. In other words, duplicate elements are not counted in a set. For instance, the set $\{1, 1, 2, 2, 3, 3\}$ is the same as the set $\{1, 2, 3\}$. Both sets are said to contain three elements.

2.2 Set of integers and its subsets

Recall from Chapter 1 that the set of *integers* includes the counting numbers (1, 2, 3, 4, ...), their negatives ($-1, -2, -3, -4, \ldots$), and zero (0). This set is denoted by $\mathbb{Z}$, which stands for the German word *Zahlen* ("numbers"). In other words,

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

Unlike the sets we encountered in Section 2.1, $\mathbb{Z}$ is an *infinite set*; i.e., it contains infinitely many elements.

We use the shorthand $3 \in \mathbb{Z}$ to say that 3 is an element of $\mathbb{Z}$ (or 3 is contained in $\mathbb{Z}$). We also write $\frac{2}{5} \notin \mathbb{Z}$ to say that $\frac{2}{5}$ is *not* an element of $\mathbb{Z}$ (or $\frac{2}{5}$ is *not* contained in $\mathbb{Z}$).

Example 2.7. We have $\frac{2}{5}, \frac{8}{5} \notin \mathbb{Z}$. But their sum $\frac{2}{5} + \frac{8}{5} = 2$ is an element of $\mathbb{Z}$. Compare this with the notion of *closure*, which is discussed in Section 2.3.

Example 2.8. Consider the following subset of $\mathbb{Z}$:

$$A = \{3, 5, 7, \ldots, 17, 19\}.$$

Suppose you were asked to write down all of its elements. You might say

$$A = \{3, 5, 7, 9, 11, 13, 15, 17, 19\},$$

i.e., the set of odd integers between 3 and 19. But it turns out that

$$A = \{3, 5, 7, 11, 13, 17, 19\},$$

i.e., the set of primes between 3 and 19. You would be justified in thinking that this is unfair, since the ellipsis (...) left some ambiguity about the remaining elements of the set.

Example 2.8 invokes a need for a more precise way to describe a set, which we will examine below.

Example 2.9. Consider the following subset of $\mathbb{Z}$:

$$B = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97\}.$$

Set B contains all primes less than 100, which is more descriptive (and succinct!) than listing all of its elements. Like set B, we will often consider a set of elements that satisfy a certain *property*. Here is a way to describe set B symbolically:

$$B = \{n \in \mathbb{Z} \mid 2 \leq n \leq 100 \text{ and } n \text{ is prime}\}.$$

The "$n \in \mathbb{Z}$" tells us where the elements of set B come from; i.e., it implies that B is a subset of $\mathbb{Z}$. The symbol | means "such that." But be careful: This symbol should be used as "such that" inside $\{\cdots\}$ only. The property that must be satisfied by all elements of B is "$2 \leq n \leq 100$ and n is prime." For example, $45 \notin B$, because it is not prime; and $101 \notin B$, because it is greater than 100.

Example 2.10. Consider the set $S = \{n \in \mathbb{Z} \mid 0 < n < 25 \text{ and } n \text{ is a multiple of } 3\}$. Hence, S contains integers n with the property that $0 < n < 25$ and n is a multiple of 3. If we list all of its elements, we have $S = \{3, 6, 9, 12, 15, 18, 21, 24\}$. Note that $S \subseteq \mathbb{Z}$.

Example 2.11. Consider the set $T = \{n \in \mathbb{Z} \mid 0 < 4n - 1 < 20\}$, which is also a subset of $\mathbb{Z}$. To find the elements of T, consider the following calculations:

- With $n = 1$, we have $4 \cdot 1 - 1 = 3$, which is between 0 and 20.
- With $n = 2$, we have $4 \cdot 2 - 1 = 7$, which is between 0 and 20.
- With $n = 3$, we have $4 \cdot 3 - 1 = 11$, which is between 0 and 20.
- With $n = 4$, we have $4 \cdot 4 - 1 = 15$, which is between 0 and 20.
- With $n = 5$, we have $4 \cdot 5 - 1 = 19$, which is between 0 and 20.

Note that the condition $0 < 4 \cdot 2 - 1 < 20$ does not imply that $4 \cdot 2 - 1 \in T$. Rather, we have $2 \in T$. Thus, when we list the elements of T, we obtain $T = \{1, 2, 3, 4, 5\}$.

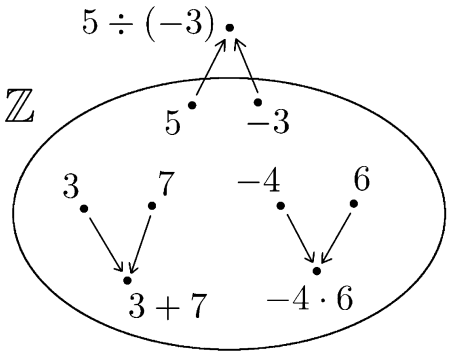
Example 2.12. An important subset of $\mathbb{Z}$ is the set of *natural numbers*, denoted $\mathbb{N}$, containing the positive integers. Thus, $\mathbb{N} = \{n \in \mathbb{Z} \mid n > 0\} = \{1, 2, 3, 4, \ldots\}$. We will use both $\mathbb{Z}$ and $\mathbb{N}$ throughout the textbook.

2.3 Closure

Any two elements in $\mathbb{Z}$ can be added together to obtain another element of $\mathbb{Z}$. For example, given $3, 7 \in \mathbb{Z}$, we find that their sum $3 + 7 = 10$ is also an element of $\mathbb{Z}$. More generally, we have the following: If $a, b \in \mathbb{Z}$, then $a + b \in \mathbb{Z}$. Therefore, we say that the set $\mathbb{Z}$ is *closed* under addition.

Similarly, $\mathbb{Z}$ is closed under multiplication. For instance, given $-4, 6 \in \mathbb{Z}$, we see that their product $-4 \cdot 6 = -24$ is also in $\mathbb{Z}$. More generally, we have the following: If a, $b \in \mathbb{Z}$, then $a \cdot b \in \mathbb{Z}$. However, $\mathbb{Z}$ is *not* closed under division. As a counterexample,

note that $5, -3 \in \mathbb{Z}$, but $5 \div (-3) \notin \mathbb{Z}$. This concept of *closure* is depicted by the figure below:



Closure is an important property of a set and operation that will play a prominent role in our study of groups and rings later in the textbook.

Example 2.13. Consider the following subset of $\mathbb{Z}$:

$$3\mathbb{Z} = \{n \in \mathbb{Z} \mid n = 3k \text{ where } k \in \mathbb{Z}\}.$$

Therefore, $3\mathbb{Z}$ contains all integer multiples of 3; i.e.,

$$3\mathbb{Z} = \{\ldots, -12, -9, -6, -3, 0, 3, 6, 9, 12, \ldots\}.$$

We claim that $3\mathbb{Z}$ is closed under addition; i.e., the sum of any two elements of $3\mathbb{Z}$ is also contained in $3\mathbb{Z}$. For instance, we have $15, 24 \in 3\mathbb{Z}$. Their sum is $15 + 24 = 39$, which is also an element of $3\mathbb{Z}$.

Digging further into Example 2.13, we see that $15 \in 3\mathbb{Z}$, because $15 = 3 \cdot 5$, and we see that $24 \in 3\mathbb{Z}$, because $24 = 3{\cdot}8$. Their sum is $15{+}24 = 3{\cdot}5{+}3{\cdot}8 = 3{\cdot}(5{+}8) = 3{\cdot}13$ so that $15{+}24$ is also a multiple of 3. Using this example as a guide, we prove the following theorem.

Theorem 2.14. *The set $3\mathbb{Z}$ is closed under addition.*

PROOF. Assume $m, n \in 3\mathbb{Z}$. Then $m = 3k$ and $n = 3j$ where $k, j \in \mathbb{Z}$. We have

$$m + n = 3k + 3j = 3(k + j) \in 3\mathbb{Z},$$

since $k + j \in \mathbb{Z}$. Thus, $m + n \in 3\mathbb{Z}$. ■

Proof know-how. The statement "the set S is closed under addition" is synonymous with the implication "if $m, n \in S$, then $m + n \in S$." Thus, to show that S is closed under addition:

(1) Assume $m, n \in S$.

(2) Show that $m + n \in S$.

Here, addition can be replaced by another operation that S permits.

Before proceeding, we give two more remarks about the proof of Theorem 2.14.

- In the second sentence, you should not write, "Then $m = 3k$ and $n = 3k$ where $k \in \mathbb{Z}$." That would imply that m and n are equal (since they're both equal to $3k$), which is not necessarily true.

- In the third sentence of the proof, you should not write, "We have $m+n = 3(k+j) = 3k + 3j \in 3\mathbb{Z}$." (The explanation is left for you as an exercise at the end of the chapter.)

Example 2.15. Similarly to $3\mathbb{Z}$, we can define the set $2\mathbb{Z} = \{n \in \mathbb{Z} \mid n = 2k \text{ where } k \in \mathbb{Z}\}$, which contains all integer multiples of 2 (or all even integers). Likewise, we can define $6\mathbb{Z} = \{n \in \mathbb{Z} \mid n = 6k \text{ where } k \in \mathbb{Z}\}$ that contains all integer multiples of 6.

More generally, fix an integer m. Then $m\mathbb{Z} = \{n \in \mathbb{Z} \mid n = mk \text{ where } k \in \mathbb{Z}\}$ is the set of all integer multiples of m.

Example 2.16. Let's look at a couple of extreme cases for the set $m\mathbb{Z}$:

- When $m = 1$, we have $m\mathbb{Z} = \mathbb{Z}$.

- When $m = 0$, we have $0\mathbb{Z} = \{0\}$, i.e., the set with just one element 0.

2.4 Showing set equality

Example 2.17. Consider the following subset of $\mathbb{Z}$:

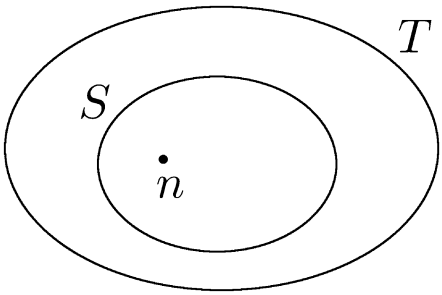
$$C = \{n \in \mathbb{Z} \mid n \in 2\mathbb{Z} \text{ and } n \in 3\mathbb{Z}\}.$$

Thus, set C contains all integers that are multiples of both 2 and 3. For example, $18 \in C$, because $18 = 2 \cdot 9$ and $18 = 3 \cdot 6$. However, $20 \notin C$. Although 20 is a multiple of 2, it is *not* a multiple of 3. Other elements of set C include 6, 12, 24, 30, ..., as well as their negatives (and also 0). We conjecture that $C = 6\mathbb{Z}$.

Remark. In mathematics, a *conjecture* is a statement that is not yet proven but that we suspect is true based on some initial evidence. When used as a verb, "to conjecture" means "to make a conjecture."

In Example 2.17 above, we claimed that $C = 6\mathbb{Z}$. But how can we *prove* it? More generally, how can we prove that two sets are equal, i.e., that they contain the same elements?

Proof know-how. To show that sets S and T are equal, we must show two things: (1) $S \subseteq T$, i.e., S is a subset of T, and (2) $T \subseteq S$, i.e., T is a subset of S. To show that $S \subseteq T$, we recall that $S \subseteq T$ means that every element of S is an element of T. Written as an implication, this becomes: If $n \in S$, then $n \in T$. Thus, to prove $S \subseteq T$, we assume that $n \in S$ and then show that $n \in T$. Below is a visual depiction of $S \subseteq T$.



Proving that $T \subseteq S$ is done similarly, with the roles of S and T swapped.

Below, we will prove that $6\mathbb{Z} \subseteq C$. The proof of $C \subseteq 6\mathbb{Z}$ is left as an exercise at the end of this chapter. Thus, we conclude that $C = 6\mathbb{Z}$.

Now, proving $6\mathbb{Z} \subseteq C$ is the same as proving: If $n \in 6\mathbb{Z}$, then $n \in C$. Thus, the first sentence of the proof should be: Assume $n \in 6\mathbb{Z}$. Our goal is to show that $n \in C$, so the last sentence should be: Therefore, $n \in C$. To proceed with the rest of the proof, let's look at an example. Suppose $n = 30$. We know that $30 \in 6\mathbb{Z}$, because 30 is a multiple of 6; i.e., $30 = 6 \cdot 5$. But a multiple of 6 is also a multiple of 2, because 6 itself is a multiple of 2. More precisely, we have $30 = 6 \cdot 5 = (2 \cdot 3) \cdot 5 = 2 \cdot (3 \cdot 5)$. Similarly, $30 = 6 \cdot 5 = (3 \cdot 2) \cdot 5 = 3 \cdot (2 \cdot 5)$, so that 30 is a multiple of 3. Therefore, $30 \in C$, because it is a multiple of both 2 and 3.

In this example, there is nothing special about $n = 30$. The same argument can be made with $30 = 6 \cdot 5$ replaced by any multiple of 6, i.e., $n = 6 \cdot k$ (where k is an integer). Now, we're ready for the proof.

Theorem 2.18. *Let $C = \{n \in \mathbb{Z} \mid n \in 2\mathbb{Z}$ and $n \in 3\mathbb{Z}\}$. Then $6\mathbb{Z} \subseteq C$.*

PROOF. Assume $n \in 6\mathbb{Z}$. We must show that $n \in C$. Then $n = 6 \cdot k$ for some integer k. We have $n = 6 \cdot k = (2 \cdot 3) \cdot k = 2 \cdot (3 \cdot k)$, so that $n \in 2\mathbb{Z}$. Similarly, $n = 6 \cdot k = (3 \cdot 2) \cdot k = 3 \cdot (2 \cdot k)$, so that $n \in 3\mathbb{Z}$. Thus, $n \in 2\mathbb{Z}$ and $n \in 3\mathbb{Z}$. Therefore, $n \in C$. ■

Proof know-how. The second sentence of the proof says, "We must show that $n \in C$." (We could also write, "We will show that $n \in C$.") Although not required, it is a helpful way to state the goal of the proof and indicate where the proof is headed. Be careful, though: We must *not* write "Therefore, $n \in C$" as the second sentence, since we haven't shown that yet.

Exercises

1. Find all subsets of the set {Sarah, Elizabeth, Anita, Ryota}. (Don't forget the empty set!)

2. List the elements of the following subsets of $\mathbb{Z}$:

   (a) $\{n \in \mathbb{Z} \mid n = \frac{1}{n}\}$.

   (b) $\{n \in \mathbb{Z} \mid n^2 = n\}$.

   (c) $\{n \in \mathbb{Z} \mid n < 100$ and $\sqrt{n}$ is an integer$\}$.

3. List the elements of the following subsets of $\mathbb{Z}$. If a subset contains infinitely many elements, then list *some* of its elements and use the ellipsis (...) appropriately.
   (a) $\{n \in \mathbb{Z} \mid n^3 \text{ is odd}\}$.
   (b) $\{n \in \mathbb{Z} \mid n = k^2 \text{ where } k \in \mathbb{Z}\}$.
   (c) $\{n \in \mathbb{Z} \mid 2 < 3n + 1 < 20\}$.

4. Describe each subset of $\mathbb{Z}$ using the notation $\{n \in \mathbb{Z} \mid \text{some property satisfied by } n\}$.
   Note: For parts (a) and (b), you may assume that the patterns you expect continue to hold.
   (a) $\{0, 5, 10, 15, 20, 25, \ldots\}$.
   (b) $\{\ldots, -19, -15, -11, -7, -3, 1, 5, 9, 13, 17, 21, \ldots\}$.
   (c) $\{3, 7, 11, 19, 23, 31, 43, 47, 59, 67, 71, 79, 83\}$.

5. Describe each subset of $\mathbb{Z}$ using the notation $\{n \in \mathbb{Z} \mid \text{some property satisfied by } n\}$.
   (a) The set of all positive multiples of 7.
   (b) The set of all negative multiples of 4.
   (c) The set of all 2-digit positive integers that end with a 3.

6. Suppose $S = \{n \in \mathbb{Z} \mid \text{some property satisfied by } n\}$. Find "some property satisfied by n" so that:
   (a) $S = \mathbb{Z}$.
   (b) $S = \emptyset$.

7. Let S be a subset of $\mathbb{Z}$. Define precisely what it means for S to be *closed* under addition.

8. Consider the set $\mathbb{Q} = \left\{\frac{m}{n} \mid m, n \in \mathbb{Z},\ n \neq 0\right\}$.
   (a) Write down a few elements of this set.
   (b) Choose two elements of $\mathbb{Q}$ and add them together. Is the sum still in $\mathbb{Q}$?
   (c) Repeat part (b) with a few more pairs of elements in $\mathbb{Q}$.

9. **Prove:** The set $\mathbb{Q} = \left\{\frac{m}{n} \mid m, n \in \mathbb{Z},\ n \neq 0\right\}$ is closed under addition.
   Note: First work through Exercise #8 above. Also, be careful: $\frac{3}{7} + \frac{2}{5} \neq \frac{3+2}{7+5}$.

10. Fix an integer m. Prove that $m\mathbb{Z}$ is closed under addition.

11. Elizabeth and Anita wrote the following proof of Theorem 2.14:

    Proof: Assume $m, n \in 3\mathbb{Z}$.
    Then $m = 3k$ and $n = 3j$ where $k, j \in \mathbb{Z}$.
    Thus, $m + n = 3(k + j) = 3k + 3j \in 3\mathbb{Z}$, since $k + j \in \mathbb{Z}$.
    Hence, $m + n \in 3\mathbb{Z}$.

    There is a (subtle) logical error in their proof. Find it and fix it.

12. **Prove:** The set $3\mathbb{Z}$ is closed under multiplication.

13. Consider the statement: If $r \in \mathbb{Z}$ and $a \in 5\mathbb{Z}$, then $r \cdot a \in 5\mathbb{Z}$.
    (a) Create an example that illustrates this statement.
    (b) Prove the statement.

14. (a) **Prove:** $12\mathbb{Z} \subseteq 4\mathbb{Z}$, i.e., $12\mathbb{Z}$ is a subset of $4\mathbb{Z}$. (This exercise is referenced in Exercise #20a.)
    (b) Give a counterexample to show that $4\mathbb{Z} \subseteq 12\mathbb{Z}$ is false.

15. Consider the statement: If $R \subseteq S$ and $S \subseteq T$, then $R \subseteq T$. (Here, R, S, and T are sets.)
    (a) Create an example that illustrates this statement.
    (b) Prove the statement.

16. Let S and T be sets. For each statement, if it's true, prove it; if it's false, give a counterexample.
    (a) If $n \in S$ and $S \subseteq T$, then $n \in T$.
    (b) If $n \in T$ and $S \subseteq T$, then $n \in S$.

17. **Prove:** Let $C = \{n \in \mathbb{Z} \mid n \in 2\mathbb{Z} \text{ and } n \in 3\mathbb{Z}\}$. Then $C \subseteq 6\mathbb{Z}$.
    Note: Combined with $6\mathbb{Z} \subseteq C$ (Theorem 2.18), we conclude that $C = 6\mathbb{Z}$.
    Hint: Chapter 1, Exercise #2 may be useful for this proof.

18. Consider the statement: Fix an integer m. Then $m\mathbb{Z} = (-m)\mathbb{Z}$.
    (a) Create an example that illustrates this statement.
    (b) Prove the statement.

19. Let S and T be sets. Define their *intersection*, denoted $S \cap T$, to be the set containing elements that are in both S and T. For instance, we saw in Example 2.17 that $2\mathbb{Z} \cap 3\mathbb{Z} = 6\mathbb{Z}$.
    (a) Find a value of m such that $4\mathbb{Z} \cap 5\mathbb{Z} = m\mathbb{Z}$.
    (b) Find a value of m such that $7\mathbb{Z} \cap 10\mathbb{Z} = m\mathbb{Z}$.
    (c) Find a value of m such that $10\mathbb{Z} \cap 15\mathbb{Z} = m\mathbb{Z}$.
    (d) Any conjectures?

20. (a) In Exercise #14, you showed that $12\mathbb{Z} \subseteq 4\mathbb{Z}$. Compute $12\mathbb{Z} \cap 4\mathbb{Z}$.
    (b) Find a pair of sets S and T such that $S \subseteq T$. Then compute $S \cap T$.
    (c) Repeat part (b) with another pair of sets S and T.
    (d) What conjecture do you have?

21. **Prove:** Let S and T be sets. If $S \subseteq T$, then $S \cap T = S$.

22. Find sets R, S, and T with the following properties:
    - There is no element that is in all three sets.
    - $R \cap S$, $R \cap T$, and $S \cap T$ are all *not* the empty set.

3

Divisors

In the first two chapters, we examined various properties of the set of integers $\mathbb{Z}$. For instance, given an integer n, we considered its *parity*, i.e., whether it is odd or even. We also saw that $\mathbb{Z}$ is *closed* under addition and multiplication. We then explored the various subsets of $\mathbb{Z}$.

In this chapter, we will study the notion of *divisors*, which indicates how a pair of integers relate to one another. For example, we say 4 is a divisor of 24, which is synonymous with saying 24 is a multiple of 4. Divisors play an important role in various topics in this textbook, including (but certainly not limited to) modular arithmetic, order of a group element, and Lagrange's theorem.

3.1 Divisor

The following mean the same thing:

- 24 is a multiple of 4.
- 4 is a divisor of 24.

Digging a bit deeper, we know that 4 is a divisor of 24 (or synonymously, 24 is a multiple of 4), because $24 = 4 \cdot 6$. Moreover, we write $4 \mid 24$ as a shorthand for "4 is a divisor of 24."

Example 3.1. We know that 7 is a divisor of 91, because $91 = 7 \cdot 13$. As a shorthand, we write $7 \mid 91$. However, 7 is *not* a divisor of 32 (shorthand: $7 \nmid 32$), since there is no integer k such that $32 = 7 \cdot k$.

Generalizing from these examples, we obtain the following.

Definition 3.2. Let $d, n \in \mathbb{Z}$. We say that d is a *divisor* of n when $n = d \cdot k$ for some integer k. We write $d \mid n$ as a shorthand for "d is a divisor of n."

Remark. As discussed above, the notation $4 \mid 24$ is a *statement* which says "4 is a divisor of 24." This should not be confused with $\frac{24}{4}$, which is a (rational) *number* that is equal to 6.

Example 3.3. Here are some divisor relations involving the integer 0:

- 17 is a divisor of 0 (i.e., $17 \mid 0$), because $0 = 17 \cdot 0$.
- 0 is *not* a divisor of 17 (i.e., $0 \nmid 17$), because there is no integer k such that $17 = 0 \cdot k$.
- 0 is a divisor of 0 (i.e., $0 \mid 0$), because $0 = 0 \cdot 23$. Here, 23 can be replaced by *any* integer. Observe that $0 \mid 0$ is a true statement, while the expression $\frac{0}{0}$ is undefined.

Our next goal is to prove the statement: *If $a \mid b$ and $b \mid c$, then $a \mid c$.* (Here, a, b, $c \in \mathbb{Z}$.) As usual, we begin by writing down the first and last sentences of the proof. Here is the first sentence, where we assume the hypothesis: Assume $a \mid b$ and $b \mid c$. Our last sentence is the conclusion: Thus, $a \mid c$. For the steps in between, let's consider an example where $a = 4$, $b = 12$, and $c = 60$. Then $4 \mid 12$, because $12 = 4 \cdot 3$. Also $12 \mid 60$, because $60 = 12 \cdot 5$. Combining these two, we obtain $60 = 12 \cdot 5 = (4 \cdot 3) \cdot 5 = 4 \cdot (3 \cdot 5)$, which implies that $4 \mid 60$. The proof is obtained by replacing 4, 12, and 60 with a, b, and c.

Theorem 3.4. *Let $a, b, c \in \mathbb{Z}$. If $a \mid b$ and $b \mid c$, then $a \mid c$.*

PROOF. Assume $a \mid b$ and $b \mid c$. We must show that $a \mid c$. First, $a \mid b$ implies $b = a \cdot k$ for some integer k. Likewise, $b \mid c$ means $c = b \cdot j$ for some $j \in \mathbb{Z}$. Then,

$$c = b \cdot j = (a \cdot k) \cdot j = a \cdot (k \cdot j),$$

so that $c = a \cdot (k \cdot j)$, where $k \cdot j$ is an integer. Thus, $a \mid c$ as desired. ■

Proof know-how. The last sentence of the above proof includes the phrase "as desired." Although not required, this is a common way of ending the proof by emphasizing that what was intended to be proved has been proved.

3.2 GCD theorem

The following example introduces the notion of a common divisor and the greatest common divisor.

Example 3.5. Let $a = 12$ and $b = 18$. Then $d = 3$ is a *common divisor* of a and b, because $d \mid a$ and $d \mid b$ (i.e., d is a divisor of both a and b). The list of all common divisors of $a = 12$ and $b = 18$ includes ± 1, ± 2, ± 3, and ± 6. Of these common divisors, 6 is the greatest (or the largest). Therefore, the *greatest common divisor* of a and b is 6, written $\gcd(a, b) = 6$.

Definition 3.6. Let $a, b \in \mathbb{Z}$. An integer d is the *greatest common divisor* of a and b, written $d = \gcd(a, b)$, if it satisfies the following properties:

- $d > 0$ (i.e., d must be positive).
- $d \mid a$ and $d \mid b$ (i.e., d is a divisor of both a and b).
- If $e \mid a$ and $e \mid b$, then $d \geq e$ (see remark below).

Remark. The third property in the above definition means, "If e is a common divisor of a and b, then d is at least as large as e." This ensures that d is the *greatest* common divisor of a and b. Some textbooks, especially those in number theory, use "if $e \mid a$ and $e \mid b$, then $e \mid d$" as its third property. We will use the version in Definition 3.6, since it emphasizes how d is the *greatest* among the common divisors of a and b.

Example 3.7. Let $a = 10$ and $b = 27$. Then $\gcd(10, 27) = 1$. We say that 10 and 27 are *relatively prime*, since their gcd is 1. Neither 10 nor 27 is a prime number, but "relatively prime" means that they do not share any common divisor except for ± 1.

On the surface, the next example seems unrelated to the notion of divisors. But it plays an important role in the theorems that we will prove in the rest of this chapter and beyond.

Example 3.8. Again, let $a = 10$ and $b = 27$. Consider the equation $10x + 27y = 1$. This equation has an *integer solution*; i.e., we can find integers x and y that satisfy the equation. With $x = -8$ and $y = 3$, we have $10 \cdot (-8) + 27 \cdot 3 = -80 + 81 = 1$. Note that $(x, y) = (-8, 3)$ is *not* the only integer solution to this equation. For instance, we have $10 \cdot 19 + 27 \cdot (-7) = 1$, so that $(x, y) = (19, -7)$ is another integer solution.

Examples 3.7 and 3.8 suggest the following theorem. Its proof is given much later in the textbook (in Chapter 35), using future concepts such as *rings* and *principal ideals*. If you'd rather not wait that long, however, there's also a more accessible proof in Appendix A.

Theorem 3.9 (GCD theorem)**.** *Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) = 1$, then there exist integers x and y such that $ax + by = 1$.*

The utility of the GCD theorem is that it allows us to convert a conceptual relationship between a and b, namely that a and b are relatively prime, to a concrete equation, namely $ax + by = 1$. As we will see in the next section, the equation $ax + by = 1$ is much easier to use in a proof involving divisors.

Example 3.10. Let $a = 5$ and $b = 8$. Then $\gcd(5, 8) = 1$, and thus by the GCD theorem, we expect to find an integer solution to the equation $5x + 8y = 1$. Indeed, we have $5 \cdot (-3) + 8 \cdot 2 = 1$, so that $(x, y) = (-3, 2)$ is an integer solution. We also have $5 \cdot 13 + 8 \cdot (-8) = 1$, so that $(x, y) = (13, -8)$ is another integer solution.

Example 3.11. Let $a = 10$ and $b = 34$. Then $\gcd(10, 34) = 2$. The GCD theorem does not make any conclusion when $\gcd(a, b) \neq 1$. But we can see that the equation $10x + 34y = 1$ does *not* have an integer solution. If x and y were integers, then $10x + 34y = 2 \cdot (5x + 17y)$ is a multiple of 2, and thus it cannot be equal to 1. You will work on a generalization of this example as an exercise at the end of the chapter.

3.3 Proofs involving the GCD theorem

In this section, we will prove a pair of theorems whose hypotheses include the condition $\gcd(a, b) = 1$. Recall that the GCD theorem allows us to convert this to a more usable form, i.e., the equation $ax + by = 1$.

Example 3.12. Consider the statement: *If $a \mid bc$, then $a \mid c$.* (Here, $a, b, c \in \mathbb{Z}$.) This statement is false. As a counterexample, suppose $a = 4$, $b = 6$, $c = 2$, so that $bc = 12$. We have $4 \mid 12$ (i.e., $a \mid bc$) but 4 is *not* a divisor of 2 (i.e., $a \nmid c$).

Example 3.13. We can salvage the statement in Example 3.12 by adding the condition $\gcd(a, b) = 1$. We thus obtain: *If $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.* We will soon see that this is a true statement. A possible example is $a = 4$, $b = 5$, $c = 12$, so that $bc = 60$. We have $4 \mid 60$ and $\gcd(4, 5) = 1$; and $4 \mid 12$, as desired.

Let's brainstorm how to prove the statement: *If $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.* Our hypotheses are $a \mid bc$ and $\gcd(a, b) = 1$, which translate to $bc = ak$ and $ax + by = 1$, respectively. Our goal is to show that $c = a \cdot$ (some integer), from which we can conclude $a \mid c$. We might be tempted to divide both sides of $bc = ak$ by b to obtain $c = a \cdot \frac{k}{b}$. However, $\frac{k}{b}$ may not be an integer. Also, the equation $ax + by = 1$ (which was derived from $\gcd(a, b) = 1$) should play a role in the proof, since we saw in Example 3.12 that the statement is false without the condition $\gcd(a, b) = 1$.

Our goal $c = a \cdot$ (some integer) involves the integers a and c. The equation $bc = ak$ already has a and c in it. But $ax + by = 1$ only involves a, so we must somehow introduce c into it. We accomplish this by multiplying both sides of $ax + by = 1$ by c, which yields $acx + (bc)y = c$. Upon substituting $bc = ak$, we obtain $acx + (ak)y = c$, from which we can conclude that c is a multiple of a.

Theorem 3.14. *Let $a, b, c \in \mathbb{Z}$. If $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.*

PROOF. Assume $a \mid bc$ and $\gcd(a, b) = 1$. Then $bc = ak$ for some integer k, and there exist $x, y \in \mathbb{Z}$ with $ax + by = 1$. Multiplying both sides of $ax + by = 1$ by c, we obtain $acx + (bc)y = c$. Then substituting $bc = ak$ yields $acx + (ak)y = c$, and thus $a \cdot (cx + ky) = c$ where $cx + ky$ is an integer. Therefore, $a \mid c$. ∎

Remark. The key to the proof above is to multiply both sides of $ax + by = 1$ by c. It may seem that the step was pulled out of thin air, and that's what makes proof writing a creative, challenging, and sometimes frustrating endeavor. Coming up with such insights is not an easy task. It requires lots of practice writing proofs, perseverance, and even luck. But you're not alone. Even mathematicians with decades of experience can and often do struggle deriving a key step to a proof.

Example 3.15. Consider the statement: *If $a \mid c$ and $b \mid c$, then $ab \mid c$.* (Here, a, b, $c \in \mathbb{Z}$.) This statement is false. As a counterexample, suppose $a = 4$, $b = 6$, $c = 12$, so that $ab = 24$. We have $4 \mid 12$ and $6 \mid 12$ (i.e., $a \mid c$ and $b \mid c$), but 24 is not a divisor of 12 (i.e., $ab \nmid c$).

Example 3.16. We can salvage the statement in Example 3.15 by adding the condition $\gcd(a, b) = 1$. We thus obtain: *If $a \mid c$, $b \mid c$, and $\gcd(a, b) = 1$, then $ab \mid c$.* This is now a true statement, which we will verify shortly. A possible example is $a = 4$, $b = 5$, $c = 60$, so that $ab = 20$. Note that $4 \mid 60$, $5 \mid 60$, and $\gcd(4, 5) = 1$; and $20 \mid 60$, as desired.

Here is the theorem that we studied in Example 3.16. Notice how its proof uses the same insight as in the proof of Theorem 3.14, namely, multiplying both sides of $ax + by = 1$ by c. (Experience helps!)

Theorem 3.17. *Let $a, b, c \in \mathbb{Z}$. If $a \mid c$, $b \mid c$, and $\gcd(a, b) = 1$, then $ab \mid c$.*

PROOF. Assume $a \mid c$, $b \mid c$, and $\gcd(a, b) = 1$. Then $c = ak$ and $c = bj$ where k and j are integers. Moreover, there exist $x, y \in \mathbb{Z}$ with $ax + by = 1$. Multiplying both sides of $ax + by = 1$ by c yields $acx + bcy = c$. Substitute $c = bj$ and $c = ak$ to obtain $c = a(bj)x + b(ak)y = ab \cdot (jx + ky)$. Thus, ab is a divisor of c. ■

To conclude this chapter, we will prove the *converse* of the GCD theorem, obtained by swapping the if-part and the then-part. Thus, the statement we will prove is: *If there exist $x, y \in \mathbb{Z}$ with $ax + by = 1$, then* $\gcd(a, b) = 1$. (Here, $a, b \in \mathbb{Z}$.) One strategy is to prove the contrapositive of this statement, and you will do this as an exercise at the end of the chapter.

But here, we will take a more direct approach. Hence, we begin by assuming that $ax + by = 1$ where x and y are integers. Our goal is to show that $\gcd(a, b) = 1$. We proceed by letting $d = \gcd(a, b)$, and so we must show that $d = 1$. This seemingly simple step of assigning a name/variable d to $\gcd(a, b)$ can facilitate our thinking. Here is a brief outline that summarizes the key steps of the proof:

- Since $d = \gcd(a, b)$ is a common divisor of a and b, we have $d \mid a$ and $d \mid b$.
- Then we show that d is a divisor of $ax + by$.
- But $ax + by = 1$, and thus d is a divisor of 1. Then, $d = 1$ or $d = -1$ (i.e., the only divisors of 1).
- By the definition of gcd, d is positive. Hence, d must be 1.

Theorem 3.18 (Converse of the GCD theorem)**.** *Let $a, b \in \mathbb{Z}$. If there exist $x, y \in \mathbb{Z}$ with $ax + by = 1$, then* $\gcd(a, b) = 1$.

PROOF. Assume there exist $x, y \in \mathbb{Z}$ with $ax + by = 1$. Let $d = \gcd(a, b)$, noting that $d > 0$ by the definition of the gcd. Then $d \mid a$ and $d \mid b$, so that $a = dk$ and $b = dj$ for some integers k and j. Substituting these into $ax + by = 1$, we get $(dk)x + (dj)y = 1$, and thus $d \cdot (kx + jy) = 1$. Hence, d is a positive divisor of 1, which implies that $d = 1$. Therefore, $\gcd(a, b) = 1$. ■

Example 3.19. Let $a = 31{,}415{,}926$ and $b = 31{,}415{,}927$. Then $a \cdot (-1) + b \cdot 1 = 1$, so that $ax + by = 1$ has an integer solution $(x, y) = (-1, 1)$. By Theorem 3.18, we conclude that $\gcd(a, b) = 1$.

Exercises

1. Determine if each of these is true or false. Explain your reasoning.

   (a) $17 \mid 17$.

   (b) $1 \mid 17$.

   (c) $17 \mid 1$.

2. Let n be an integer.
   (a) Explain why $n \mid n$.
   (b) Explain why $1 \mid n$.
3. **Prove:** Let m and n be positive integers. If $m \mid n$ and $n \mid m$, then $m = n$.
4. Explain why $\gcd(0, 0)$ does not exist.
5. Find each of the following:
   (a) $\gcd(4, 12)$.
   (b) $\gcd(20, 80)$.
   (c) $\gcd(17, 85)$.
   (d) $\gcd(a, ak)$ where $a, k \in \mathbb{Z}$. (Assume that $a \neq 0$.)
6. Find each of the following. Explain your reasoning.
   (a) $\gcd(0, 17)$.
   (b) $\gcd(0, 314)$.
   (c) $\gcd(0, n)$ where n is a positive integer.
7. Compute and compare each pair of GCDs:
   (a) $\gcd(12, 30)$ and $\gcd(12, 18)$.
   (b) $\gcd(156, 228)$ and $\gcd(156, 72)$.
   (c) $\gcd(35, 21)$ and $\gcd(14, 21)$.
   (d) $\gcd(182, 52)$ and $\gcd(130, 52)$.

   Based on these examples, what conjectures do you have?
8. Computations in Exercise #7 illustrate how
   $$\gcd(a, b) = \gcd(a, b - a) \text{ and } \gcd(a, b) = \gcd(a - b, b).$$
   (a) Appendix A explains *why* these relationships are true. For now, use these relationships to come up with a procedure for finding $\gcd(a, b)$.
   (b) Use your procedure in part (a) to find $\gcd(391, 582)$ and to find $\gcd(873,\ 3{,}642)$.
9. Consider the statement:

   Let $a, b, c \in \mathbb{Z}$. If $a \mid c$ and $b \mid c$, then $ab \mid c$.

   Show that this is false by exhibiting a counterexample. Be sure to explain why your counterexample invalidates the statement.
10. Consider the statement: Let $m, n, a, b \in \mathbb{Z}$. If $m \mid a$ and $n \mid b$, then $mn \mid ab$.
    (a) Create an example to illustrate this statement.
    (b) Prove the statement.
11. **Prove:** Let $d, m, n \in \mathbb{Z}$. If $d \mid m$ and $d \mid n$, then $d \mid (m + n)$.

12. Let $m, n \in \mathbb{Z}$ and consider the statement: *If $n \mid m$, then $m\mathbb{Z} \subseteq n\mathbb{Z}$.*

    (a) Create an example to illustrate this statement.

    (b) Prove the statement.

    (This exercise and Exercise #13 below are referenced in Chapter 31.)

13. **Prove:** Let $m, n \in \mathbb{Z}$. If $m\mathbb{Z} \subseteq n\mathbb{Z}$, then $n \mid m$.

    Note: This is the converse of the statement in Exercise #12.

14. For each pair of integers a and b, determine whether or not $ax + by = 1$ has an integer solution. If it does, then find at least *three* integer solutions (x, y). If it doesn't, then explain why not.

    (a) $a = 7, b = 10$.

    (b) $a = 8, b = 10$.

    (c) $a = 15, b = 21$.

    (d) $a = 15, b = 16$.

15. Find *all* integer solutions to $7x + 10y = 1$. How do you know that you've found them all?

16. **Prove:** Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) \neq 1$, then $ax + by = 1$ does *not* have an integer solution.

    Note: You must prove this statement as it's stated and not work with its contrapositive, which is Theorem 3.18. Example 3.11 should help.

17. Explain why $\gcd(3n + 2,\ 5n + 3) = 1$ for any integer n.

18. Explain why $\gcd(n,\ n + 1) = 1$ for any integer n.

19. **Prove:** Let $a, b, c \in \mathbb{Z}$. If $\gcd(a, b) = 1$ and $c \mid a$, then $\gcd(c, b) = 1$.

20. Let a and b be consecutive odd integers (e.g., 43 and 45). Show that $\gcd(a, b) = 1$.

    Hint: Let $d = \gcd(a, b)$ so that $d \mid a$ and $d \mid b$. Then show that $d \mid 2$.

21. **Prove:** 3, 5, 7 is the only "prime triple," i.e., three consecutive odd integers that are all prime.

Unit II: Examples of Groups

The next few chapters are devoted to examples of groups:

- In Chapter 4, we will study *integers modulo n*, a number system in which we add and multiply as if we're using a clock. For example, if it's currently 9:00 AM and 5 hours pass, then it will be 2:00 PM.
- We will explore *symmetries* in Chapter 5. These are motions of a square (or other regular polygons) that, when applied to the square, place the square in the same space that it originally occupied.
- In Chapter 6, we will investigate *permutations*, which are functions on the set $\{1, 2, 3, \ldots, n\}$ that "shuffle" these numbers. Historically, permutations paved the way for the development of group theory.
- Finally, in Chapter 7, we will study *matrices*, which are rectangular arrays of numbers that play an important role in many areas of mathematics (including abstract algebra, of course).

Our main goal is to identify features that are common to these seemingly different sets of objects. Indeed, these concrete examples introduce the fundamental notion of "group properties," which include (1) closure, (2) associative law, (3) identity, and (4) inverses. Other essential ideas such as the *order* of a group element and *subgroups* are also foreshadowed through these examples. By the time the concept of a *group* is formally defined in Chapter 8, we hope it will feel familiar to you!

Here is a taste of what you'll be able to accomplish in this unit:

- Very quickly determine which elements in $\mathbb{Z}_{35}$ have multiplicative inverses. For example, 8 has a multiplicative inverse in $\mathbb{Z}_{35}$, i.e., an element x such that $8 \cdot x = 1 \pmod{35}$.
- Very quickly determine whether or not a 2×2 matrix with entries in $\mathbb{Z}_{10}$ has a multiplicative inverse.
- Understand why $(\alpha \cdot \beta)^{-1} = \beta^{-1} \cdot \alpha^{-1}$ by drawing an analogy to the act of putting on and taking off your socks and shoes. (Here, α and β could be a pair of symmetries, permutations, or matrices.)

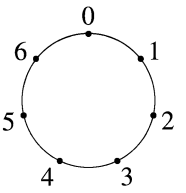
4

Modular Arithmetic

If it's currently 7:00 AM, what time will it be in 2,000 hours? Since there are 24 hours in a day and $2{,}000 = 24 \cdot 83 + 8$, we see that 2,000 hours equal 83 days and 8 hours. Thus, it will be 3:00 PM. This scenario is an illustration of *modular arithmetic*, which is the focus of this chapter. We will work with number systems such as $\mathbb{Z}_7$, which, along with $\mathbb{Z}$, are perhaps the most important examples in this textbook.

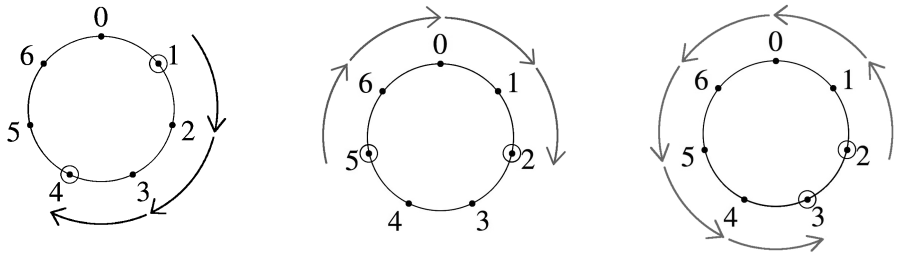
4.1 Number system $\mathbb{Z}_7$

Consider the set $\mathbb{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$. We will be using this and related number systems throughout the book. Unlike the set of integers $\mathbb{Z}$, our set $\mathbb{Z}_7$ contains finitely many elements (7 of them, in fact). To add, subtract, and multiply in $\mathbb{Z}_7$, we use the following picture called the $\mathbb{Z}_7$ clock:



Example 4.1. These sums are computed in $\mathbb{Z}_7$:

- To compute $1 + 3$, start at 1 on the $\mathbb{Z}_7$ clock and move 3 units clockwise (see the left figure below). We land on 4, so that $1 + 3 = 4$ in $\mathbb{Z}_7$.
- To compute $5 + 4$, start at 5 on the $\mathbb{Z}_7$ clock and move 4 units clockwise (see the center figure below). We land on 2, so that $5 + 4 = 2$ in $\mathbb{Z}_7$.
- To compute $2 - 6$, start at 2 on the $\mathbb{Z}_7$ clock and move 6 units *counterclockwise* (see the right figure below). We land on 3, so that $2 - 6 = 3$ in $\mathbb{Z}_7$.



Example 4.2. To find -3 in $\mathbb{Z}_7$, we view -3 as $0-3$. Thus, we start at 0 on the $\mathbb{Z}_7$ clock and move 3 units *counterclockwise*. We land on 4, so that $-3 = 4$ in $\mathbb{Z}_7$.

Example 4.3. We can also multiply in $\mathbb{Z}_7$. For instance, $3 \cdot 5 = 15$. To find what 15 equals in $\mathbb{Z}_7$, we can view 15 as $0 + 15$. To compute this sum, start at 0 on the $\mathbb{Z}_7$ clock and move 15 units clockwise. You should land on 1. Try it! Thus, $3 \cdot 5 = 15 = 1$ in $\mathbb{Z}_7$.

Example 4.4. To find 3,258 in $\mathbb{Z}_7$, we start at 0 on the $\mathbb{Z}_7$ clock and move 3,258 units clockwise. Every movement by 7 units brings us back to 0. Since $3{,}258 = 7 \cdot 465 + 3$, moving 3,258 units yields 465 full revolutions around the $\mathbb{Z}_7$ clock, plus 3 more units. Thus, $3{,}258 = 3$ in $\mathbb{Z}_7$.

Example 4.5. To find $-3{,}258$ in $\mathbb{Z}_7$, we note that $3{,}258 = 3$ (Example 4.4) and that $-3 = 4$ (Example 4.2). Combining these, we obtain $-3{,}258 = -3 = 4$ in $\mathbb{Z}_7$.

When we find $3{,}258 = 3$ in $\mathbb{Z}_7$, we say that 3,258 has been *simplified* or *reduced* to 3 in $\mathbb{Z}_7$. Here, recall that $3{,}258 = 7 \cdot 465 + 3$, so that 3 is the remainder when dividing 3,258 by 7.

Example 4.6. We have $10{,}000 = 7 \cdot 1{,}428 + 4$, so that $10{,}000 = 4$ in $\mathbb{Z}_7$. In other words, 10,000 is reduced to 4 in $\mathbb{Z}_7$. Again, we observe that 4 is the remainder when dividing 10,000 by 7.

Example 4.7. To find 2^{101} in $\mathbb{Z}_7$, we begin by computing smaller powers of 2:

$$2^1 = 2,\ \ 2^2 = 4,\ \ 2^3 = 8 = 1,\ \ 2^4 = 16 = 2,\ \ 2^5 = 32 = 4,\ \ 2^6 = 64 = 1, \ldots .$$

The successive powers of 2 form a pattern: 2, 4, 1, 2, 4, 1, ... In particular, we see that $2^n = 1$ in $\mathbb{Z}_7$ whenever the exponent n is a multiple of 3. Thus, we have $2^{99} = 1$, $2^{100} = 2$, and $2^{101} = 4$ in $\mathbb{Z}_7$.

Remark. Here is a word of caution about notation. As a shorthand for "$-3 = 4$ in $\mathbb{Z}_7$," we might be tempted to write "$-3 = 4 \in \mathbb{Z}_7$." After all, the symbol $\in$ stands for "in." However, this is an incorrect use of $\in$, which is actually a shorthand for "is an element of." When we write "$-3 = 4$ in $\mathbb{Z}_7$" (see Example 4.2), the word "in" refers to the fact that the equality $-3 = 4$ is **taking place in** the number system $\mathbb{Z}_7$. We do *not* mean to say that 4 is an element of the set $\mathbb{Z}_7$, which is what "$-3 = 4 \in \mathbb{Z}_7$" would indicate.

4.2 Equality in $\mathbb{Z}_7$

Example 4.8. Consider the integers $a = 16$ and $b = 30$. We will determine whether or not $a = b$ in $\mathbb{Z}_7$. One way to do this is to simplify (or reduce) each of a and b in

$\mathbb{Z}_7$. As in Example 4.3, we can view 16 as $0 + 16$. Then start at 0 on the $\mathbb{Z}_7$ clock and move 16 units clockwise. You should land on 2. Similar work shows that $30 = 2$ in $\mathbb{Z}_7$. Therefore, $a = b$ in $\mathbb{Z}_7$.

Alternatively, observe that $30 = 16 + 7 \cdot 2$. Adding multiples of 7 does not change the location of a number on the $\mathbb{Z}_7$ clock. Thus, we determined that $a = b$ in $\mathbb{Z}_7$ *without* first simplifying a and b.

Example 4.9. In each pair of a and b below, we determine whether or not $a = b$ in $\mathbb{Z}_7$:

- $a = 24$ and $b = 45$: YES, because 45 is $21 = 7 \cdot 3$ more than 24.
- $a = 3{,}258$ and $b = 3{,}288$: NO, because the difference between a and b is 30, which is *not* a multiple of 7.
- $a = -710$ and $b = -731$: YES, because -731 is $21 = 7 \cdot 3$ less than -710.
- $a = 98{,}765{,}123{,}4\mathbf{06}$ and $b = 98{,}765{,}123{,}4\mathbf{76}$: YES, because b is $70 = 7 \cdot 10$ more than a.

Examples 4.8 and 4.9 suggest the following generalization.

Definition 4.10 (Equality in $\mathbb{Z}_m$)**.** Let $a, b \in \mathbb{Z}$. Then $a = b$ in $\mathbb{Z}_7$ whenever $7 \mid (a-b)$, i.e., 7 is a divisor of $a - b$. More generally, $a = b$ in $\mathbb{Z}_m$ whenever $m \mid (a - b)$.

Example 4.11. Consider the statement: *Let $m \mid n$. If $a = b$ in $\mathbb{Z}_n$, then $a = b$ in $\mathbb{Z}_m$.* This statement is true, as we'll prove soon. For an example, suppose $m = 4$ and $n = 24$ so that $m \mid n$. Let $a = 59$ and $b = 11$, whence $a = b$ in $\mathbb{Z}_{24}$ (both reduce to 11 in $\mathbb{Z}_{24}$); and we have $a = b$ in $\mathbb{Z}_4$ (both reduce to 3 in $\mathbb{Z}_4$), as desired.

We now brainstorm how to prove the statement: *Let $m \mid n$. If $a = b$ in $\mathbb{Z}_n$, then $a = b$ in $\mathbb{Z}_m$.* We start the proof by assuming the hypotheses, namely: Assume $m \mid n$ and $a = b$ in $\mathbb{Z}_n$. Our goal, or the last sentence of the proof, is: Thus, $a = b$ in $\mathbb{Z}_m$. For the intermediate steps, we revisit Example 4.11 for insights. We have $59 = 11$ in $\mathbb{Z}_{24}$, because $59 - 11 = 48 = 24 \cdot 2$; i.e., 24 is a divisor of $59 - 11$. We also have $4 \mid 24$, as $24 = 4 \cdot 6$. Combining these, we see that

$$59 - 11 = 24 \cdot 2 = (4 \cdot 6) \cdot 2 = 4 \cdot (6 \cdot 2),$$

which implies that 4 is a divisor of $59 - 11$. Therefore, $59 = 11$ in $\mathbb{Z}_4$.

Theorem 4.12. *Let $m \mid n$. If $a = b$ in $\mathbb{Z}_n$, then $a = b$ in $\mathbb{Z}_m$.*

PROOF. Assume $m \mid n$ and $a = b$ in $\mathbb{Z}_n$. Then $n \mid (a - b)$ so that $a - b = nk$ for some $k \in \mathbb{Z}$. Moreover, $m \mid n$ implies that $n = mj$ for some $j \in \mathbb{Z}$. Therefore, $a - b = nk = (mj)k = m(jk)$, where jk is an integer. Hence, $m \mid (a - b)$. Thus, $a = b$ in $\mathbb{Z}_m$. ■

Example 4.13. The converse of Theorem 4.12 is false. With $m \mid n$, the condition $a = b$ in $\mathbb{Z}_m$ does *not* imply $a = b$ in $\mathbb{Z}_n$. As a counterexample, let $m = 4$ and $n = 24$ again, and consider $a = 11$ and $b = 3$. Then $a = b$ in $\mathbb{Z}_4$ (both reduce to 3 in $\mathbb{Z}_4$), but $a \neq b$ in $\mathbb{Z}_{24}$ (both are already reduced in $\mathbb{Z}_{24}$).

4.3 Multiplicative inverses

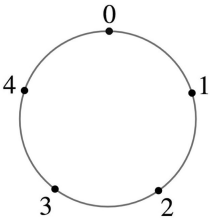
We saw in Example 4.3 that $3 \cdot 5 = 1$ in $\mathbb{Z}_7$. Because of this, we say that 3 and 5 are *multiplicative inverses* of each other in $\mathbb{Z}_7$. (**Note:** This is analogous to $3 \cdot \frac{1}{3} = 1$ and $\frac{1}{5} \cdot 5 = 1$ with the real numbers.) The element $0 \in \mathbb{Z}_7$ does not have a multiplicative inverse, because $0 \cdot x = 0$ for all $x \in \mathbb{Z}_7$. However, all other elements of $\mathbb{Z}_7$ have multiplicative inverses. The *inverse pairs* (i.e., pairs a and b such that $a \cdot b = 1$) are

$$1 \cdot 1 = 1, \quad 2 \cdot 4 = 1, \quad 3 \cdot 5 = 1, \quad 6 \cdot 6 = 1.$$

The multiplicative inverse of 6 is itself. Thus, we say that 6 is a *self-inverse*. Similarly, 1 is a self-inverse.

Definition 4.14 (Multiplicative inverse)**.** Let $a, b \in \mathbb{Z}_m$. We say that a and b are *multiplicative inverses* of each other if $a \cdot b = 1$ in $\mathbb{Z}_m$. Together, a and b form an *inverse pair*. If $a \cdot a = 1$ in $\mathbb{Z}_m$, then the element a is said to be a *self-inverse*.

Example 4.15. Let's switch gears and consider a new number system $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$, where the computation is done on the $\mathbb{Z}_5$ clock:



In $\mathbb{Z}_5$, all elements except 0 have multiplicative inverses. The inverse pairs are $1 \cdot 1 = 1$, $2 \cdot 3 = 1$, and $4 \cdot 4 = 1$ (which means that 1 is a self-inverse, and so is 4).

Example 4.16. Now consider $\mathbb{Z}_{15} = \{0, 1, 2, 3, \ldots, 12, 13, 14\}$ (fifteen elements), where the computation is done on the $\mathbb{Z}_{15}$ clock. Let's find the elements in $\mathbb{Z}_{15}$ that have multiplicative inverses:

- 0 does not have a multiplicative inverse, because $0 \cdot x = 0$ for all $x \in \mathbb{Z}_{15}$.
- We have $1 \cdot 1 = 1$, $4 \cdot 4 = 1$, $11 \cdot 11 = 1$, and $14 \cdot 14 = 1$. Thus, 1, 4, 11, and 14 are self-inverses.
- Other inverse pairs are: $2 \cdot 8 = 1$ and $7 \cdot 13 = 1$.

The remaining elements of $\mathbb{Z}_{15}$ do not have multiplicative inverses. Consider $3 \in \mathbb{Z}_{15}$, for instance. The only multiples of 3 in $\mathbb{Z}_{15}$ are 0, 3, 6, 9, 12, and so there is no element $x \in \mathbb{Z}_{15}$ such that $3 \cdot x = 1$. Therefore, the elements of $\mathbb{Z}_{15}$ with multiplicative inverses are 1, 2, 4, 7, 8, 11, 13, 14.

Example 4.17. Here is one reason why multiplicative inverses are useful. Suppose we want to solve the equation $7 \cdot x = 9$ in $\mathbb{Z}_{15}$. We might divide both sides by 7 to obtain $x = \frac{9}{7}$. However, $\frac{9}{7}$ is not an element in $\mathbb{Z}_{15}$, nor is division a valid operation in $\mathbb{Z}_{15}$. Since there are only 15 elements in $\mathbb{Z}_{15}$, we could substitute each of them for x and see if any of them satisfies the equation.

Alternatively, recall from Example 4.16 that 7 and 13 form an inverse pair; i.e., $7 \cdot 13 = 1$ and $13 \cdot 7 = 1$. We multiply both sides of $7 \cdot x = 9$ by 13 to obtain

$$\begin{aligned} 13 \cdot (7 \cdot x) = 13 \cdot 9 &\implies (13 \cdot 7) \cdot x = 13 \cdot 9 \\ &\implies 1 \cdot x = 13 \cdot 9 \\ &\implies x = 13 \cdot 9. \end{aligned}$$

Thus, $x = 13 \cdot 9 = 117 = 12$ in $\mathbb{Z}_{15}$. (You should verify that $7 \cdot 12 = 9$ in $\mathbb{Z}_{15}$.) Here, multiplying by 13 has the same effect as dividing by 7, but in a way that is still valid in $\mathbb{Z}_{15}$.

Remark. In the example above, the symbol $\implies$ denotes an implication. For instance,

$$13 \cdot (7 \cdot x) = 13 \cdot 9 \implies (13 \cdot 7) \cdot x = 13 \cdot 9$$

is a shorthand for "If $13 \cdot (7 \cdot x) = 13 \cdot 9$ is true, then $(13 \cdot 7) \cdot x = 13 \cdot 9$ is also true."

In $\mathbb{Z}_7$ and $\mathbb{Z}_5$, we found that every non-zero element has a multiplicative inverse. But this was *not* the case in $\mathbb{Z}_{15}$. In general, we ask the question: Which elements in $\mathbb{Z}_m$ have multiplicative inverses? To find patterns and make conjectures, here are additional data (try to find inverse pairs for each $\mathbb{Z}_m$):

$\mathbb{Z}_m$	Elements with multiplicative inverses
$\mathbb{Z}_5$	1, 2, 3, 4
$\mathbb{Z}_6$	1, 5
$\mathbb{Z}_7$	1, 2, 3 ,4, 5, 6
$\mathbb{Z}_8$	1, 3, 5, 7
$\mathbb{Z}_{10}$	1, 3, 7, 9
$\mathbb{Z}_{15}$	1, 2, 4, 7, 8, 11, 13, 14
$\mathbb{Z}_{21}$	1, 2, 4, 5, 8, 10, 11, 13, 16, 17, 19, 20

What's going on here? Looking at $\mathbb{Z}_{15}$, for instance, we see that the elements with multiplicative inverses are precisely those that are *relatively prime* to 15. In other words:

- If $\gcd(a, 15) = 1$, then a has a multiplicative inverse in $\mathbb{Z}_{15}$.
- If $\gcd(a, 15) \neq 1$, then a does not have a multiplicative inverse in $\mathbb{Z}_{15}$.

Example 4.18. In $\mathbb{Z}_{35}$, does 8 have a multiplicative inverse? What about 10? Applying the conjecture we found by examining $\mathbb{Z}_{15}$, we have:

- Since $\gcd(8, 35) = 1$, we believe 8 has a multiplicative inverse in $\mathbb{Z}_{35}$. (In fact, $8 \cdot 22 = 1$ in $\mathbb{Z}_{35}$.)
- Since $\gcd(10, 35) \neq 1$, we believe 10 does not have a multiplicative inverse in $\mathbb{Z}_{35}$.

And there is nothing special about $\mathbb{Z}_{15}$ or $\mathbb{Z}_{35}$ here. This conjecture is true in any $\mathbb{Z}_m$, which we will prove soon. Here is the generalization:

Conjecture (first draft). Let $a \in \mathbb{Z}_m$.

- If $\gcd(a, m) = 1$, then a has a multiplicative inverse in $\mathbb{Z}_m$.
- If $\gcd(a, m) \neq 1$, then a does not have a multiplicative inverse in $\mathbb{Z}_m$.

The contrapositive of the second statement is: If a has a multiplicative inverse in $\mathbb{Z}_m$, then $\gcd(a, m) = 1$. So, these two statements can be combined into one "if and only if" statement, namely:

Conjecture (second draft). Let $a \in \mathbb{Z}_m$. Then a has a multiplicative inverse in $\mathbb{Z}_m$ if and only if $\gcd(a, m) = 1$.

We also introduce a notation that will be used often:

$$U_m = \{a \in \mathbb{Z}_m \mid a \text{ has a multiplicative inverse in } \mathbb{Z}_m\}.$$

In other words, U_m is the subset of elements in $\mathbb{Z}_m$ that have multiplicative inverses. For example,

$$\begin{aligned} U_7 &= \{1, 2, 3, 4, 5, 6\}, \\ U_8 &= \{1, 3, 5, 7\}, \\ U_{15} &= \{1, 2, 4, 7, 8, 11, 13, 14\}. \end{aligned}$$

Remark. We use the notation U_m, because an element of $\mathbb{Z}_m$ with a multiplicative inverse (such as $7 \in \mathbb{Z}_{15}$) is called a *unit*. In fact, the set U_m is often called the "group of *units* modulo m." Here, "modulo m" indicates that the computation is done in $\mathbb{Z}_m$, i.e., by using the $\mathbb{Z}_m$ clock. The word "group" will be described in much more depth later in this book.

Using this new notation, we state our conjecture as follows:

Theorem 4.19 (Multiplicative inverses in $\mathbb{Z}_m$)**.** *Let $a \in \mathbb{Z}_m$. Then $a \in U_m$ if and only if* $\gcd(a, m) = 1$.

Since this theorem is an "if and only if" statement, there are two implications we must prove:

(a) If $a \in U_m$, then $\gcd(a, m) = 1$.

(b) If $\gcd(a, m) = 1$, then $a \in U_m$.

Below, we will prove implication (a). As usual, we begin with the hypothesis: Assume $a \in U_m$. Starting with this hypothesis, we repeatedly unravel its meaning. First, $a \in U_m$ means that a has a multiplicative inverse in $\mathbb{Z}_m$; and that means $ax = 1$ in $\mathbb{Z}_m$ for some $x \in \mathbb{Z}_m$; and that means m is a divisor of $ax - 1$; and so on. Eventually, we will show that the equation $ax + my = 1$ has an integer solution (x, y). Then Theorem 3.18 (the converse of the GCD theorem) implies that $\gcd(a, m) = 1$.

The proof of implication (b) is left as an exercise at the end of this chapter.

PROOF. Assume $a \in U_m$. Then, a has a multiplicative inverse in $\mathbb{Z}_m$. Thus, there exists $x \in \mathbb{Z}_m$ such that $ax = 1$ in $\mathbb{Z}_m$. Hence, $m \mid (ax - 1)$, so that $ax - 1 = mk$ for some $k \in \mathbb{Z}$. Rewriting this equation, we obtain $ax + m(-k) = 1$. Thus $\gcd(a, m) = 1$ by Theorem 3.18. ∎

Exercises

1. Compute each of the following in $\mathbb{Z}_7$. Simplify your result as much as possible.

   (a) $5 + 6$. (b) $6 \cdot 6$. (c) $3 - 5$. (d) 5^3. (e) 57,298.

2. (a) Describe all integers n such that $n = 0$ in $\mathbb{Z}_7$.
   (b) Describe all integers n such that $n = 2$ in $\mathbb{Z}_7$.

3. For each pair a and b, determine whether or not $a = b$ in $\mathbb{Z}_7$.

   (a) $a = 124$ and $b = 152$.
   (b) $a = 51$ and $b = 38$.
   (c) $a = 300$ and $b = 312$.
   (d) $a = 400{,}000$ and $b = 400{,}070$.

4. (a) Describe all integers n such that $n = 0$ in $\mathbb{Z}_9$.
   (b) Consider the following method for reducing the integer $n = 4{,}189{,}536$ in $\mathbb{Z}_9$:

   $$\begin{aligned}
   &4{,}189{,}536\\
   &= 4 \cdot \mathbf{10}^6 + 1 \cdot \mathbf{10}^5 + 8 \cdot \mathbf{10}^4 + 9 \cdot \mathbf{10}^3 + 5 \cdot \mathbf{10}^2 + 3 \cdot \mathbf{10}^1 + 6 \cdot \mathbf{10}^0\\
   &= 4 \cdot \mathbf{1}^6 + 1 \cdot \mathbf{1}^5 + 8 \cdot \mathbf{1}^4 + 9 \cdot \mathbf{1}^3 + 5 \cdot \mathbf{1}^2 + 3 \cdot \mathbf{1}^1 + 6 \cdot \mathbf{1}^0 \quad \leftarrow \text{ since } 10 = 1 \text{ in } \mathbb{Z}_9\\
   &= 4 + 1 + 8 + 9 + 5 + 3 + 6\\
   &= 36\\
   &= 0.
   \end{aligned}$$

   Based on this, is 9 a divisor of the integer $n = 4{,}189{,}536$? Explain your reasoning.
   (c) Use the method in part (b) to determine if 9 is a divisor of $n = 573{,}921$, $n = 123{,}456$, $n = 234{,}567$.
   (d) Describe a general method for determining if 9 is a divisor of an integer n.

5. Devise a method for determining if 3 is a divisor of an integer n.

6. Devise a method for determining if 11 is a divisor of an integer n. (**Hint:** $10 = -1$ in $\mathbb{Z}_{11}$.)

7. For each element in $\mathbb{Z}_{13}$, find its multiplicative inverse or explain why one does not exist.

8. Repeat Exercise #7 with $\mathbb{Z}_6$; with $\mathbb{Z}_{10}$; with $\mathbb{Z}_{16}$; with $\mathbb{Z}_{21}$.

9. For each of these, feel free to use Theorem 4.19.

   (a) List the elements of $U_{20} = \{a \in \mathbb{Z}_{20} \mid a \text{ has a multiplicative inverse in } \mathbb{Z}_{20}\}$. There should be 8 of them.
   (b) List the elements of U_{24}.
   (c) List the elements of U_p where p is prime.

10. (a) Consider the set $U_7 = \{1, 2, 3, 4, 5, 6\}$. For each $a \in U_7$, compute a^6.
    (b) For each $a \in U_5$, compute a^4.
    (c) For each $a \in U_{15}$, compute a^8.
    (d) For each $a \in U_{20}$, compute a^8.
    (e) For each $a \in U_{28}$, compute a^{12}.
    (f) Any conjectures?

    (This exercise is referenced in Chapter 20, Exercise #6(a).)

11. Consider $2 \in U_7$. The *order* of 2 refers to the smallest positive exponent n such that $2^n = 1$.

    (a) Verify that the order of 2 is 3.
    (b) Find the order of all other elements in U_7.
    (c) Find the order of each element in U_{10}.
    (d) Find the order of each element in U_{15}.
    (e) Any conjectures?

    (This exercise is referenced in Chapter 11, Example 12.24, and Section 20.3.)

12. (a) Complete the addition and multiplication tables below for U_{10}.

+	1	3	7	9
1				
3				
7				
9				

×	1	3	7	9
1				
3				
7				
9				

    (b) Is U_{10} closed under addition? Under multiplication? Why or why not?
    (c) Pick any row or column of the multiplication table. Notice anything? Can you explain it?

    (This exercise is referenced in Section 5.2.)

13. Recall that an element $a \in \mathbb{Z}_m$ is called a *self-inverse* if $a \cdot a = 1$ in $\mathbb{Z}_m$.

    (a) Verify that $4 \in \mathbb{Z}_5$ is a self-inverse.
    (b) Verify that $6 \in \mathbb{Z}_7$ is a self-inverse.
    (c) Verify that $9 \in \mathbb{Z}_{10}$ is a self-inverse.
    (d) Verify that $14 \in \mathbb{Z}_{15}$ is a self-inverse.
    (e) Explain why $m - 1 \in \mathbb{Z}_m$ is a self-inverse.

14. In $\mathbb{Z}_7$, compute 6^{231}. Also compute 3^{146}.

15. For each computation below, simplify your answer as much as possible.

    (a) $13^{2,000}$ in $\mathbb{Z}_{12}$.
    (b) $12^{2,001}$ in $\mathbb{Z}_{13}$.

16. (a) In $\mathbb{Z}_{79}$, verify that the multiplicative inverse of 9 is 44 and that the multiplicative inverse of 5 is 16.
    (b) Use the results in part (a) to find the multiplicative inverse of $9 \cdot 5$ in $\mathbb{Z}_{79}$.
    (c) Create your own example like the one in parts (a) and (b), but this time in $\mathbb{Z}_{101}$.
    Tip: In wolframalpha.com, try typing something like multiplicative inverse of 9 mod 79.

17. **Prove:** U_m is closed under multiplication; i.e., the product of any two elements of U_m is still in U_m. (This exercise is referenced in Example 8.6.)
    Note: Your proof may *not* use Theorem 4.19. But Exercise #16 above should help!

18. Complete the proof of Theorem 4.19 by proving its implication (b).

19. (a) Find all the elements in U_7.
    (b) Find all the elements in U_{13}.
    (c) Find all the elements in U_{101}.

20. **Prove:** Let p be a prime number. If $a \in \mathbb{Z}_p$ with $a \neq 0$, then $a \in U_p$.

21. Consider the equation $x^2 - 6x + 8 = 0$.
    (a) Working in $\mathbb{Z}$, find its solutions. How many are there?
    (b) Now let's work in $\mathbb{Z}_{15}$. Substitute the values $x = 0$, $x = 1$, $x = 2, \ldots, x = 14$ into the equation $x^2 - 6x + 8 = 0$. How many solutions did you find?
    (c) Elizabeth says,
    > "In $\mathbb{Z}_{15}$, the equation $x^2 - 6x + 8 = 0$ can be rewritten as $x^2 + 9x + 8 = 0$."
    What might she mean, and how does it relate to the solutions you found in part (b)?

22. A non-zero element a of $\mathbb{Z}_m$ is said to be a *zero divisor* if there exists a non-zero element b in $\mathbb{Z}_m$ such that $ab = 0$. For example, 5 is a zero divisor in $\mathbb{Z}_{20}$ because $5 \cdot 8 = 0$.
    (a) Find all zero divisors in $\mathbb{Z}_{20}$.
    (b) Find all zero divisors in $\mathbb{Z}_{12}$.
    (c) Find all zero divisors in $\mathbb{Z}_{13}$.
    (d) Any conjectures?

23. Consider the statement: *If $a = b$ in $\mathbb{Z}_m$ and $a = b$ in $\mathbb{Z}_n$, then $a = b$ in $\mathbb{Z}_{mn}$.* This is false. Find a counterexample that invalidates the statement.

24. **Prove:** If $a = b$ in $\mathbb{Z}_m$, $a = b$ in $\mathbb{Z}_n$, and $\gcd(m, n) = 1$, then $a = b$ in $\mathbb{Z}_{mn}$.

25. **(Some food for thought)** Compute
    $$a^6 + a^5 + a^4 + a^3 + a^2 + a + 1$$
    for each $a \in \mathbb{Z}_7$ with $a \neq 1$. Can you explain what's going on and why?

5

Symmetries

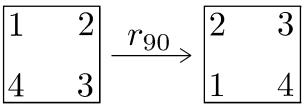
The set of integers $\mathbb{Z}$ is an example of a set whose elements can be added to obtain other elements in the same set (i.e., other integers). Recall that we say $\mathbb{Z}$ is *closed* under addition. The set $\mathbb{Z}$ together with the addition operation has other useful features. For instance, the zero element $0 \in \mathbb{Z}$ has the property that $0 + a = a$ and $a + 0 = a$ for all $a \in \mathbb{Z}$; i.e., when added to any integer, 0 keeps that integer unchanged. The number system $\mathbb{Z}_m$, which we explored in Chapter 4, has similar features to those of $\mathbb{Z}$. Perhaps this is not too surprising, since addition in $\mathbb{Z}_m$ is based on addition in $\mathbb{Z}$. (**Note:** While multiplication is also a valid operation in $\mathbb{Z}$ and in $\mathbb{Z}_m$, we will later see why we are only considering addition in this context.)

In this chapter, we will study certain motions of squares called *symmetries*. On the surface, these symmetries, which are geometric in nature, seem quite different from the integers, which we typically associate with arithmetic. However, these symmetries share many of the same useful features enjoyed by $\mathbb{Z}$ and $\mathbb{Z}_m$. In fact, the symmetries, $\mathbb{Z}$, and $\mathbb{Z}_m$ are all examples of *groups*, which we will formally introduce in Chapter 8. This process of extracting structural similarities that arise in different contexts is called *abstraction*, and it's what makes mathematics so powerful and beautiful.

5.1 Symmetries of a square

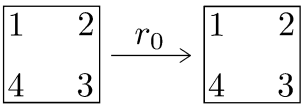
Suppose we take a square and move it in a way so that the square occupies the same space. In how many ways can this be done? We will study such motions in this chapter, with an emphasis on how those motions *interact* with each other.

Example 5.1. The motion r_{90} is a *counterclockwise* rotation about the center of the square by 90° (see figure below). Labels on the vertices indicate how the square was moved. Note how after the rotation, the square takes up the same space as it did before the rotation.

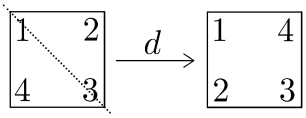


Definition 5.2. A *symmetry* of a square is a motion that, when applied to the square, places the square in the same space that it originally occupied.

Example 5.3. The motion r_0 is a counterclockwise rotation about the center of the square by 0° (see figure below). In other words, this "motion" does not move the square at all! While it may seem a bit boring, the motion r_0 will play a critical role in our work with symmetries.



Example 5.4. The symmetry denoted d (for "diagonal") is a reflection across the main diagonal of the square (see figure below). Labels on the vertices indicate how the square was moved. Note how after the reflection, the square takes up the same space as it did before the reflection.



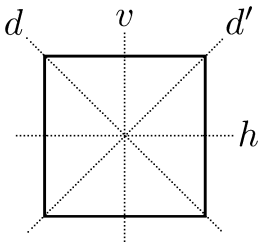
Example 5.5 (Non-examples)**.** Here are a couple of motions that are *not* symmetries of a square. Try drawing pictures to show how the square does *not* occupy the same space after applying each of these motions.

- A counterclockwise rotation about the center of the square by 45°.
- A translation by 2 units to the right. (Assume that the side length of the square is 1 unit.)

In all, there are 8 symmetries of a square:

- 4 rotations: r_0, r_{90}, r_{180}, r_{270}.

  Note: The motion r_0 is often denoted ε and is called the *identity*.
- 4 reflections: h, v, d, d', where the axes of reflections are shown below.



Remark. You might suggest r_{450} as another symmetry, i.e., a full 360° rotation followed by another 90°. However, we are interested in the net effect of the motion, rather than the motion itself. Thus, we will consider r_{450} to be equal to r_{90}, since both motions have the same net effect on the square.

We let D_4 be the set of symmetries of a square; i.e.,

$$D_4 = \{\varepsilon,\ r_{90},\ r_{180},\ r_{270},\ h,\ v,\ d,\ d'\}.$$

The set D_4 is often called the *dihedral group*, hence the use of the letter D in its name. As described in the beginning of this chapter, our goal is to identify structural similarities between D_4 and $\mathbb{Z}$. To do that, we must equip D_4 with an operation, i.e., a way of combining symmetries to obtain other symmetries, just as we can add integers to obtain other integers. The operation for D_4 is *composition*, as shown in the following example.

Example 5.6. Consider the elements $d, r_{90} \in D_4$. To compute $d \circ r_{90}$ (read "d composed with r_{90}"), we apply the composite motion $d \circ r_{90}$ onto the square:

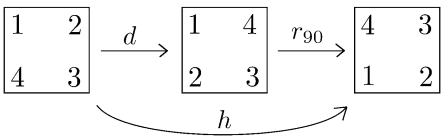
$$(d \circ r_{90}) \left(\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \right) = d \left(r_{90} \left(\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \right) \right) = d \left(\boxed{\begin{matrix} 2 & 3 \\ 1 & 4 \end{matrix}} \right) = \boxed{\begin{matrix} 2 & 1 \\ 3 & 4 \end{matrix}}$$

This has the same net effect as applying v onto the initial square:

$$v \left(\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \right) = \boxed{\begin{matrix} 2 & 1 \\ 3 & 4 \end{matrix}}$$

Thus, we conclude $d \circ r_{90} = v$. With $d \circ r_{90}$, note how the "inside" function r_{90} gets applied to the square first. This is analogous to $(f \circ g)(x) = f(g(x))$, which you may have seen in a calculus course.

Example 5.7. We compute $r_{90} \circ d$ by applying it onto the square. Noting that d is the "inside" function that gets applied to the square first, we use the following diagram:



The net effect is the same as applying h onto the initial square, and thus $r_{90} \circ d = h$. Comparing with Example 5.6 above, we see that $r_{90} \circ d \neq d \circ r_{90}$. This is different from what we are used to seeing in $\mathbb{Z}$, where $a + b = b + a$ for all $a, b \in \mathbb{Z}$.

Example 5.8. The figure below shows the composite motion $\varepsilon \circ r_{270}$ applied onto the square:

$$(\varepsilon \circ r_{270}) \left(\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \right) = \varepsilon \left(r_{270} \left(\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \right) \right) = \varepsilon \left(\boxed{\begin{matrix} 4 & 1 \\ 3 & 2 \end{matrix}} \right) = \boxed{\begin{matrix} 4 & 1 \\ 3 & 2 \end{matrix}}$$

Notice how in the last step, the motion ε keeps the input square unchanged. The composition $\varepsilon \circ r_{270}$ thus has the same net effect as applying just r_{270} onto the initial square:

$$r_{270} \left(\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \right) = \boxed{\begin{matrix} 4 & 1 \\ 3 & 2 \end{matrix}}$$

Therefore, we have $\varepsilon \circ r_{270} = r_{270}$. We can similarly show that $r_{270} \circ \varepsilon = r_{270}$. An analogous calculation in $\mathbb{Z}$ is $0 + a = a$ and $a + 0 = a$. In other words, the element $\varepsilon \in D_4$ behaves like $0 \in \mathbb{Z}$.

Example 5.9. The figure below shows the composite motion $r_{90} \circ r_{270}$ applied onto the square:

$$(r_{90} \circ r_{270}) \left(\begin{array}{|cc|} \hline 1 & 2 \\ 4 & 3 \\ \hline \end{array} \right) = r_{90} \left(r_{270} \left(\begin{array}{|cc|} \hline 1 & 2 \\ 4 & 3 \\ \hline \end{array} \right) \right) = r_{90} \left(\begin{array}{|cc|} \hline 4 & 1 \\ 3 & 2 \\ \hline \end{array} \right) = \begin{array}{|cc|} \hline 1 & 2 \\ 4 & 3 \\ \hline \end{array}$$

Thus, $r_{90} \circ r_{270} = \varepsilon$. We can similarly show that $r_{270} \circ r_{90} = \varepsilon$. We say that r_{90} and r_{270} are *inverses* of each other, because their composition is the identity element ε. An analogous calculation in $\mathbb{Z}$ is $a + (-a) = 0$ and $(-a) + a = 0$, where a and $-a$ are *additive inverses* of each other.

Example 5.10. Consider the element $h \in D_4$. The figure below shows the computation of $h \circ h = \varepsilon$. Thus, h is said to be a *self-inverse*, because the inverse of h is itself.

$$(h \circ h) \left(\begin{array}{|cc|} \hline 1 & 2 \\ 4 & 3 \\ \hline \end{array} \right) = h \left(h \left(\begin{array}{|cc|} \hline 1 & 2 \\ 4 & 3 \\ \hline \end{array} \right) \right) = h \left(\begin{array}{|cc|} \hline 4 & 3 \\ 1 & 2 \\ \hline \end{array} \right) = \begin{array}{|cc|} \hline 1 & 2 \\ 4 & 3 \\ \hline \end{array}$$

5.2 Group properties of D_4

Recall that $D_4 = \{\varepsilon, r_{90}, r_{180}, r_{270}, h, v, d, d'\}$ is the set of symmetries of a square. The composition table for D_4 is shown below. For $\sigma, \tau \in D_4$, the composition $\sigma \circ \tau$ is the entry in row σ and column τ. For example, $d \circ r_{90} = v$ from Example 5.6 is shown in a bold font.

$\circ$	ε	$\boldsymbol{r_{90}}$	r_{180}	r_{270}	h	v	d	d'
ε	ε	r_{90}	r_{180}	r_{270}	h	v	d	d'
r_{90}	r_{90}	r_{180}	r_{270}	ε	d'	d	h	v
r_{180}	r_{180}	r_{270}	ε	r_{90}	v	h	d'	d
r_{270}	r_{270}	ε	r_{90}	r_{180}	d	d'	v	h
h	h	d	v	d'	ε	r_{180}	r_{90}	r_{270}
v	v	d'	h	d	r_{180}	ε	r_{270}	r_{90}
$\boldsymbol{d}$	d	$\boldsymbol{v}$	d'	h	r_{270}	r_{90}	ε	r_{180}
d'	d'	h	d	v	r_{90}	r_{270}	r_{180}	ε

We say that D_4 (with the operation $\circ$) is a *group*, because it satisfies the following "group properties." In an exercise at the end of the chapter, you will verify that $\mathbb{Z}$ (with the operation $+$) is also a group.

(1) D_4 **is closed under composition.** For any pair of elements $\sigma, \tau \in D_4$, the composition $\sigma \circ \tau$ is also in D_4. We can see this from the table above, since every entry in the table (i.e., all possible compositions) is an element of D_4.

(2) **The associative law holds.** For any three elements $\sigma, \tau, \mu \in D_4$, we have $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$. For example, suppose our three elements are r_{90}, h, and d. Then

$$(r_{90} \circ h) \circ d = d' \circ d = r_{180} \qquad \text{and} \qquad r_{90} \circ (h \circ d) = r_{90} \circ r_{90} = r_{180},$$

so that $(r_{90} \circ h) \circ d = r_{90} \circ (h \circ d)$. In essence, the associative law says that we can change how the elements are grouped together (no pun intended!) without changing the result of the composition. We'll soon see that this law holds for any three elements in D_4.

(3) D_4 **has an identity element ε that keeps all elements in D_4 unchanged.** Note that $\varepsilon \circ \sigma = \sigma$ (the first row of the table) and $\sigma \circ \varepsilon = \sigma$ (the first column of the table) for all $\sigma \in D_4$. As discussed in Example 5.8, this is analogous to the calculation $0 + a = a$ and $a + 0 = a$ in $\mathbb{Z}$.

(4) **Every element in D_4 has an inverse.** We saw in Example 5.9 that r_{90} and r_{270} are inverses of each other, because $r_{90} \circ r_{270} = \varepsilon$ and $r_{270} \circ r_{90} = \varepsilon$. From Example 5.10, h is a *self-inverse*; i.e., $h \circ h = \varepsilon$. In an exercise at the end of the chapter, you will find the inverse of each element in D_4.

Notice how in each row or column of the table, every element of D_4 shows up exactly once. We saw the same feature in the multiplication table of U_{10} in Chapter 4, Exercise #12. In fact, D_4 and U_{10} also have many structural similarities, as we will soon learn.

5.3 Centralizer

In Chapter 2, we studied various subsets of $\mathbb{Z}$. In this section, we will examine an interesting subset of D_4. Fix $h \in D_4$, i.e., the horizontal reflection. (There is nothing special about h here. We could have fixed any element of D_4 for this process.) Then define the set

$$C(h) = \{\sigma \in D_4 \mid \sigma \circ h = h \circ \sigma\},$$

which is called the *centralizer* of h in D_4. In other words, $C(h)$ is the set of elements in D_4 that *commute* with h. The following elements are contained $C(h)$:

- $\varepsilon \in C(h)$, because $\varepsilon \circ h = h \circ \varepsilon$.
- $r_{180} \in C(h)$, because $r_{180} \circ h = h \circ r_{180}$.
- $h \in C(h)$, because $h \circ h = h \circ h$.
- $v \in C(h)$, because $v \circ h = h \circ v$.

On the other hand, $r_{90} \notin C(h)$, because $r_{90} \circ h \neq h \circ r_{90}$. You should verify for yourself that r_{270}, d, and d' are also not in $C(h)$. Therefore, we conclude that $C(h) = \{\varepsilon,\ r_{180},\ h,\ v\}$.

Here is the table for $C(h)$:

$\circ$	ε	r_{180}	h	v
ε	ε	r_{180}	h	v
r_{180}	r_{180}	ε	v	h
h	h	v	ε	r_{180}
v	v	h	r_{180}	ε

Let's check the group properties for $C(h)$:

(1) $C(h)$ is closed under composition. We can see this from the table above, since every entry in the table (i.e., all possible compositions) is an element of $C(h)$.

(2) The associative law holds. We've already discussed how $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$ for any three elements $\sigma, \tau, \mu \in D_4$. Since $C(h)$ is a subset of D_4 (i.e., any element of $C(h)$ is an element of D_4), the same must also hold for any three elements in $C(h)$.

(3) $C(h)$ contains the identity element ε.

(4) Every element in $C(h)$ has an inverse that is also in $C(h)$. In fact, we can see from the table above that each element of $C(h)$ is a self-inverse.

Thus, $C(h)$ is also a group. And since $C(h)$ is a subset of D_4, we say that $C(h)$ is a *subgroup* of D_4.

In the table for $C(h)$, we see, for instance, that $r_{180} \circ h = h \circ r_{180}$ and $v \circ r_{180} = r_{180} \circ v$. In fact, such a relationship holds for any pair of elements in $C(h)$. (How can we tell this fairly quickly by looking at the table?) This feature of $C(h)$ is captured by the following definition.

Definition 5.11 (Commutative group)**.** We say that $C(h)$ is a *commutative* group, since $\sigma \circ \tau = \tau \circ \sigma$ for all $\sigma, \tau \in C(h)$.

Commutativity should be a familiar concept. For instance, addition of integers is commutative, because $a + b = b + a$ for all $a, b \in \mathbb{Z}$. However, D_4 is a *non-commutative* group, because $r_{90} \circ d \neq d \circ r_{90}$.

Remark. In mathematics, commutative groups are more commonly referred to as *abelian* groups. In this textbook, however, we will continue to use the term "commutative," since this is likely more familiar to you from your prior experiences.

Example 5.12. Fix the identity element $\varepsilon \in D_4$. Then the centralizer of ε in D_4 is given by

$$C(\varepsilon) = \{\sigma \in D_4 \mid \sigma \circ \varepsilon = \varepsilon \circ \sigma\}.$$

But for all $\sigma \in D_4$, we have $\sigma \circ \varepsilon = \sigma$ and $\varepsilon \circ \sigma = \sigma$, so that $\sigma \circ \varepsilon = \varepsilon \circ \sigma$. Therefore, every element of D_4 is contained in $C(\varepsilon)$, and hence $C(\varepsilon) = D_4$.

For the set $C(h)$, we used its table to verify that every element in $C(h)$ has an inverse that is also in $C(h)$. Let's generalize this observation to $C(\tau)$, where τ is an arbitrary fixed element of D_4. Then the set $C(\tau)$ is defined by $C(\tau) = \{\sigma \in D_4 \mid \sigma \circ \tau = \tau \circ \sigma\}$. Now consider an element $\alpha \in C(\tau)$. We will show that α has an inverse that is also in $C(\tau)$. To start, we know that $\alpha \in D_4$, as $C(\tau)$ is a subset of D_4. Since every element in D_4 has an inverse, α must have one as well. Thus, let β be the inverse of α, so that $\alpha \circ \beta = \varepsilon$ and $\beta \circ \alpha = \varepsilon$. We must show that β is contained in $C(\tau)$.

Theorem 5.13. *Fix $\tau \in D_4$. If $\alpha \in C(\tau)$ and β is the inverse of α, then $\beta \in C(\tau)$.*

To prove it, we begin as usual by assuming the hypotheses: Assume $\alpha \in C(\tau)$ and β is the inverse of α. From these, we obtain the following equations:

- $\alpha \in C(\tau) \implies \alpha \circ \tau = \tau \circ \alpha$.
- β is the inverse of $\alpha \implies \alpha \circ \beta = \varepsilon$ and $\beta \circ \alpha = \varepsilon$.

The conclusion is $\beta \in C(\tau)$, which means we must show that $\beta \circ \tau = \tau \circ \beta$. Now that we have expressed the hypotheses and conclusion using equations, which are easier to manipulate, we are ready to write the proof.

PROOF. Assume $\alpha \in C(\tau)$ and β is the inverse of α. Since $\alpha \in C(\tau)$, we have $\alpha \circ \tau = \tau \circ \alpha$. Furthermore, $\alpha \circ \beta = \varepsilon$ and $\beta \circ \alpha = \varepsilon$, because β is the inverse of α. We take $\alpha \circ \tau = \tau \circ \alpha$ and left-compose by β on both sides to obtain $\beta \circ (\alpha \circ \tau) = \beta \circ (\tau \circ \alpha)$. The left side of this equation simplifies to

$$\beta \circ (\alpha \circ \tau) = (\beta \circ \alpha) \circ \tau = \varepsilon \circ \tau = \tau,$$

so that $\tau = \beta \circ (\tau \circ \alpha)$. Right-compose by β on both sides of this new equation to get $\tau \circ \beta = \beta \circ (\tau \circ \alpha) \circ \beta$, whose right side equals $\beta \circ (\tau \circ \alpha) \circ \beta = (\beta \circ \tau) \circ (\alpha \circ \beta) = (\beta \circ \tau) \circ \varepsilon = \beta \circ \tau$. Thus, $\tau \circ \beta = \beta \circ \tau$, which implies that $\beta \in C(\tau)$ as desired. ■

Remark. The above proof occurs in D_4, which is non-commutative. Thus, we must compose on the *same side* of an equation. For instance, we cannot take $\alpha \circ \tau = \tau \circ \alpha$ and obtain $\beta \circ (\alpha \circ \tau) = (\tau \circ \alpha) \circ \beta$, i.e., left-compose by β on one side of the equation and right-compose by β on the other.

Exercises

1. (a) Compute $\upsilon \circ r_{270}$ by drawing a figure like the one in Example 5.6.
   (b) Compute $r_{270} \circ \upsilon$ by drawing a figure like the one in Example 5.7.
   (c) Verify that $\upsilon \circ r_{270} \neq r_{270} \circ \upsilon$.

2. Verify that $\mathbb{Z}$ with addition satisfies the four group properties described in Section 5.2.

3. (a) Verify that $\mathbb{Z}_7$ with addition satisfies the four group properties described in Section 5.2.
   (b) Do the same with $\mathbb{Z}_5$; with $\mathbb{Z}_{12}$; with $\mathbb{Z}_{20}$; with $\mathbb{Z}_m$.

4. Find the inverse of each element in D_4. (Feel free to use the table for D_4.)

5. Fix $h \in D_4$ and let $C(h)$ be the centralizer of h in D_4. Verify that r_{270}, d, d' are not contained in $C(h)$.

6. Fix $d \in D_4$, i.e., the reflection across the main diagonal of the square. The centralizer of d in D_4 is defined by $C(d) = \{\sigma \in D_4 \mid \sigma \circ d = d \circ \sigma\}$. Find all the elements of $C(d)$.

7. The *center* of D_4 is defined by $Z(D_4) = \{\sigma \in D_4 \mid \sigma \circ \tau = \tau \circ \sigma \text{ for all } \tau \in D_4\}$. For example, we have $\varepsilon \in Z(D_4)$ because $\varepsilon \circ \tau = \tau \circ \varepsilon$ for all $\tau \in D_4$.

   Note: The notation $Z(D_4)$ originates from the German word Zentrum ("center").

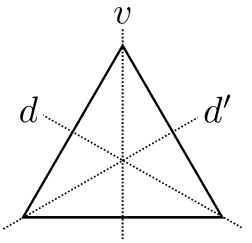
   (a) Describe the set $Z(D_4)$ in your own words. How does the center differ from a centralizer?
   (b) Find all the elements of $Z(D_4)$.
   (c) Recall that $C(r_{180}) = \{\sigma \in D_4 \mid \sigma \circ r_{180} = r_{180} \circ \sigma\}$. Find all elements of $C(r_{180})$.
   (d) Elizabeth says, "After doing part (b) of this problem, I could do part (c) *without* looking at the composition table for D_4." Why might she say that?

   (This exercise is referenced in Example 11.15.)

8. Let D_3 denote the set of symmetries of an equilateral triangle. We have

$$D_3 = \{\varepsilon, r_{120}, r_{240}, v, d, d'\},$$

where the rotations are counterclockwise and the axes for the reflections are as shown below.



   (a) Construct a composition table for D_3. (This exercise is referenced in Chapter 9, Exercise #20.)

   (b) Use the table created to check the group properties for D_3. (**Note:** Technically, $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$ should be checked for all $\sigma, \tau, \mu \in D_3$. But you can just do one example here. To make it interesting, though, choose σ, τ, μ to be three *different* elements of D_3.)

   (c) Is D_3 commutative or non-commutative?

9. For $n \geq 3$, let D_n denote the set of symmetries of a regular n-sided polygon.

   (a) Describe the elements of D_n. How many are there?

   (b) Explain why D_n is non-commutative. (The operation is composition.)

10. For an element $\sigma \in D_4$, the *order* of σ refers to the smallest positive exponent n such that

$$\sigma^n = \underbrace{\sigma \circ \sigma \circ \cdots \circ \sigma}_{n \text{ copies}} = \varepsilon.$$

    Find the order of each element in D_4. Any conjectures? (This exercise is referenced in Chapter 11 and Section 20.3.)

11. Find the order of each element in D_7. Any conjectures?

12. Find the order of each element in D_{10}. Any conjectures?

13. In D_n, explain the following:

   (a) Why a rotation composed with a rotation is a rotation.

   (b) Why a reflection composed with a reflection is a rotation.

   (c) Why a rotation composed with a reflection (in either order) is a reflection.

    (This exercise is referenced in Example 23.8.)

14. Consider the subset $H = \{0, 3, 6, 9\}$ of $\mathbb{Z}_{12}$.
    (a) Construct an addition table for H. Here, addition is done in $\mathbb{Z}_{12}$.
    (b) Use the table created to check the group properties for H. (**Note:** You may simply assume that the associative law for addition holds in $\mathbb{Z}_{12}$, i.e., $(a+b)+c = a+(b+c)$ for all $a, b, c \in \mathbb{Z}_{12}$.)
    (c) Is H commutative or non-commutative?

15. **Prove:** Fix $\tau \in D_n$. Then $C(\tau)$ is closed under composition.
    Note: In other words, prove the statement: *If* $\alpha, \beta \in C(\tau)$, *then* $\alpha \circ \beta \in C(\tau)$.
    (This exercise is referenced in Example 23.9.)

16. **Prove:** $Z(D_n)$ is closed under composition.
    Note: Here, $Z(D_n)$ is the *center* of D_n, as defined (for D_4) in Exercise #7.

17. **Prove:** Let $\alpha, \beta \in D_n$ be inverses of each other. If $\alpha \in Z(D_n)$, then $\beta \in Z(D_n)$.

18. Draw a figure whose symmetries include only r_0 and r_{180}.

19. Draw a figure whose symmetries include only r_0, r_{90}, r_{180}, and r_{270}.

20. Draw a figure whose symmetries include only r_0, r_{180}, h, and v.

21. Find all symmetries of a tetrahedron. How many are there?

22. Find all symmetries of a cube. How many are there?

23. **(Challenge)** Recall from Section 5.3 that $C(h)$ is a *subgroup* of D_4. This means that $C(h)$ is a subset of D_4 that also satisfy the four group properties described in Section 5.2. Find all subgroups of D_4. How do you know that you've found them all?

6

Permutations

In the last chapter, we studied D_4, the set of *symmetries* of a square. With the operation $\circ$ (composition), these symmetries satisfy the same group properties that are satisfied by the set of integers $\mathbb{Z}$ with addition (and also $\mathbb{Z}_m$ with addition). In this chapter, we will examine another example of a *group*, i.e., a set with an operation that satisfies the four group properties described in Section 5.2. Specifically, we will study a special type of functions called *permutations*. As we did with symmetries, we are interested in how these permutations *interact* with each other.

In the history of mathematics, concrete examples of groups preceded the formal definition of a group. Thus, the approach that we are taking in this textbook mimics this historical development. In particular, the group of permutations was the first kind of groups that were studied by mathematicians.

6.1 Permutations of the set $\{1, 2, 3\}$

Example 6.1. Let $\sigma : \{1, 2, 3\} \to \{1, 2, 3\}$ be a function defined by

$$\sigma(1) = 3,\ \sigma(2) = 1,\ \sigma(3) = 2.$$

We say that σ is a *permutation* of the set $\{1, 2, 3\}$; i.e., it "shuffles" the numbers 1, 2, and 3 so that the outputs $\sigma(1)$, $\sigma(2)$, and $\sigma(3)$ are all different. But a similar function defined by

$$f(1) = 3,\ f(2) = 1,\ f(3) = 1$$

is *not* a permutation, since $f(2)$ and $f(3)$ are equal.

Remark. In Example 6.1 above, the notation $\sigma : \{1, 2, 3\} \to \{1, 2, 3\}$ means the following:

- The name of the function is σ.
- The inputs into the function are 1, 2, and 3.
- The possible outputs are also 1, 2, and 3.

A more general discussion of functions is given in Chapter 15.

Example 6.2. Suppose τ is a permutation of $\{1, 2, 3\}$ where $\tau(1) = 2$ and $\tau(2) = 1$. Since $\tau(3)$ must be different from $\tau(1)$ and $\tau(2)$, we conclude that $\tau(3) = 3$.

Example 6.3. We define α to be a permutation of the set $\{1, 2, 3, 4\}$ given by

$$\alpha(1) = 2,\ \alpha(2) = 3,\ \alpha(3) = 4,\ \alpha(4) = 1.$$

Based on the examples above, here is a general definition of a permutation.

Definition 6.4 (Permutation)**.** A function $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ is called a *permutation* of the set $\{1, 2, \ldots, n\}$ when the outputs $\sigma(1), \sigma(2), \ldots, \sigma(n)$ are all different. Moreover, the set of all permutations of $\{1, 2, \ldots, n\}$ is denoted S_n.

Remark. The set S_n is often called the *symmetric group*, hence the use of the letter S in its name.

Example 6.5. Suppose $\sigma \in S_3$; i.e., σ is a permutation of the set $\{1, 2, 3\}$. Then there are 3 choices for $\sigma(1)$, since $\sigma(1)$ can equal either 1, 2, or 3. For $\sigma(2)$, there are only 2 choices, since $\sigma(2)$ cannot equal $\sigma(1)$. And for $\sigma(3)$, there is only 1 choice left, namely the number that was not taken up by $\sigma(1)$ or $\sigma(2)$. Thus, there are $3! = 3 \times 2 \times 1 = 6$ permutations in the set S_3.

To turn S_n into a group, we must equip it with an operation, i.e., a way of combining permutations to obtain other permutations, just as we can add integers to obtain other integers. As we did with the group of symmetries D_n, we will use the operation $\circ$ (composition) for S_n.

Example 6.6. Let $\sigma, \tau \in S_3$ be permutations of $\{1, 2, 3\}$ given by

$$\sigma(1) = 3,\ \sigma(2) = 1,\ \sigma(3) = 2 \quad \text{and} \quad \tau(1) = 2,\ \tau(2) = 1,\ \tau(3) = 3.$$

To compute $\sigma \circ \tau$, we apply the composite function to the inputs 1, 2, and 3:

- $(\sigma \circ \tau)(1) = \sigma(\tau(1)) = \sigma(2) = 1$.
- $(\sigma \circ \tau)(2) = \sigma(\tau(2)) = \sigma(1) = 3$.
- $(\sigma \circ \tau)(3) = \sigma(\tau(3)) = \sigma(3) = 2$.

Notice how τ is the "inside" function that gets applied first to 1, 2, and 3. We observe that $\sigma \circ \tau$ does indeed shuffle 1, 2, and 3, so it is a permutation. Thus, $\sigma \circ \tau \in S_3$.

Example 6.7. With $\sigma, \tau \in S_3$ as defined in Example 6.6, let's compute $\tau \circ \sigma$.

- $(\tau \circ \sigma)(1) = \tau(\sigma(1)) = \tau(3) = 3$.
- $(\tau \circ \sigma)(2) = \tau(\sigma(2)) = \tau(1) = 2$.
- $(\tau \circ \sigma)(3) = \tau(\sigma(3)) = \tau(2) = 1$.

Note how σ is now the "inside" function. We see that $\tau \circ \sigma$ is a permutation, since it shuffles the inputs 1, 2, and 3. Thus, $\tau \circ \sigma \in S_3$. However, comparing with Example 6.6, we see that $\sigma \circ \tau \neq \tau \circ \sigma$, because they shuffle 1, 2, and 3 in different ways. Therefore, S_3 is *non-commutative*.

6.2 Group properties of S_n

In this section, we will verify that S_n (with the operation $\circ$) satisfies the group properties. Let's begin by showing that S_n is closed under composition. Thus, assume σ, $\tau \in S_n$. We must show that $\sigma \circ \tau \in S_n$, i.e., that the outputs $(\sigma \circ \tau)(1)$, $(\sigma \circ \tau)(2)$, …, $(\sigma \circ \tau)(n)$ are all different.

Proof know-how. To show that $(\sigma\circ\tau)(1), (\sigma\circ\tau)(2), \ldots, (\sigma\circ\tau)(n)$ are all different, show that any two of them are different. More specifically, consider $(\sigma \circ \tau)(i)$ and $(\sigma \circ \tau)(j)$, where i, j are integers between 1 and n with $i \neq j$. Then show that $(\sigma\circ\tau)(i) \neq (\sigma\circ\tau)(j)$.

The proof of the following theorem is left for you as an exercise at the end of the chapter.

Theorem 6.8. *S_n is closed under composition.*

To introduce the notion of the identity element and inverses in S_n, consider the following examples.

Example 6.9. Let $\sigma,\ \gamma,\ \varepsilon \in S_3$ be given by the following:

- $\sigma(1) = 3,\ \sigma(2) = 1,\ \sigma(3) = 2.$
- $\gamma(1) = 2,\ \gamma(2) = 3,\ \gamma(3) = 1.$
- $\varepsilon(1) = 1,\ \ \varepsilon(2) = 2,\ \varepsilon(3) = 3.$

Here is the computation of $\varepsilon \circ \sigma$:

- $(\varepsilon \circ \sigma)(1) = \varepsilon(\sigma(1)) = \varepsilon(3) = 3.$
- $(\varepsilon \circ \sigma)(2) = \varepsilon(\sigma(2)) = \varepsilon(1) = 1.$
- $(\varepsilon \circ \sigma)(3) = \varepsilon(\sigma(3)) = \varepsilon(2) = 2.$

Notice how in the last step the permutation ε keeps the inputs unchanged. Thus, the composition $\varepsilon \circ \sigma$ shuffles 1, 2, and 3 in the same manner as σ. Therefore, we have $\varepsilon\circ\sigma = \sigma$. We can similarly show that $\sigma\circ\varepsilon = \sigma$ and that these relationships hold with σ replaced by any other element of S_3. Hence, ε is the identity element of S_3. Composing with ε is analogous to $0 + a = a$ and $a + 0 = a$ in $\mathbb{Z}$; i.e., the element $\varepsilon \in S_3$ behaves like $0 \in \mathbb{Z}$.

Here is the generalization of Example 6.9 above. In an exercise, you'll show that the element ε has the property that $\varepsilon \circ \alpha = \alpha$ and $\alpha \circ \varepsilon = \alpha$ for all $\alpha \in S_n$.

Definition 6.10 (Identity permutation)**.** The element $\varepsilon \in S_n$ defined by $\varepsilon(1) = 1$, $\varepsilon(2) = 2, \ldots, \varepsilon(n) = n$ is called the *identity permutation* in S_n.

Example 6.11. With $\sigma,\ \gamma,\ \varepsilon \in S_3$ as defined in Example 6.9, consider the following computations:

- $(\sigma \circ \gamma)(1) = \sigma(\gamma(1)) = \sigma(2) = 1$.
- $(\sigma \circ \gamma)(2) = \sigma(\gamma(2)) = \sigma(3) = 2$.
- $(\sigma \circ \gamma)(3) = \sigma(\gamma(3)) = \sigma(1) = 3$.

Thus, $\sigma\circ\gamma$ shuffles 1, 2, and 3 in the same manner as ε, so that $\sigma\circ\gamma = \varepsilon$. We can similarly show that $\gamma\circ\sigma = \varepsilon$. We say σ and γ are *inverses* of each other, because their composition is the identity element ε. This is analogous to $a + (-a) = 0$ and $(-a) + a = 0$ in $\mathbb{Z}$.

Definition 6.12 (Inverses in S_n)**.** Let $\sigma, \gamma \in S_n$ such that $\sigma \circ \gamma = \varepsilon$ and $\gamma \circ \sigma = \varepsilon$. We say that γ is the *inverse* of σ, and we write $\gamma = \sigma^{-1}$. Similarly, σ is the inverse of γ, and we write $\sigma = \gamma^{-1}$.

We conclude this section by verifying the group properties for S_n (with the operation $\circ$).

(1) S_n **is closed under composition.** For any pair of elements $\sigma, \tau \in S_n$, the composition $\sigma \circ \tau$ is also in S_n. This is Theorem 6.8.

(2) **The associative law holds.** For any three elements $\sigma, \tau, \mu \in S_n$, we have $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$. This is justified in Section 6.4.

(3) S_n **has an identity element** ε **that keeps all elements in** S_n **unchanged.** See Definition 6.10.

(4) **Every element in** S_n **has an inverse.** In Example 6.11, we saw that $\sigma, \gamma \in S_3$ are inverses of each other, since $\sigma \circ \gamma = \varepsilon$ and $\gamma \circ \sigma = \varepsilon$. You will generalize this in an exercise at the end of the chapter.

6.3 Computations in S_n

In this section, we will perform various computations in S_n that will help us better understand its group properties. Before proceeding, we introduce a useful notation to describe these permutations.

Definition 6.13 (Matrix notation)**.** Consider $\sigma \in S_3$ defined by $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 2$. We can write σ in *matrix form* like this:

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

For a permutation in S_3, the top row is always 1, 2, 3. The bottom row contains the corresponding outputs.

Example 6.14. We revisit Example 6.6 using this new matrix notation. Thus, let σ, $\tau \in S_3$ be defined by

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

To compute $\sigma \circ \tau$, note that τ is the "inside" function. To find $(\sigma \circ \tau)(1)$, for instance, we see that τ maps 1 to 2, and then σ maps 2 to 1. Thus, $(\sigma \circ \tau)(1) = 1$. Finding $(\sigma \circ \tau)(2)$ and $(\sigma \circ \tau)(3)$ similarly, we obtain

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

Example 6.15. Let $\alpha \in S_5$ be defined by

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

We begin by finding α^{-1}, the inverse of α. The matrix notation tells us that $\alpha(1) = 4$. Since α^{-1} undoes the effect of α, we must have $\alpha^{-1}(4) = 1$. Thus, we have $(\alpha^{-1} \circ \alpha)(1) = \alpha^{-1}(\alpha(1)) = \alpha^{-1}(4) = 1$; i.e., $\alpha^{-1} \circ \alpha$ maps 1 to 1, just as ε does. Continuing in this manner, we obtain

$$\alpha^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}.$$

We can also obtain α^{-1} by swapping the rows of the matrix of α, which has the effect of swapping the roles of inputs and outputs of α. After swapping the rows, we must also rearrange the columns, so that the new top row would read 1, 2, 3, 4, 5:

$$\alpha^{-1} = \begin{pmatrix} 4 & 1 & 3 & 5 & 2 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} \Longrightarrow \alpha^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}.$$

Example 6.16. Let $\alpha \in S_5$ be defined as in Example 6.15, and also let $\beta \in S_5$ be given by

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

We wish to find $\chi \in S_5$ such that $\alpha \circ \chi = \beta$.

- We have $\alpha(\chi(1)) = (\alpha \circ \chi)(1) = \beta(1) = 1$. Thus, α maps $\chi(1)$ to 1. We must have $\chi(1) = 2$.
- We have $\alpha(\chi(2)) = (\alpha \circ \chi)(2) = \beta(2) = 3$. Thus, α maps $\chi(2)$ to 3. We must have $\chi(2) = 3$.
- We have $\alpha(\chi(3)) = (\alpha \circ \chi)(3) = \beta(3) = 4$. Thus, α maps $\chi(3)$ to 4. We must have $\chi(3) = 1$.
- We have $\alpha(\chi(4)) = (\alpha \circ \chi)(4) = \beta(4) = 5$. Thus, α maps $\chi(4)$ to 5. We must have $\chi(4) = 4$.
- We have $\alpha(\chi(5)) = (\alpha \circ \chi)(5) = \beta(5) = 2$. Thus, α maps $\chi(5)$ to 2. We must have $\chi(5) = 5$.

Therefore, we conclude that

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

Below is an alternative approach to finding χ. The equation $\alpha \circ \chi = \beta$ in S_5 is analogous to $6x = 17$ (let's view 6 and 17 as real numbers), where α and β are known and χ is the unknown. With $6x = 17$, we can solve for x by multiplying both sides

by $\frac{1}{6}$ (i.e., the multiplicative inverse of 6). We obtain $\frac{1}{6} \cdot 6x = \frac{1}{6} \cdot 17$, so that $x = \frac{17}{6}$. Analogously, we left-compose both sides of $\alpha \circ \chi = \beta$ by α^{-1} to obtain $\alpha^{-1} \circ (\alpha \circ \chi) = \alpha^{-1} \circ \beta$, so that $\chi = \alpha^{-1} \circ \beta$. Using the matrix form of α^{-1} from Example 6.15, we obtain

$$\chi = \alpha^{-1} \circ \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix},$$

as before. You should verify that $\alpha \circ \chi$ does indeed equal β.

The next theorem and its proof explain how $\alpha \circ \chi = \beta$ implies $\chi = \alpha^{-1} \circ \beta$ using the group properties of S_n. Also, compare this with how we solved the equation $7 \cdot x = 9$ in $\mathbb{Z}_{15}$ in Example 4.17.

Theorem 6.17. *Let $\alpha, \beta, \chi \in S_n$. If $\alpha \circ \chi = \beta$, then $\chi = \alpha^{-1} \circ \beta$.*

PROOF. Assume $\alpha \circ \chi = \beta$. Since $\alpha \in S_5$, it has an inverse element α^{-1} such that $\alpha^{-1} \circ \alpha = \varepsilon$. We left-compose both sides of $\alpha \circ \chi = \beta$ by α^{-1} to obtain $\alpha^{-1} \circ (\alpha \circ \chi) = \alpha^{-1} \circ \beta$. Working with the left-hand side of this equation, we obtain

$$\begin{aligned} \alpha^{-1} \circ (\alpha \circ \chi) &= (\alpha^{-1} \circ \alpha) \circ \chi && \text{(associative law)} \\ &= \varepsilon \circ \chi && (\alpha^{-1} \text{ is the inverse of } \alpha) \\ &= \chi && (\varepsilon \text{ is the identity element}). \end{aligned}$$

Thus, we obtain $\chi = \alpha^{-1} \circ \beta$ as desired. ■

6.4 Associative law in S_n (and in D_n)

For groups D_n and S_n, the operation is function composition. We have claimed—and verified with concrete examples—that for any three elements σ, τ, μ, we have

$$(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu).$$

Let's show that this is true in general. We'll make an argument in S_n, but a similar argument can be made for D_n as well. So, suppose $\sigma, \tau, \mu \in S_n$. Then these are permutations of $\{1, 2, 3, \ldots, n\}$. To show that $(\sigma \circ \tau) \circ \mu$ and $\sigma \circ (\tau \circ \mu)$ are equal, we must show that they "shuffle" the numbers $1, 2, 3, \ldots, n$ in the same way. Let's see where $(\sigma \circ \tau) \circ \mu$ and $\sigma \circ (\tau \circ \mu)$ map 1:

- $[(\sigma \circ \tau) \circ \mu](1) = (\sigma \circ \tau)(\mu(1)) = \sigma(\tau(\mu(1)))$.
- $[\sigma \circ (\tau \circ \mu)](1) = \sigma((\tau \circ \mu)(1)) = \sigma(\tau(\mu(1)))$.

So, the two functions agree on the input 1. There's nothing special about 1 here, and it can be replaced by any of the numbers $1, 2, 3, \ldots, n$. Therefore, $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$.

Exercises

1. Find the number of elements in S_4; in S_5; in S_6; in S_n.
2. Define $\varepsilon, \sigma, \gamma, \tau, \mu, \delta \in S_3$, respectively, by

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

(a) Construct the composition table for S_3.
(b) Use the table created to verify the group properties for S_3.
(c) Is S_3 commutative or non-commutative?

(This exercise is referenced in Chapter 24, Exercise #18.)

3. Here, $\gamma, \tau \in S_3$ are as defined in Exercise #2.
   (a) The *centralizer* of γ in S_3 is $C(\gamma) = \{\alpha \in S_3 \mid \alpha \circ \gamma = \gamma \circ \alpha\}$. Find all the elements of $C(\gamma)$.
   (b) Find all the elements of $C(\tau)$.

4. Here, $\sigma, \gamma \in S_3$ are as defined in Exercise #2.
   (a) Construct a composition table for the set $H = \{\varepsilon, \sigma, \gamma\}$.

$\circ$	ε	σ	γ
ε			
σ			
γ			

   (b) Use the table created to verify the group properties for H.
   (c) Is H commutative or non-commutative?

   (This exercise is referenced in Chapter 24, Exercise #18.)

5. The *center* of S_3 is $Z(S_3) = \{\alpha \in S_3 \mid \alpha \circ \beta = \beta \circ \alpha \text{ for all } \beta \in S_3\}$. Find all the elements of $Z(S_3)$.

6. Here, $\sigma, \tau \in S_3$ are as defined in Exercise #2.
   (a) Verify that $(\sigma \circ \tau)^{-1} \neq \sigma^{-1} \circ \tau^{-1}$, but $(\sigma \circ \tau)^{-1} = \tau^{-1} \circ \sigma^{-1}$.
   (b) Draw an analogy between the statement $(\sigma \circ \tau)^{-1} = \tau^{-1} \circ \sigma^{-1}$ and the act of putting on and taking off your socks and shoes.

   (This exercise is referenced in Section 8.2.)

7. For an element $\alpha \in S_3$, the *order* of α refers to the smallest positive exponent n such that
$$\alpha^n = \underbrace{\alpha \circ \alpha \circ \cdots \circ \alpha}_{n \text{ copies}} = \varepsilon.$$
   Find the order of each element in S_3. Any conjectures? (This exercise is referenced in Chapter 11 and Section 20.3.)

8. **Prove:** S_n is non-commutative for $n \geq 3$.

9. Let $\varepsilon \in S_n$ be the element defined by $\varepsilon(1) = 1, \varepsilon(2) = 2, \ldots, \varepsilon(n) = n$. Show that $\varepsilon \circ \alpha = \alpha$ and $\alpha \circ \varepsilon = \alpha$ for all $\alpha \in S_n$.

   Note: This shows that $\varepsilon \in S_n$ has the desired behavior as an identity element of S_n.

10. Let $\sigma \in S_n$. Explain why there exists $\gamma \in S_n$ such that $\sigma \circ \gamma = \varepsilon$ and $\gamma \circ \sigma = \varepsilon$.

    Note: This shows that every element in S_n has an inverse (which is also in S_n).

11. Consider the following elements in S_4:

$$\gamma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \text{ and } \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}.$$

    (a) Compute $\gamma \circ \tau$ and $\tau \circ \gamma$.

    (b) Compute $(\gamma \circ \sigma) \circ \tau$ and $\gamma \circ (\sigma \circ \tau)$ and verify that they are equal.

    (c) Find the inverse of each of these elements.

12. Let $\gamma, \sigma, \tau \in S_4$ be defined as in Exercise #11.

    (a) Find the order of each of these elements.

    (b) Find an element in S_4 of order 4.

13. (a) Can an element of S_4 have order greater than 4? Either find such an element or explain why not.

    (b) Can an element of S_5 have order greater than 5? Either find such an element or explain why not.

14. **Prove:** Let $\sigma, \tau, \mu \in S_n$. If $\sigma \circ \tau = \sigma \circ \mu$, then $\tau = \mu$. (This exercise is referenced in Section 8.2.)

15. Let $\sigma, \tau, \mu \in S_n$. Does $\sigma \circ \tau = \mu \circ \sigma$ imply that $\tau = \mu$? Support your answer.

16. Each motion of a square in D_4 can be viewed as an element of S_4 that shuffles the vertices of the square. For instance, here is a visual depiction of $r_{90} \in D_4$:

$$\boxed{\begin{matrix} 1 & 2 \\ 4 & 3 \end{matrix}} \xrightarrow{r_{90}} \boxed{\begin{matrix} 2 & 3 \\ 1 & 4 \end{matrix}}$$

    Then, r_{90} maps vertex 1 to the position that was previously occupied by vertex 4. It maps vertex 2 to the position that was previously occupied by vertex 1, and similarly for vertices 3 and 4. Thus, $r_{90} \in D_4$ corresponds to $\sigma \in S_4$ given by

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

    (a) For each element in D_4, find a corresponding element in S_4.

    (b) Here's another element of S_4:

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

    Either find an element in D_4 that corresponds to τ or explain why none exists.

    (c) We've seen that D_4 and S_4 have eight and twenty-four elements, respectively. So not every element in S_4 can have a corresponding element in D_4. Explain *geometrically* why these sixteen remaining elements of S_4 do not permit a corresponding element in D_4.

17. Let H be the subset of S_5 defined by $H = \{\sigma \in S_5 \mid \sigma(3) = 3\}$. For example, suppose

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 5 & 2 \end{pmatrix}.$$

Then $\alpha \in H$, because $\alpha(3) = 3$. But $\beta \notin H$, since $\beta(3) \neq 3$.

(a) Find several elements of S_5 that are in H.
(b) Find several elements of S_5 that are *not* in H.
(c) Choose two elements in H that you found in part (a) and call them σ and τ. Compute $\sigma \circ \tau$, $\tau \circ \sigma$, σ^{-1}, and τ^{-1}, and verify that all of these are still in H.
(d) Repeat part (c) with two more elements in H.

18. Let $H = \{\sigma \in S_5 \mid \sigma(3) = 3\}$, as defined in Exercise #17. Find the number of elements in H.

19. Let $H = \{\sigma \in S_5 \mid \sigma(3) = 3\}$. **Prove:** H is closed under composition.

20. Let $H = \{\sigma \in S_5 \mid \sigma(3) = 3\}$. **Prove:** If $\alpha \in H$, then $\alpha^{-1} \in H$.

21. Let $\alpha, \beta \in S_4$ where

$$\alpha \circ \beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix} \quad \text{and} \quad \beta \circ \alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

Moreover, suppose $\beta(3) = 1$. Find α and β.

22. Prove Theorem 6.8. Be sure to read the discussion before the theorem.

23. **(Challenge)** In Exercise #4, H is a *subgroup* of S_3, because H is a subset of S_3 that also satisfies the group properties. Find all subgroups of S_3. How do you know that you've found them all?

7

Matrices

This is the last chapter before the concept of a *group* is formally introduced in Chapter 8. We will explore one more example of a group, namely, the set of *matrices*, which are rectangular arrays of numbers. In this textbook, we will work primarily with 2×2 matrices with 2 rows and 2 columns. Just like the integers, matrices can be added or multiplied, and both operations will be considered in our study.

Beyond serving as an example of a group, matrices play an important role in many areas of mathematics, particularly in a branch of mathematics called *linear algebra*. If you have studied linear algebra in the past, some of the concepts in this chapter may seem familiar.

7.1 Matrix arithmetic

Example 7.1. Consider a pair of 2×2 matrices $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]$. For an extra twist, assume that the entries of these matrices are in $\mathbb{Z}_{10}$. (**Note:** An "entry" of a matrix refers to one of its numbers.) To add these matrices, we add the corresponding entries. Therefore,

$$\alpha + \beta = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 0 & 2 \end{bmatrix}.$$

Note that we reduced the entries in the bottom row; i.e., $3+7 = 0$ and $4+8 = 2$ in $\mathbb{Z}_{10}$.

Example 7.2. Let $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]$ be the matrices from Example 7.1, again with the assumption that the entries are in $\mathbb{Z}_{10}$. Then the difference $\alpha - \beta$ is given by

$$\alpha - \beta = \begin{bmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix}.$$

In the last step, note that $-4 = 6$ in $\mathbb{Z}_{10}$.

To multiply a pair of matrices, we do *not* multiply the corresponding entries. Instead, we perform the process that is described in the next example.

Example 7.3. Let α and β be the matrices from Example 7.1, with the entries in $\mathbb{Z}_{10}$. The product $\alpha \cdot \beta$ is also a 2×2 matrix whose entries are computed as follows:

- To compute its entry in the first row / first column, we take the first row from α (1 and 2) and the first column from β (5 and 7). Then we compute their *dot product* $1 \cdot 5 + 2 \cdot 7 = 9$ in $\mathbb{Z}_{10}$.
- To compute its entry in the first row / second column, we take the first row from α (1 and 2) and the second column from β (6 and 8). Then we compute their dot product $1 \cdot 6 + 2 \cdot 8 = 2$ in $\mathbb{Z}_{10}$.
- To compute its entry in the second row / first column, we take the second row from α (3 and 4) and the first column from β (5 and 7). Then we compute their dot product $3 \cdot 5 + 4 \cdot 7 = 3$ in $\mathbb{Z}_{10}$.
- To compute its entry in the second row / second column, we take the second row from α (3 and 4) and the second column from β (6 and 8). Then we compute their dot product $3 \cdot 6 + 4 \cdot 8 = 0$ in $\mathbb{Z}_{10}$.

Thus the product $\alpha \cdot \beta$ is given by

$$\alpha \cdot \beta = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{bmatrix} = \begin{bmatrix} 9 & 2 \\ 3 & 0 \end{bmatrix}.$$

Remark. Matrix multiplication is defined in this manner for the following reason. In linear algebra, matrices are used to represent a type of functions called *linear transformations*. And matrix multiplication, as defined in Example 7.3, corresponds to composition of linear transformations.

Example 7.4. Once again, let α and β be the matrices from Example 7.1, with the entries in $\mathbb{Z}_{10}$. The calculation below shows that $\alpha \cdot \beta \neq \beta \cdot \alpha$; i.e., matrix multiplication is *not* commutative.

$$\beta \cdot \alpha = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 + 6 \cdot 3 & 5 \cdot 2 + 6 \cdot 4 \\ 7 \cdot 1 + 8 \cdot 3 & 7 \cdot 2 + 8 \cdot 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 6 \end{bmatrix}.$$

Example 7.5. Here is an example of *scalar multiplication*, where we multiply a number by a matrix. Again, the entries of the matrix are in $\mathbb{Z}_{10}$.

$$8 \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 8 \cdot 1 & 8 \cdot 2 \\ 8 \cdot 3 & 8 \cdot 4 \end{bmatrix} = \begin{bmatrix} 8 & 6 \\ 4 & 2 \end{bmatrix}.$$

7.2 Matrix group $M(\mathbb{Z}_{10})$

To form a group using matrices, we need to put them into a set.

Definition 7.6. Let $M(\mathbb{Z}_{10})$ be the set of 2×2 matrices with entries in $\mathbb{Z}_{10}$.

The following observations about $M(\mathbb{Z}_{10})$ are based in part on the examples in Section 7.1.

- $M(\mathbb{Z}_{10})$ is closed under addition and under multiplication. See Examples 7.1 and 7.3, respectively.

- Addition in $M(\mathbb{Z}_{10})$ is commutative. You will prove this in an exercise at the end of the chapter.
- Multiplication in $M(\mathbb{Z}_{10})$ is *not* commutative. See Example 7.4.
- Both addition and multiplication in $M(\mathbb{Z}_{10})$ are associative. See the explanation below.

Since addition in $M(\mathbb{Z}_{10})$ is based on addition in $\mathbb{Z}_{10}$, the associativity of matrix addition follows from the associativity of addition in $\mathbb{Z}_{10}$. For the proof, we consider three arbitrary matrices $\alpha, \beta, \gamma \in M(\mathbb{Z}_{10})$ and show that $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

Theorem 7.7. *Addition in $M(\mathbb{Z}_{10})$ is associative.*

PROOF. Let $\alpha, \beta, \gamma \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, $\beta = \left[\begin{smallmatrix} q & r \\ s & t \end{smallmatrix}\right]$, $\gamma = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]$, and all the entries shown are in $\mathbb{Z}_{10}$. We have

$$
\begin{aligned}
(\alpha+\beta)+\gamma &= \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} q & r \\ s & t \end{bmatrix}\right) + \begin{bmatrix} w & x \\ y & z \end{bmatrix} \\
&= \begin{bmatrix} a+q & b+r \\ c+s & d+t \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} \\
&= \begin{bmatrix} (a+q)+w & (b+r)+x \\ (c+s)+y & (d+t)+z \end{bmatrix} \\
&= \begin{bmatrix} a+(q+w) & b+(r+x) \\ c+(s+y) & d+(t+z) \end{bmatrix} \quad \leftarrow \text{ addition in } \mathbb{Z}_{10} \text{ is associative} \\
&= \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} q+w & r+x \\ s+y & t+z \end{bmatrix} \\
&= \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \left(\begin{bmatrix} q & r \\ s & t \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix}\right) \\
&= \alpha + (\beta + \gamma).
\end{aligned}
$$

Thus, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ as desired. ■

The proof that multiplication in $M(\mathbb{Z}_{10})$ is associative is left for you as an exercise.

Example 7.8. Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ with its entries in $\mathbb{Z}_{10}$. We have $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right] + \alpha = \alpha$ as shown:

$$
\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \alpha = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0+a & 0+b \\ 0+c & 0+d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \alpha.
$$

Similarly, we can show that $\alpha + \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right] = \alpha$. Therefore, $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ is the additive identity element of $M(\mathbb{Z}_{10})$. This is analogous to $0 + a = a$ and $a + 0 = 0$ in $\mathbb{Z}_{10}$.

Example 7.9. Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 9 & 8 \\ 7 & 6 \end{smallmatrix}\right]$. We have

$$
\alpha + \beta = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 9 & 8 \\ 7 & 6 \end{bmatrix} = \begin{bmatrix} 1+9 & 2+8 \\ 3+7 & 4+6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
$$

Therefore, $\alpha + \beta = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Similarly, we can show that $\beta + \alpha = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Therefore, α and β are *additive inverses* of each other, since their sum is the additive identity element $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

We conclude this section by verifying the group properties for $M(\mathbb{Z}_{10})$ with addition.

(1) $M(\mathbb{Z}_{10})$ **is closed under addition.** For any pair of matrices $\alpha, \beta \in M(\mathbb{Z}_{10})$, their sum $\alpha + \beta$ is also in $M(\mathbb{Z}_{10})$. After all, addition in $M(\mathbb{Z}_{10})$ is based on addition in $\mathbb{Z}_{10}$.

(2) **The associative law (for addition) holds.** This is Theorem 7.7.

(3) $M(\mathbb{Z}_{10})$ **has an additive identity element** $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. See Example 7.8.

(4) **Every element in** $M(\mathbb{Z}_{10})$ **has an additive inverse.** In Example 7.9, we saw that $\alpha, \beta \in M(\mathbb{Z}_{10})$ are additive inverses of each other. You will generalize this in an exercise at the end of the chapter.

7.3 Multiplicative inverses

In Section 7.2, we saw that $M(\mathbb{Z}_{10})$ is a group under addition. We'll soon see that $M(\mathbb{Z}_{10})$ with multiplication is *not* a group, because not every matrix has a multiplicative inverse. Nonetheless, multiplication plays an important role in our work with matrices. In the examples below, we will discuss the notions of *identity* and *inverse* when the matrix operation is multiplication.

Example 7.10. Let $\varepsilon, \alpha \in M(\mathbb{Z}_{10})$ where $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ with its entries in $\mathbb{Z}_{10}$. We have

$$\varepsilon \cdot \alpha = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 \cdot a + 0 \cdot c & 1 \cdot b + 0 \cdot d \\ 0 \cdot a + 1 \cdot c & 0 \cdot b + 1 \cdot d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Therefore, $\varepsilon \cdot \alpha = \alpha$. A similar calculation shows that $\alpha \cdot \varepsilon = \alpha$. The matrix ε is called the *multiplicative identity* matrix, because it keeps all the matrices unchanged via multiplication. This is analogous to multiplying a number by 1; i.e., $1 \cdot a = a$ and $a \cdot 1 = a$ in $\mathbb{Z}_{10}$.

In $\mathbb{Z}_{10}$, we have $3 \cdot 7 = 1$ and $7 \cdot 3 = 1$, where 1 is the multiplicative identity of $\mathbb{Z}_{10}$. Recall that we say 3 and 7 are *multiplicative inverses* of each other. Let's apply the same concept to matrices in $M(\mathbb{Z}_{10})$.

Example 7.11. Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 4 & 9 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 9 & 8 \\ 6 & 1 \end{smallmatrix}\right]$. We have

$$\alpha \cdot \beta = \begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix} \cdot \begin{bmatrix} 9 & 8 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 9 + 2 \cdot 6 & 1 \cdot 8 + 2 \cdot 1 \\ 4 \cdot 9 + 9 \cdot 6 & 4 \cdot 8 + 9 \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Therefore, $\alpha \cdot \beta = \varepsilon$. Similarly, we can show that $\beta \cdot \alpha = \varepsilon$. Therefore, α and β are *multiplicative inverses* of each other, since their product is the multiplicative identity element $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.

Example 7.12. Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 6 & 5 \\ 5 & 6 \end{smallmatrix}\right]$. We have

$$\alpha \cdot \alpha = \begin{bmatrix} 6 & 5 \\ 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 6 & 5 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 6 \cdot 6 + 5 \cdot 5 & 6 \cdot 5 + 5 \cdot 6 \\ 5 \cdot 6 + 6 \cdot 5 & 5 \cdot 5 + 6 \cdot 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, $\alpha \cdot \alpha = \varepsilon$, so that α is a *self-inverse* in $M(\mathbb{Z}_{10})$.

Below are examples of matrices in $M(\mathbb{Z}_{10})$ that do *not* have multiplicative inverses. Again, this is why $M(\mathbb{Z}_{10})$ is *not* a group under multiplication.

Example 7.13. Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 2 & 3 \\ 0 & 0 \end{smallmatrix}\right]$. Then for $\beta \in M(\mathbb{Z}_{10})$ where $\beta = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, we have

$$\alpha \cdot \beta = \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2 \cdot a + 3 \cdot c & 2 \cdot b + 3 \cdot d \\ 0 \cdot a + 0 \cdot c & 0 \cdot b + 0 \cdot d \end{bmatrix} = \begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix},$$

where $x = 2 \cdot a + 3 \cdot c$ and $y = 2 \cdot b + 3 \cdot d$. Then $\alpha \cdot \beta \neq \varepsilon$, since the bottom row of $\alpha \cdot \beta$ will always contain 0 and 0. Thus, α does *not* have a multiplicative inverse in $M(\mathbb{Z}_{10})$.

Example 7.14. Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 6 \end{smallmatrix}\right]$. Then for $\beta \in M(\mathbb{Z}_{10})$ where $\beta = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, we have

$$\alpha \cdot \beta = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 \cdot a + 2 \cdot c & 1 \cdot b + 2 \cdot d \\ 3 \cdot a + 6 \cdot c & 3 \cdot b + 6 \cdot d \end{bmatrix}.$$

Now, assume for contradiction that $\alpha \cdot \beta = \varepsilon$. Then, equating the entries in the first columns of $\alpha \cdot \beta$ and ε, we obtain $1 \cdot a + 2 \cdot c = 1$ and $3 \cdot a + 6 \cdot c = 0$. If we multiply both sides of $1 \cdot a + 2 \cdot c = 1$ by 3, we get $3 \cdot a + 6 \cdot c = 3$. Comparing that with $3 \cdot a + 6 \cdot c = 0$ yields $3 = 0$ in $\mathbb{Z}_{10}$, which is a contradiction. Hence, $\alpha \cdot \beta$ cannot be equal to ε. Thus, α does *not* have a multiplicative inverse in $M(\mathbb{Z}_{10})$.

7.4 Determinant

In the previous section, we found some matrices in $M(\mathbb{Z}_{10})$ that have multiplicative inverses and others that do not. How can we determine whether or not a matrix $\alpha \in M(\mathbb{Z}_{10})$ has a multiplicative inverse? And if it does, how can we find its multiplicative inverse? These questions will be addressed in this section.

Example 7.15. Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ with its entries in $\mathbb{Z}_{10}$. If possible, we'd like to find the multiplicative inverse of α, i.e., a matrix β such that $\alpha \cdot \beta = \varepsilon$ and $\beta \cdot \alpha = \varepsilon$. Consider $\beta \in M(\mathbb{Z}_{10})$ where $\beta = \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$. Multiplying the two, we obtain

$$\begin{aligned} \alpha \cdot \beta &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\ &= \begin{bmatrix} a \cdot d + b \cdot (-c) & a \cdot (-b) + b \cdot a \\ c \cdot d + d \cdot (-c) & c \cdot (-b) + d \cdot a \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix}. \end{aligned}$$

(**Note:** You should verify that $\beta \cdot \alpha$ yields the same result.) The product $\alpha \cdot \beta$ is *almost* equal to the multiplicative identity $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. In fact, if $ad - bc = 1$, then we do have $\alpha \cdot \beta = \varepsilon$. Example 7.11 illustrates this scenario. With $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 4 & 9 \end{smallmatrix}\right]$, we have $ad - bc = 1 \cdot 9 - 2 \cdot 4 = 1$, so that its multiplicative inverse is given by $\beta = \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right] = \left[\begin{smallmatrix} 9 & -2 \\ -4 & 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 9 & 8 \\ 6 & 1 \end{smallmatrix}\right]$, where $-2 = 8$ and $-4 = 6$ in $\mathbb{Z}_{10}$.

Remark. In Example 7.15, it might seem that matrix $\beta = \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$ was pulled out of thin air. But here is the thought process involved in finding it. Given matrix $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, we want to find its multiplicative inverse (if possible), i.e., a matrix $\beta = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]$ such that $\alpha \cdot \beta = \varepsilon$. Let's compute $\alpha \cdot \beta$ and set it equal to ε:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The matrix ε has zeros in two entries, in the first row / second column and in the second row / first column. The zero in the first row / second column is obtained by the following dot product: $a \cdot x + b \cdot z = 0$. What could x and z be? One possibility is to set $x = 0$ and $z = 0$, but then the entry in the second row / second column in the product would be $c \cdot x + d \cdot z = c \cdot 0 + d \cdot 0 = 0$, whereas we'd like to obtain 1 in that entry.

Another option is to set $x = -b$ and $z = a$, which gives us $a \cdot x + b \cdot z = a \cdot (-b) + b \cdot a = 0$. (We could have also set $x = b$ and $z = -a$, but then we would have defined the determinant of α in Definition 7.16 as $bc - ad$ instead.) Likewise, the zero in the second row / first column of ε is obtained by $c \cdot w + d \cdot y = 0$, so we can set $w = d$ and $y = -c$. Therefore, we find that

$$\beta = \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

As we saw in Example 7.15, this matrix β is not quite the inverse of α, but it gives us a good starting point.

Example 7.15 above motivates the following definition.

Definition 7.16 (Determinant)**.** Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ with its entries in $\mathbb{Z}_{10}$. The *determinant* of α, denoted $\det \alpha$, is given by $\det \alpha = ad - bc$. Note that $\det \alpha$ is a number in $\mathbb{Z}_{10}$.

Example 7.17. Let $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$ be the multiplicative identity matrix. Then $\det \varepsilon = 1 \cdot 1 - 0 \cdot 0 = 1$, which is the multiplicative identity element of $\mathbb{Z}_{10}$.

Example 7.18. Let $\alpha \in M(\mathbb{Z}_{10})$, where $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right]$. Then $\det \alpha = 2 \cdot 4 - 1 \cdot 5 = 3$. Motivated by the calculation in Example 7.15, let's define $\beta = \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right] = \left[\begin{smallmatrix} 4 & -1 \\ -5 & 2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 4 & 9 \\ 5 & 2 \end{smallmatrix}\right]$, since $-1 = 9$ and $-5 = 5$ in $\mathbb{Z}_{10}$. From Example 7.15, we know that

$$\alpha \cdot \beta = \begin{bmatrix} \det \alpha & 0 \\ 0 & \det \alpha \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = 3 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

so that $\alpha \cdot \beta = 3 \cdot \varepsilon$, which *almost* equals the multiplicative identity ε. (If only we could "divide" by 3....) Noting that 3 and 7 are multiplicative inverses of each other in $\mathbb{Z}_{10}$, let's consider multiplying α not by β, but by $7 \cdot \beta$. Then we would have

$$\alpha \cdot (7 \cdot \beta) = 7 \cdot (\alpha \cdot \beta) = 7 \cdot (3 \cdot \varepsilon) = (7 \cdot 3) \cdot \varepsilon = 1 \cdot \varepsilon.$$

Thus, $\alpha \cdot (7 \cdot \beta) = \varepsilon$, so that $7 \cdot \beta = 7 \cdot \left[\begin{smallmatrix} 4 & 9 \\ 5 & 2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 8 & 3 \\ 5 & 4 \end{smallmatrix}\right]$ is the multiplicative inverse of α. Let's verify by actually computing the product of α and $7 \cdot \beta$:

$$\alpha \cdot (7 \cdot \beta) = \begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} \cdot \begin{bmatrix} 8 & 3 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 2 \cdot 8 + 1 \cdot 5 & 2 \cdot 3 + 1 \cdot 4 \\ 5 \cdot 8 + 4 \cdot 5 & 5 \cdot 3 + 4 \cdot 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \varepsilon.$$

We can similarly verify that $(7 \cdot \beta) \cdot \alpha = \varepsilon$.

In Example 7.18 above, we used the following property of scalar multiplication: $\alpha \cdot (7 \cdot \beta) = 7 \cdot (\alpha \cdot \beta)$. Here is the generalization, which you will prove in an exercise at the end of the chapter. Intuitively, the theorem says that when multiplying a scalar and two matrices, the scalar k can be moved around freely in the product. But we may *not* change the order of α and β, since matrix multiplication is not commutative.

Theorem 7.19. *Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ and $k \in \mathbb{Z}_{10}$. Then $k \cdot (\alpha \cdot \beta) = (k \cdot \alpha) \cdot \beta = \alpha \cdot (k \cdot \beta)$.*

Revisiting Example 7.18, we note the $\det \alpha = 3$ having a multiplicative inverse in $\mathbb{Z}_{10}$ was the key to α having a multiplicative inverse in $M(\mathbb{Z}_{10})$. This is captured in the following theorem, whose proof is a generalization of the calculation in Example 7.18.

Theorem 7.20. *Let $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$ and $\Delta = \det \alpha \in \mathbb{Z}_{10}$. If Δ has a multiplicative inverse in $\mathbb{Z}_{10}$, then α has a multiplicative inverse in $M(\mathbb{Z}_{10})$. Moreover, the multiplicative inverse of α is given by*

$$\alpha^{-1} = \Delta^{-1} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

where Δ^{-1} is the multiplicative inverse of Δ in $\mathbb{Z}_{10}$.

PROOF. Assume $\Delta = \det \alpha$ has a multiplicative inverse in $\mathbb{Z}_{10}$, namely $\Delta^{-1} \in \mathbb{Z}_{10}$. Then $\alpha^{-1} = \Delta^{-1} \cdot \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$ is a matrix in $M(\mathbb{Z}_{10})$. To prove that α^{-1} is the multiplicative inverse of α, we must show that $\alpha \cdot \alpha^{-1} = \varepsilon$ and $\alpha^{-1} \cdot \alpha = \varepsilon$. We have

$$\begin{aligned}
\alpha \cdot \alpha^{-1} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \left(\Delta^{-1} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \right) \\
&= \Delta^{-1} \cdot \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \right) \\
&= \Delta^{-1} \cdot \begin{bmatrix} \det \alpha & 0 \\ 0 & \det \alpha \end{bmatrix} \\
&= \Delta^{-1} \cdot \left(\Delta \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\
&= (\Delta^{-1} \cdot \Delta) \cdot \varepsilon \\
&= 1 \cdot \varepsilon.
\end{aligned}$$

Therefore, $\alpha \cdot \alpha^{-1} = \varepsilon$. A similar calculation shows that $\alpha^{-1} \cdot \alpha = \varepsilon$. ■

The converse of Theorem 7.20 is also true, as you will prove in an exercise at the end of the chapter. We will state it as a theorem here.

Theorem 7.21 (Converse of Theorem 7.20)**.** *Let $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$ and $\Delta = \det \alpha \in \mathbb{Z}_{10}$. If α has a multiplicative inverse in $M(\mathbb{Z}_{10})$, then Δ has a multiplicative inverse in $\mathbb{Z}_{10}$.*

In practice, the contrapositive of Theorem 7.21 is more useful; namely: *If $\Delta = \det \alpha$ does* not *have a multiplicative inverse in $\mathbb{Z}_{10}$, then α does* not *have a multiplicative inverse in $M(\mathbb{Z}_{10})$.*

Example 7.22 (Example 7.14 revisited)**.** Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 6 \end{smallmatrix}\right]$. Then $\det \alpha = 1 \cdot 6 - 2 \cdot 3 = 0$, which does *not* have a multiplicative inverse in $\mathbb{Z}_{10}$. Thus, α does not have a multiplicative inverse in $M(\mathbb{Z}_{10})$.

Example 7.23. Let $\alpha \in M(\mathbb{Z}_{10})$, where $\alpha = \left[\begin{smallmatrix} 2 & 1 \\ 3 & 4 \end{smallmatrix}\right]$. Then $\det \alpha = 2 \cdot 4 - 1 \cdot 3 = 5$, which does *not* have a multiplicative inverse in $\mathbb{Z}_{10}$. Therefore, α does not have a multiplicative inverse in $M(\mathbb{Z}_{10})$.

We end the chapter with a couple of theorems about determinants, which you will explore (by creating your own examples) and prove in the exercises at the end of this chapter.

Theorem 7.24. *Let $\alpha, \beta \in M(\mathbb{Z}_{10})$. Then* $\det(\alpha \cdot \beta) = \det \alpha \cdot \det \beta$.

In words, Theorem 7.24 says, "The determinant of the product (of matrices) is equal to the product of the determinants." In the expression $\det(\alpha \cdot \beta)$, the multiplication takes place in $M(\mathbb{Z}_{10})$; and in $\det \alpha \cdot \det \beta$, the multiplication occurs in $\mathbb{Z}_{10}$.

Theorem 7.25. *Let $\alpha \in M(\mathbb{Z}_{10})$. If its multiplicative inverse α^{-1} exists, then* $\det \alpha$ *and* $\det(\alpha^{-1})$ *are multiplicative inverses of each other in $\mathbb{Z}_{10}$.*

Remark. Throughout this chapter, we used $\mathbb{Z}_{10}$ as the number system in which the entries of our matrices exist. But there is nothing special about $\mathbb{Z}_{10}$ here. We can replace $\mathbb{Z}_{10}$ with any number system that permits both addition and multiplication. Examples of such number systems (called *rings*, which we'll explore much later in this textbook) include $\mathbb{Z}_m$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and more.

Exercises

1. Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 2 & 4 \\ 5 & 6 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 8 & 3 \\ 1 & 7 \end{smallmatrix}\right]$. Compute each of the following:

   (a) $\alpha + \beta$. (b) $\beta + \alpha$. (c) $\alpha - \beta$. (d) $\beta - \alpha$.

2. **Prove:** Addition in $M(\mathbb{Z}_{10})$ is commutative.

3. Is subtraction in $M(\mathbb{Z}_{10})$ commutative? Why or why not?

4. (a) Find $\alpha, \beta \in M(\mathbb{Z}_{10})$ such that $\alpha \cdot \beta = \beta \cdot \alpha$.
   (b) Find $\alpha, \beta \in M(\mathbb{Z}_{10})$ such that $\alpha \cdot \beta \neq \beta \cdot \alpha$.
   (c) Anita wonders, "Doesn't part (a) prove to us that multiplication in $M(\mathbb{Z}_{10})$ is commutative?" How would you respond to her?

5. Let $\alpha, \beta, \gamma \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$, $\beta = \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]$, and $\gamma = \left[\begin{smallmatrix} 9 & 0 \\ 4 & 7 \end{smallmatrix}\right]$.

   (a) Describe how to compute the product $(\alpha \cdot \beta) \cdot \gamma$. Which matrices are multiplied first?
   (b) Describe how to compute the product $\alpha \cdot (\beta \cdot \gamma)$.
   (c) Compute $(\alpha \cdot \beta) \cdot \gamma$ and $\alpha \cdot (\beta \cdot \gamma)$, and verify that they are equal.

6. **Prove:** Multiplication in $M(\mathbb{Z}_{10})$ is associative. (This exercise is referenced in Sections 10.2 and 10.3.)

7. (a) Let $\alpha \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 4 & 7 \\ 1 & 5 \end{smallmatrix}\right]$. Find the additive inverse of α.
   (b) Repeat part (a), but with $\alpha = \left[\begin{smallmatrix} 8 & 0 \\ 3 & 9 \end{smallmatrix}\right]$.
   (c) Repeat part (a), but with $\alpha = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

8. **Prove:** Every element in $M(\mathbb{Z}_{10})$ has an additive inverse.

9. For each matrix in $M(\mathbb{Z}_{10})$ shown below, determine whether or not it has a multiplicative inverse. If it does, find the multiplicative inverse. If it does not, explain why not.

   (a) $\begin{bmatrix} 1 & 5 \\ 2 & 7 \end{bmatrix}$. (b) $\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$. (c) $\begin{bmatrix} 4 & 4 \\ 3 & 6 \end{bmatrix}$. (d) $\begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix}$.

10. Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 4 \\ 3 & 2 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 2 & 1 \\ 3 & 4 \end{smallmatrix}\right]$. Without using the determinant, explain why each of these matrices does not have a multiplicative inverse. (See Example 7.14.)

11. Let $\alpha \in M(\mathbb{Z}_{10})$ whose second row is a constant multiple of the first row. Explain why α does not have a multiplicative inverse.

    Note: Consider the matrix $\alpha = \left[\begin{smallmatrix} 1 & 4 \\ 3 & 2 \end{smallmatrix}\right]$ in Exercise #10, for instance. The second row (3 and 2) is obtained by multiplying each entry the first row (1 and 4) by the same number 3.

12. Find five different self-inverses in $M(\mathbb{Z}_{10})$, i.e., matrices $\alpha \in M(\mathbb{Z}_{10})$ such that $\alpha \cdot \alpha = \varepsilon$.

13. (a) Verify that $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$ has a multiplicative inverse.
    (b) For an element $\alpha \in M(\mathbb{Z}_{10})$ with a multiplicative inverse, the *order* of α refers to the smallest positive exponent n such that

$$\alpha^n = \underbrace{\alpha \cdot \alpha \cdot \cdots \cdot \alpha}_{n \text{ copies}} = \varepsilon,$$

    where $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is the multiplicative identity matrix. Find the order of $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$.

14. **Prove:** Let $\alpha \in M(\mathbb{Z}_{10})$. If there exists a positive integer n such that $\alpha^n = \varepsilon$, then α has a multiplicative inverse in $M(\mathbb{Z}_{10})$.

15. Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 1 & 5 \\ 2 & 7 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 5 & 2 \\ 3 & 1 \end{smallmatrix}\right]$. Compute each of the following products/inverses:

    (a) $(\alpha \cdot \beta)^{-1}$; i.e., first find the product $\alpha \cdot \beta$ and then find its inverse.
    (b) $\alpha^{-1} \cdot \beta^{-1}$; i.e., first find the inverses α^{-1} and β^{-1} and then multiply them.
    (c) $\beta^{-1} \cdot \alpha^{-1}$; i.e., this is similar to $\alpha^{-1} \cdot \beta^{-1}$ but in different order.

    You should find that $(\alpha \cdot \beta)^{-1} \neq \alpha^{-1} \cdot \beta^{-1}$, but $(\alpha \cdot \beta)^{-1} = \beta^{-1} \cdot \alpha^{-1}$. (This exercise is referenced in Section 8.2.)

16. Draw an analogy between the equation $(\alpha \cdot \beta)^{-1} = \beta^{-1} \cdot \alpha^{-1}$ (from Exercise #15) and the act of putting on and taking off your socks and shoes.

17. Theorem 7.19 states: Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ and $k \in \mathbb{Z}_{10}$. Then $k \cdot (\alpha \cdot \beta) = (k \cdot \alpha) \cdot \beta = \alpha \cdot (k \cdot \beta)$.

    (a) Create an example that illustrates this theorem.

    (b) Prove the theorem.

18. Theorem 7.21 states: Let $\alpha = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$ and $\Delta = \det \alpha \in \mathbb{Z}_{10}$. If α has a multiplicative inverse in $M(\mathbb{Z}_{10})$, then Δ has a multiplicative inverse in $\mathbb{Z}_{10}$.

    (a) Create an example that illustrates this theorem.

    (b) Prove the theorem.

19. Theorem 7.24 states: Let $\alpha, \beta \in M(\mathbb{Z}_{10})$. Then $\det(\alpha \cdot \beta) = \det \alpha \cdot \det \beta$.

    (a) Create an example that illustrates this theorem.

    (b) Prove the theorem.

20. Theorem 7.25 states: Let $\alpha \in M(\mathbb{Z}_{10})$. If its multiplicative inverse α^{-1} exists, then $\det \alpha$ and $\det(\alpha^{-1})$ are multiplicative inverses of each other in $\mathbb{Z}_{10}$.

    (a) Create an example that illustrates this theorem.

    (b) Prove the theorem.

21. In a linear algebra course, students typically work with matrices whose entries are real numbers. They learn the following theorem: A matrix α has a multiplicative inverse if and only if $\det \alpha \neq 0$. Discuss how this linear algebra theorem is consistent with our Theorems 7.20 and 7.21.

22. (a) Find the number of matrices in $M(\mathbb{Z}_{10})$.

    (b) Find the number of matrices in $M(\mathbb{Z}_7)$.

    (c) Find the number of matrices in $M(\mathbb{Z}_{12})$.

    (d) Find the number of matrices in $M(\mathbb{Z}_m)$.

23. **(Challenge)** Find the number of matrices in $M(\mathbb{Z}_{10})$ that have multiplicative inverses.

Unit III: Introduction to Groups

Chapter 8 will formally define the notion of a *group*, which, we hope, will feel familiar to you. The next several chapters are devoted to articulating and proving properties that apply to *all* groups, thereby engaging in *abstraction* or the process of extracting structural similarities that arise in different scenarios. (Don't worry, we'll still keep referring back to concrete examples!) For instance, we will see in Chapter 9 that every group with three elements is essentially the same and that a four-element group can be categorized into one of two types. Subgroups, which are subsets of groups that happen to be groups themselves, are introduced in Chapter 11. Two chapters are set aside for *cyclic groups*, an important type of groups that is generated by a single element. For example, the additive group of integers $\mathbb{Z}$ is cyclic, because every integer can be expressed as a sum of 1's or -1's. Our excursion into cyclic groups will also foreshadow the notion of an *isomorphism*, i.e., what it means for two groups to be "the same."

Here is a taste of what you'll be able to accomplish in this unit:

- Learn how to prove that a certain group is commutative.
- Given a group element g, determine which integer exponents k satisfy $g^k = \varepsilon$.
- Prove that a subgroup of a cyclic group is also cyclic.

8

Introduction to Groups

In the last unit, we explored several important examples of groups:

- The set of integers $\mathbb{Z}$ with addition.
- The number system $\mathbb{Z}_7$ with addition.
- The set of symmetries of a square D_4 with composition.
- The set of permutations S_3 with composition.
- The set of 2×2 matrices $M(\mathbb{Z}_{10})$ with addition.

These examples all share common *group properties*, which we will formalize in this chapter. We will begin proving statements not just about $\mathbb{Z}_7$ or S_3, but about a general group. Thus, our theorems will apply to *all* groups. But, of course, concrete examples will continue to ground us in our work.

8.1 Definition of a "group"

Group is one of the two fundamental structures that we will study in this textbook. (*Ring* is the other.) To review the group properties, we revisit some computations in S_3, i.e., the set of all permutations of $\{1, 2, 3\}$.

Example 8.1. Let $\sigma, \gamma, \varepsilon \in S_3$ be given by

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

In Chapter 6, we saw that $\varepsilon \circ \sigma = \sigma$ and $\sigma \circ \varepsilon = \sigma$ and that these relationships hold when σ is replaced by any element of S_3. In other words, composing with the *identity* element ε is analogous to $0 + a = a$ and $a + 0 = a$ in $\mathbb{Z}$. We also found that $\sigma \circ \gamma = \varepsilon$ and $\gamma \circ \sigma = \varepsilon$, so that σ and γ are *inverses* of each other. This is analogous to $a + (-a) = 0$ and $(-a) + a = 0$ in $\mathbb{Z}$.

Now we articulate the four group properties, i.e., the special features that make the set S_3 together with the operation $\circ$ (composition) into a group:

1. S_3 is closed under $\circ$; i.e., if $\sigma, \tau \in S_3$, then $\sigma \circ \tau \in S_3$.
2. The operation $\circ$ is associative; i.e., $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$ for all σ, τ, $\mu \in S_3$.
3. S_3 contains an identity element ε such that $\varepsilon \circ \alpha = \alpha$ and $\alpha \circ \varepsilon = \alpha$ for all $\alpha \in S_3$.
4. Each element $\sigma \in S_3$ has an inverse $\sigma^{-1} \in S_3$ such that $\sigma \circ \sigma^{-1} = \varepsilon$ and $\sigma^{-1} \circ \sigma = \varepsilon$.

Remark. Rather than "S_3 with composition is a group," we often say "S_3 is a group *under* composition." These are just two different ways of saying the same thing.

Finally, here is the long-awaited definition.

Definition 8.2 (Group)**.** When a set with an associated operation satisfies the four properties above, we call that pair (the set and the operation) a *group*.

In addition to S_3 under composition, we have seen several examples of groups already. Shown below is a partial list. When naming a group, we should indicate both the set (e.g., S_3) and the operation (e.g., $\circ$).

- D_4 (the set of symmetries of a square) under composition.
- $\mathbb{Z}$ (the set of integers) under addition.
- $\mathbb{Z}_{35}$ under addition.
- U_{35} under multiplication.
- $M(\mathbb{Z}_{10})$ under addition.

Remark. As a default, we will view the operation of a group as *multiplication*. Given a generic group G and its elements a, $b \in G$, we will refer to their *product* $a \cdot b$, or more simply ab. Even for specific groups like S_3 and D_4, we will often write $\alpha\beta$ instead of $\alpha \circ \beta$ and refer to it as a "product," rather than a "composition." The one exception to this rule is when we know that the group operation is addition, such as with $\mathbb{Z}$ or $\mathbb{Z}_{35}$. In those cases, we will continue to use the *additive notation*. Given a, $b \in \mathbb{Z}$, for instance, we will write $a+b$ and refer to it as a "sum." Moreover, a group under addition is always assumed to be commutative. Thus, in such a group, we have $a + b = b + a$ for any pair of elements a and b.

Below, we will explain why $\mathbb{Z}_{35}$ is a group under addition. This group is fairly large (with 35 elements), so constructing an addition table to check the group properties does not seem feasible. Thus, we will take a more general approach. An upside is that the explanations given below can be applied to $\mathbb{Z}_m$ as well.

Example 8.3. We verify the group properties for $\mathbb{Z}_{35}$ under addition:

(1) $\mathbb{Z}_{35}$ is closed under addition. Given a, $b \in \mathbb{Z}_{35}$, their sum $a + b$ is still in $\mathbb{Z}_{35}$. If the sum exceeds 34, we can reduce it by subtracting 35. For example, $26 + 17 = 43 = 43 - 35 = 8$ in $\mathbb{Z}_{35}$.

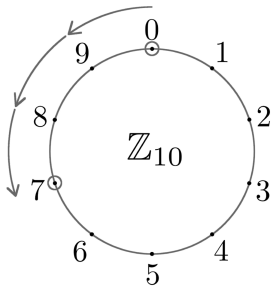
(2) Addition in $\mathbb{Z}_{35}$ is associative; i.e., $(a+b)+c = a+(b+c)$ for all $a, b, c \in \mathbb{Z}_{35}$. After all, addition in $\mathbb{Z}_{35}$ is based on addition in $\mathbb{Z}$.

(3) The element $0 \in \mathbb{Z}_{35}$ is the *additive* identity, because $0+a = a$ and $a+0 = a$ for all $a \in \mathbb{Z}_{35}$.

(4) Each element $a \in \mathbb{Z}_{35}$ has an *additive* inverse $35-a$, which is also in $\mathbb{Z}_{35}$. For example, the additive inverse of 12 is $35-12 = 23$. Note that $12+23 = 0$ and $23+12 = 0$.

Remark. We saw that $12+23 = 0$ in $\mathbb{Z}_{35}$. Thus, the additive inverse of 12 is 23, and we write $-12 = 23$. Also, the additive inverse of 23 is 12, and we write $-23 = 12$. When the operation is addition, the additive inverse of x is denoted by $-x$, rather than by x^{-1}, which is the notation for multiplicative inverse.

Example 8.4. Just like $\mathbb{Z}_{35}$, the set $\mathbb{Z}_{10}$ is a group under addition. In $\mathbb{Z}_{10}$, we have $3+7 = 0$. Thus, the additive inverse of 3 is 7, and we write $-3 = 7$. Note how this is consistent with our computation on the $\mathbb{Z}_{10}$ clock. To find -3 on the $\mathbb{Z}_{10}$ clock, we first view -3 as $0-3$. Then, we start at 0 on the $\mathbb{Z}_{10}$ clock and move 3 units *counterclockwise*. We land on 7, so that $-3 = 7$ in $\mathbb{Z}_{10}$.



Example 8.5. Let $\mathbb{R}$ be the set of all real numbers. Below, we verify that $\mathbb{R}$ is a group under addition. Also, we will simply assume that addition in $\mathbb{R}$ is closed and associative.

(1) The set $\mathbb{R}$ is closed under addition.

(2) Addition in $\mathbb{R}$ is associative.

(3) The additive identity element is $0 \in \mathbb{R}$, as $0+a = a$ and $a+0 = a$ for all $a \in \mathbb{R}$.

(4) Every $a \in \mathbb{R}$ has an additive inverse $-a \in \mathbb{R}$ such that $a+(-a) = 0$ and $(-a)+a = 0$.

Next, we will show that U_{35} is a group under multiplication. Recall that

$$U_{35} = \{a \in \mathbb{Z}_{35} \mid a \text{ has a multiplicative inverse in } \mathbb{Z}_{35}\}.$$

For example, $4 \cdot 9 = 1$ in $\mathbb{Z}_{35}$ so that 4 and 9 are multiplicative inverses of each other in $\mathbb{Z}_{35}$. Thus $4, 9 \in U_{35}$. On the other hand, $5 \notin U_{35}$, because $5 \cdot x \neq 1$ for any $x \in \mathbb{Z}_{35}$.

Using Theorem 4.19 (i.e., $a \in U_{35}$ if and only if $\gcd(a, 35) = 1$), we can quickly find its elements:

$$U_{35} = \{1, 2, 3, 4, 6, 8, 9, 11, 12, 13, 16, 17, 18, 19, 22, 23, 24, 26, 27, 29, 31, 32, 33, 34\}.$$

This group is large enough (with twenty-four elements) that constructing a multiplication table does not seem feasible. As we did with $\mathbb{Z}_{35}$, we will take a more general approach. Again, the upside is that the explanations given below can be applied to U_m as well.

Example 8.6. We consider U_{35} under multiplication:

(1) U_{35} is closed under multiplication. Given $a, b \in U_{35}$, their product ab is still in U_{35}. This is proved (by you!) in Chapter 4, Exercise #17.

(2) Multiplication in U_{35} is associative; i.e., $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in U_{35}$. After all, multiplication in U_{35} is based on multiplication in $\mathbb{Z}$.

(3) The element 1 is the *multiplicative* identity, because $1 \cdot a = a$ and $a \cdot 1 = a$ for all $a \in U_{35}$. Moreover, 1 is in U_{35}, because it has a multiplicative inverse, namely itself.

(4) By the definition of U_{35}, each element $a \in U_{35}$ has a *multiplicative* inverse a^{-1} such that $a \cdot a^{-1} = 1$ and $a^{-1} \cdot a = 1$. We must also verify that $a^{-1} \in U_{35}$, i.e., that a^{-1} has a multiplicative inverse; it does, since a is a multiplicative inverse of a^{-1}.

Example 8.7. Let $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$. For instance, $3 \cdot \frac{1}{3} = 1$ so that 3, $\frac{1}{3} \in \mathbb{R}^*$. Similarly, $\pi \in \mathbb{R}^*$, since it has a multiplicative inverse, namely $\frac{1}{\pi}$. In fact, if $a \in \mathbb{R}$ with $a \neq 0$, then a has a multiplicative inverse, namely $\frac{1}{a}$. Thus, the set $\mathbb{R}^*$ contains all *non-zero* real numbers. In an exercise at the end of the chapter, you will explain why $\mathbb{R}^*$ is a group under multiplication.

Example 8.8. The sets $\mathbb{Z}$, $\mathbb{Z}_7$, $\mathbb{Z}_{35}$, and $\mathbb{R}$ are all groups under addition. However, $\mathbb{Z}$ is not a group under multiplication, because $0 \in \mathbb{Z}$ does not have a multiplicative inverse. The same argument (i.e., 0 does not have a multiplicative inverse) can also be used in $\mathbb{Z}_7$, $\mathbb{Z}_{35}$, and $\mathbb{R}$ to explain why none of them is a group under multiplication.

8.2 Essential properties of a group

In this section, we investigate essential properties that are shared by all groups.

Uniqueness of identity and inverse. In $\mathbb{Z}_{35}$, which is a group under addition, the element 0 is the *only* additive identity. In other words, there is no other element besides $\varepsilon = 0$ that satisfies $\varepsilon + a = a$ and $a + \varepsilon = a$ for all $a \in \mathbb{Z}_{35}$. Likewise, in the multiplicative group U_{35}, the element 1 is the *only* multiplicative identity, as there is no other element ε that satisfies $\varepsilon \cdot a = a$ and $a \cdot \varepsilon = a$ for all $a \in U_{35}$. Each group we've seen so far has contained only one identity element. Not surprisingly, this is true in any group.

Theorem 8.9. *A group has a unique* (*i.e., only one*) *identity element.*

PROOF. Assume a group has two identity elements, namely ε_1 and ε_2. We will show that these two elements must be the same. Since ε_1 is an identity element, it keeps ε_2 unchanged upon multiplication: $\varepsilon_1 \cdot \varepsilon_2 = \varepsilon_2$. But ε_2 is also an identity element and thus keeps ε_1 unchanged upon multiplication: $\varepsilon_1 \cdot \varepsilon_2 = \varepsilon_1$. Therefore, $\varepsilon_1 = \varepsilon_2$, since they are both equal to $\varepsilon_1 \cdot \varepsilon_2$. ■

As we remarked earlier, we will view the operation of a group as multiplication as a default. This convention can be seen in the proof above. We also identify a proof technique that was used in this proof.

Proof know-how. To show that there is a unique element (with some property), assume that there are two such elements. Then show that those two elements must be the same.

The element $12 \in \mathbb{Z}_{35}$ has an additive inverse, namely 23, because $12 + 23 = 0$ and $23 + 12 = 0$ in $\mathbb{Z}_{35}$. Moreover, 23 is the *only* additive inverse of 12. There is no other element $b \in \mathbb{Z}_{35}$ such that $12 + b = 0$ and $b + 12 = 0$ in $\mathbb{Z}_{35}$. Likewise, the element $8 \in U_{35}$ has a *unique* (i.e., only one) multiplicative inverse, namely 22. There is no other element $b \in U_{35}$ such that $8 \cdot b = 1$ and $b \cdot 8 = 1$. This is also true in any group.

Theorem 8.10. *Let G be a group. Each element $g \in G$ has a unique inverse in G.*

The proof of this theorem is left for you as an exercise at the end of this chapter. Because of this theorem, we can refer to *the* inverse of g and refer to it unambiguously as g^{-1}.

Socks-shoes property. With $\sigma, \tau \in S_3$, we saw that $(\sigma\tau)^{-1} \neq \sigma^{-1}\tau^{-1}$, but rather $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$ (see Chapter 6, Exercise #6). We observed the same phenomenon with 2×2 matrices (see Chapter 7, Exercise #15). We made an analogy between $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$ and the act of putting on and taking off your socks and shoes:

> Think of σ as putting on your socks and τ as putting on your shoes. Then $\sigma\tau$ denotes first putting on your socks, followed by putting on your shoes. Now, $(\sigma\tau)^{-1}$ denotes the process of *undoing* the steps taken in $\sigma\tau$. And to undo $\sigma\tau$, you must first take off your shoes (denoted by τ^{-1}) and then take off your socks (i.e., σ^{-1}). Thus, we have $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$.

The proof of Theorem 8.11 below gives an algebraic justification of the socks-shoes property. The relationship we prove is $(ab)^{-1} = b^{-1}a^{-1}$, which translates to "the multiplicative inverse of ab is $b^{-1}a^{-1}$." To show this, we multiply ab by $b^{-1}a^{-1}$ and verify that the product equals the identity element ε.

Proof know-how. Suppose x and y are elements of a group. To show that $x^{-1} = y$ ("the multiplicative inverse of x is y"), we verify that $x \cdot y = \varepsilon$ and $y \cdot x = \varepsilon$.

Theorem 8.11 (Socks-shoes property)**.** *Let a and b be elements of a group. Then* $(ab)^{-1} = b^{-1}a^{-1}$.

PROOF. We will show that ab multiplied by $b^{-1}a^{-1}$ (on either side) yields the identity element. First, on the right: $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a\,\varepsilon\,a^{-1} = aa^{-1} = \varepsilon$. Here,

we used the fact that a^{-1} and b^{-1} are inverses of a and b, respectively, so that $aa^{-1} = \varepsilon$ and $bb^{-1} = \varepsilon$. Next, we multiply on the left: $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}\,\varepsilon\, b = b^{-1}b = \varepsilon$. Thus, $b^{-1}a^{-1}$ is the inverse of ab, as desired. ■

Remark. In *any* group, we have $(ab)^{-1} = b^{-1}a^{-1}$. If a group happens to be commutative, then we can further say that $b^{-1}a^{-1} = a^{-1}b^{-1}$, and hence $(ab)^{-1} = a^{-1}b^{-1}$.

Example 8.12. Let g be a group element. We will use the Proof know-how technique employed in the proof of Theorem 8.11 to show that $(g^{-1})^{-1} = g$. First, the equation $(g^{-1})^{-1} = g$ translates to "the multiplicative inverse of g^{-1} is g." To show this, we must verify $g^{-1} \cdot g = \varepsilon$ and $g \cdot g^{-1} = \varepsilon$. But these are true by the definition of g^{-1}. Therefore, $(g^{-1})^{-1} = g$ as desired.

Laws of exponents. Let g be an element of a group. Below, we will discuss the meaning of the expression g^n for different values of the integer exponent n. First, we define g^0 to be the identity element of the group; i.e., $g^0 = \varepsilon$.

For $n > 0$, we have

$$g^n = \underbrace{g \cdot g \cdot \cdots \cdot g}_{n \text{ copies}}.$$

For instance, $g^1 = g$ and $g^2 = g \cdot g$. Consider g^3. On the surface, this seems straightforward: $g^3 = g \cdot g \cdot g$. However, an operation of a group is a *binary* operation; i.e., we multiply/add/combine only two elements at a time. Thus, the expression $g \cdot g \cdot g$ actually means either $(g \cdot g) \cdot g$ or $g \cdot (g \cdot g)$. Fortunately, $(g \cdot g) \cdot g$ and $g \cdot (g \cdot g)$ are equal due to the associative law, so there is no ambiguity in writing $g^3 = g \cdot g \cdot g$.

Similarly, $g^4 = g \cdot g \cdot g \cdot g$ could mean any one of these expressions:

$$(gg)(gg),\ (g(gg))g,\ ((gg)g)g,\ g((gg)g),\ g(g(gg)),$$

where the multiplication symbol $\cdot$ was removed for simplicity. Again, we use the associative law to show that these expressions are all equal. For example, we have $g\big((gg)g\big) = \big(g(gg)\big)g$, where the larger parentheses were used to emphasize the regrouping that took place. Thus, just as with $g^3 = g \cdot g \cdot g$, there is no ambiguity when we write $g^4 = g \cdot g \cdot g \cdot g$, so we will continue to write expressions such as these.

Next, we demonstrate that the familiar laws of exponents hold with group elements:

- $g^{m+n} = g^m \cdot g^n$.
- $(g^m)^n = g^{mn}$.

We will prove $g^{m+n} = g^m \cdot g^n$ for the case when m and n are positive. But first, let's create an example.

Example 8.13. With $m = 2$ and $n = 3$, we have $m + n = 5$ so that

$$g^{2+3} = \underbrace{g \cdot g \cdot g \cdot g \cdot g}_{5 \text{ copies}} = \underbrace{(g \cdot g)}_{2 \text{ copies}} \cdot \underbrace{(g \cdot g \cdot g)}_{3 \text{ copies}} = g^2 \cdot g^3.$$

Theorem 8.14. *Let g be an element of a group, and let $m, n \in \mathbb{Z}$ with $m, n > 0$. Then $g^{m+n} = g^m \cdot g^n$.*

PROOF. We have

$$g^{m+n} = \underbrace{g \cdot g \cdot \cdots \cdot g}_{m+n \text{ copies}} = \underbrace{(g \cdot g \cdot \cdots \cdot g)}_{m \text{ copies}} \cdot \underbrace{(g \cdot g \cdot \cdots \cdot g)}_{n \text{ copies}} = g^m \cdot g^n. \qquad \blacksquare$$

The proof of $(g^m)^n = g^{mn}$ (with $m, n > 0$) is left for you as an exercise at the end of this chapter.

What if the exponents are negative? Let's first address a more fundamental question: What does an expression like g^{-5} mean in a group setting?

Example 8.15 (Meaning of g^{-5})**.** Let g be an element of a group. Then g^{-1} denotes the multiplicative inverse of g, but what might g^{-5} mean? Two potential candidates are:

- $g^{-5} = (g^{-1})^5$; i.e., first invert g and then multiply g^{-1} by itself five times.
- $g^{-5} = (g^5)^{-1}$; i.e., first multiply g by itself five times and then invert g^5.

Not too surprisingly, the two interpretations $(g^{-1})^5$ and $(g^5)^{-1}$ are equivalent. Applying the socks-shoes property five times (imagine four layers of socks, followed by shoes), we have

$$(g^5)^{-1} = (g \cdot g \cdot g \cdot g \cdot g)^{-1} = g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} \cdot g^{-1} = (g^{-1})^5.$$

Thus the expression g^{-5} may be viewed, without ambiguity, as either $(g^{-1})^5$ or $(g^5)^{-1}$.

Rather than prove the laws $g^{m+n} = g^m \cdot g^n$ and $(g^m)^n = g^{mn}$ when m or n (or possibly both) is negative, we will examine several examples below and in the exercises at the end of the chapter.

Example 8.16. Let g be a group element. We have $g^{-2} = (g^{-1})^2 = g^{-1}g^{-1}$, so that

$$g^{-2} \cdot g^5 = (g^{-1}g^{-1})(ggggg) = g^{-1}(g^{-1}g)gggg = g^{-1}\varepsilon gggg = (g^{-1}g)ggg = \varepsilon ggg = g^3.$$

Here, we used the fact that $g^{-1}g = \varepsilon$, since g^{-1} is the inverse of g. Thus, $g^{-2}g^5 = g^3$, which verifies the law $g^{m+n} = g^m \cdot g^n$ when $m = -2$ and $n = 5$.

Example 8.17. Let g be a group element. We have

$$(g^2)^{-3} = ((g^2)^3)^{-1} = \left(g^2 \cdot g^2 \cdot g^2\right)^{-1} = (g^6)^{-1} = g^{-6}.$$

Thus, $(g^2)^{-3} = g^{-6}$, which verifies the law $(g^m)^n = g^{mn}$ when $m = 2$ and $n = -3$.

Cancellation laws. Because every element of a group has an inverse, we have the cancellation laws. The theorem and proof below are more or less identical to those provided for S_n in Chapter 6, Exercise #14. The right cancellation law (i.e., if $ba = ca$, then $b = c$) can be proved similarly, and it is left for you as an exercise.

Theorem 8.18 (Left cancellation)**.** *Let a, b, c be elements of a group. If $ab = ac$, then $b = c$.*

PROOF. Assume $ab = ac$ in a group. Multiply both sides of the equation *on the left* by a^{-1} to obtain $a^{-1}(ab) = a^{-1}(ac)$. Using the associative law gives $(a^{-1}a)b = (a^{-1}a)c$. Since $a^{-1}a = \varepsilon$, we get $\varepsilon b = \varepsilon c$. Finally, ε keeps all elements of the group unchanged; i.e., $\varepsilon b = b$ and $\varepsilon c = c$. Thus, $b = c$ as desired. $\blacksquare$

Example 8.19 (Non-example)**.** In $\mathbb{Z}_{10}$, we have $2 \cdot 6 = 2 \cdot 1$, even though $6 \neq 1$. This does *not* violate Theorem 8.18, because $\mathbb{Z}_{10}$ is *not* a group under multiplication (it is a group under addition, though). In fact, 2 does not have a multiplicative inverse; i.e., 2^{-1} does not exist in $\mathbb{Z}_{10}$.

Example 8.20 (Non-example)**.** Define $\sigma, \tau, \mu \in S_3$, respectively, by

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

Then we have $\sigma\tau = \mu\sigma$. (We'll leave the calculation up to you.) It is incorrect to cancel σ from both sides of the equation and conclude $\tau = \mu$, since the cancellation must occur on the *same* side of each expression.

8.3 Proving that a group is commutative

Definition 8.21 (Commutative group)**.** A group G is *commutative* if $ab = ba$ for all a, $b \in G$. Otherwise, G is *non-commutative*.

Example 8.22. The groups $\mathbb{Z}_{35}$ (under addition) and U_{35} (under multiplication) are commutative. After all, addition and multiplication in these groups are based on addition and multiplication in $\mathbb{Z}$. We've also seen that the groups S_3 and D_4 (both under composition) are non-commutative.

If we're given a particular group, say $\mathbb{Z}_{35}$ or U_{35}, it's pretty straightforward to determine whether or not the group is commutative. But oftentimes, we're given a *property* of a group, and we must *prove* that the group is commutative. Here's one such example.

Theorem 8.23. *Let G be a group. If $(ab)^2 = a^2b^2$ for all $a, b \in G$, then G is commutative.*

PROOF. Assume that $(ab)^2 = a^2b^2$ for all $a, b \in G$. Let $a, b \in G$. We must show that $ab = ba$. From the given property, we know that $(ab)^2 = a^2b^2$. Therefore, $(ab)(ab) = (aa)(bb)$, so that $a(b(ab)) = a(a(bb))$. Left cancellation gives $b(ab) = a(bb)$, and the associative law yields $(ba)b = (ab)b$. Further canceling on the right gives $ba = ab$, as desired. ■

Proof know-how. To prove that a group G is commutative, let $a, b \in G$. Then show that $ab = ba$.

We'll start the proof of the next theorem and leave it for you to complete as an exercise.

Theorem 8.24. *Let G be a group. If $(ab)^{-1} = a^{-1}b^{-1}$ for all $a, b \in G$, then G is commutative.*

PROOF. Assume that $(ab)^{-1} = a^{-1}b^{-1}$ for all $a, b \in G$. Let $a, b \in G$. We must show that $ab = ba$. From the given property, we know that $(ab)^{-1} = a^{-1}b^{-1}$. Taking the inverse of both sides of this equation, we obtain $((ab)^{-1})^{-1} = (a^{-1}b^{-1})^{-1}$.

[We'll leave the rest of the details to you as an exercise.]

Therefore, $ab = ba$ as desired. ■

Proof know-how. If two expressions involving group elements are equal, then their inverses are equal as well. In the proof above, we have the equality $(ab)^{-1} = a^{-1}b^{-1}$; i.e., the group elements $(ab)^{-1}$ and $a^{-1}b^{-1}$ are equal. Then their inverses are equal to each other, and thus we obtain $((ab)^{-1})^{-1} = (a^{-1}b^{-1})^{-1}$.

8.4 Non-associative operations

Since we'll be working with groups (and later with rings), the operations we'll study will be associative; e.g., $(a+b)+c = a+(b+c)$ for all $a, b, c \in \mathbb{Z}_{35}$. Throughout this chapter, we saw the value of the associative law. For instance, we can write $g^4 = g \cdot g \cdot g \cdot g$ without ambiguity due to the associative law. As a fun exercise, you might find examples and proofs in this chapter where associativity was used, both explicitly and implicitly.

Every operation that we have seen thus far has been associative, so it's easy to take the property for granted. But non-associative operations exist. A familiar example is *subtraction* in $\mathbb{Z}$. We have

$$(15-2)-7 = 13-7 = 6 \quad \text{and} \quad 15-(2-7) = 15-(-5) = 20,$$

so that $(15-2)-7 \neq 15-(2-7)$.

Another example is *exponentiation*. Define $a \star b = a^b$ for $a, b \in \mathbb{N}$. We then have

$$(2 \star 3) \star 4 = (2^3) \star 4 = 8^4 = 4{,}096.$$

But $2 \star (3 \star 4) = 2 \star (3^4) = 2 \star 81 = 2^{81}$, which is *much* bigger than 4,096. Thus,

$$(2 \star 3) \star 4 \neq 2 \star (3 \star 4).$$

8.5 Direct product

When we plot points on the coordinate plane, we work with an *ordered pair* such as $(3, 4)$, which has 3 as its x-coordinate and 4 as its y-coordinate. The point $(3, 4)$ is different from the point $(4, 3)$, which has 4 and 3 as its x- and y-coordinates, respectively. If we restrict ourselves to integer coordinates, some other points on the plane include $(7, -8)$, $(-22, 10)$, $(0, 0)$, and so on. In the context of abstract algebra, we may view these points as elements of the *direct product* $\mathbb{Z} \times \mathbb{Z}$. Below are some examples.

Example 8.25. Consider the groups D_4 under composition and $\mathbb{Z}_{10}$ under addition. Their *direct product*, denoted $D_4 \times \mathbb{Z}_{10}$, is the set containing ordered pairs (σ, a) where $\sigma \in D_4$ and $a \in \mathbb{Z}_{10}$. Some elements of $D_4 \times \mathbb{Z}_{10}$ include $(r_{90}, 4)$, $(h, 7)$, and $(\varepsilon, 0)$. We "multiply" two elements of $D_4 \times \mathbb{Z}_{10}$ by *componentwise operation*, i.e., by "multiplying" each coordinate separately using the respective operations of D_4 and $\mathbb{Z}_{10}$. For instance, we have $(r_{90}, 4) \cdot (h, 7) = (r_{90} \circ h, 4+7) = (d', 1)$. This computation suggests $D_4 \times \mathbb{Z}_{10}$ is closed, because D_4 and $\mathbb{Z}_{10}$ are closed themselves. Likewise, the fact that the operations of D_4 and $\mathbb{Z}_{10}$ are associative implies that the componentwise operation of $D_4 \times \mathbb{Z}_{10}$ is also associative. The identity element of $D_4 \times \mathbb{Z}_{10}$ is $(\varepsilon, 0)$, because

$$(\varepsilon, 0) \cdot (\sigma, a) = (\varepsilon \circ \sigma, 0 + a) = (\sigma, a) \text{ and } (\sigma, a) \cdot (\varepsilon, 0) = (\sigma \circ \varepsilon, a + 0) = (\sigma, a)$$

for all $(\sigma, a) \in D_4 \times \mathbb{Z}_{10}$. Lastly, each element of $D_4 \times \mathbb{Z}_{10}$ has an inverse in $D_4 \times \mathbb{Z}_{10}$. For instance,

$$(r_{90}, 4) \cdot (r_{270}, 6) = (r_{90} \circ r_{270}, 4 + 6) = (\varepsilon, 0)$$

and

$$(r_{270}, 6) \cdot (r_{90}, 4) = (r_{270} \circ r_{90}, 6 + 4) = (\varepsilon, 0)$$

so that $(r_{90}, 4)$ and $(r_{270}, 6)$ are inverses of each other. Symbolically, we write $(r_{90}, 4)^{-1} = (r_{270}, 6)$ and $(r_{270}, 6)^{-1} = (r_{90}, 4)$. Therefore, we conclude that $D_4 \times \mathbb{Z}_{10}$ is a group under componentwise operation. A general proof showing that a direct product of two groups is a group is left for you as an exercise.

When working with direct products, we use the multiplication notation as a default. (See Example 8.25.) An exception to this rule is when *both* components are additive groups, as shown below.

Example 8.26. Consider the direct product $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$. Since both components are additive groups, the operation of $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$ is componentwise addition. Given $(6, 4)$, $(5, 10) \in \mathbb{Z}_{10} \times \mathbb{Z}_{12}$, we have

$$(6, 4) + (5, 10) = (6 + 5, 4 + 10) = (1, 2),$$

where the sums $6 + 5$ and $4 + 10$ are computed in $\mathbb{Z}_{10}$ and $\mathbb{Z}_{12}$, respectively. Closure and associativity in $\mathbb{Z}_{10}$ and $\mathbb{Z}_{12}$ ensure that these properties hold in $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$. The additive identity of $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$ is $(0, 0)$, since

$$(0, 0) + (a, b) = (0 + a, 0 + b) = (a, b) \text{ and } (a, b) + (0, 0) = (a + 0, b + 0) = (a, b)$$

for all $(a, b) \in \mathbb{Z}_{10} \times \mathbb{Z}_{12}$. Each element of $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$ has an additive inverse in $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$. For instance,

$$(3, 8) + (7, 4) = (3 + 7, 8 + 4) = (0, 0) \text{ and } (7, 4) + (3, 8) = (7 + 3, 4 + 8) = (0, 0)$$

so that $(3, 8)$ and $(7, 4)$ are additive inverses of each other. We write $-(3, 8) = (7, 4)$ and $-(7, 4) = (3, 8)$. Thus, the direct product $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$ is a group under componentwise addition.

Here is a general definition of the direct product.

Definition 8.27 (Direct product)**.** Let G and H be groups with operations $*_G$ and $*_H$, respectively. Their *direct product* is the set $G \times H = \{(g, h) \mid g \in G, h \in H\}$ with the *componentwise operation* where

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 *_G g_2, h_1 *_H h_2)$$

for all $(g_1, h_1), (g_2, h_2) \in G \times H$.

The following theorem is left for you to prove as an exercise.

Theorem 8.28. *If G and H are groups, then the direct product $G \times H$ is a group under the componentwise operation.*

Exercises

1. For each element $a \in \mathbb{Z}_{35}$, find its additive inverse.

   (a) $a = 23$. (b) $a = 14$. (c) $a = 21$. (d) $a = 0$. (e) $a = 34$.

2. **Prove:** $\mathbb{Z}_m$ is closed under addition.

3. Determine if each set with the given operation forms a group. If it does, verify the group properties. If it doesn't, describe *all* group properties that fail.

   (a) The set of *positive* integers with addition.
   (b) The set $2\mathbb{Z}$ with addition.
   (c) The set $\{\varepsilon, \sigma, \tau, \mu\} \subseteq S_4$ with composition. Here, ε is the identity element, and
   $$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \quad \mu = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}.$$
   (d) The set $\{0, 2, 4, 6\}$ with addition in $\mathbb{Z}_8$.
   (e) The set $\{4, 8, 12, 16\}$ with multiplication in $\mathbb{Z}_{20}$. (This exercise is referenced in Chapter 9, Exercise #14e.)
   (f) The set $\{q \in \mathbb{Q} \mid q > 0\}$ with multiplication. (Recall that $\mathbb{Q}$ is the set of rational numbers.)

4. Recall that $U_{10} = \{a \in \mathbb{Z}_{10} \mid a \text{ has a multiplicative inverse in } \mathbb{Z}_{10}\}$.

   (a) Find the elements of U_{10}.
   (b) Construct a multiplication table for U_{10}.
   (c) Use the table to verify the group properties for U_{10}.
   Note: You may simply assume that multiplication in U_{10} is associative.
   (d) Is U_{10} commutative or non-commutative?

5. Recall that $\mathbb{R}$ is the set of all real numbers and that $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$.

   (a) Explain why $\mathbb{R}^*$ is a group under multiplication. (This exercise is referenced in Example 9.2.)
   Hint: See Example 8.7 for a description of the elements in $\mathbb{R}^*$.
   (b) Anita says, "$\mathbb{R}$ is kind of like $\mathbb{Z}_7$. Both are groups under addition, and they're *almost* groups under multiplication." What might she mean? Name other additive groups that are like $\mathbb{R}$ and $\mathbb{Z}_7$.

6. We say that 1 is a *generator* of the additive group $\mathbb{Z}_{12}$, because its sums give all elements in the group, as shown:
   $$1 = 1, \quad 1 + 1 = 2, \quad 1 + 1 + 1 = 3, \quad 1 + 1 + 1 + 1 = 4, \ \ldots, \quad \underbrace{1 + 1 + \cdots + 1}_{12 \text{ terms}} = 0.$$
   (a) Find all the generators of $\mathbb{Z}_{12}$.
   (b) Find all the generators of $\mathbb{Z}_7$.
   (c) Find all the generators of $\mathbb{Z}_{15}$.

(d) Find all the generators of $\mathbb{Z}_{20}$.

(e) What conjectures do you have?

(This exercise is referenced in Section 9.5 and in Chapter 13.)

7. The *quaternion group* contains eight elements:

$$G = \left\{ \pm\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \pm\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \pm\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \right\},$$

where the entries of these matrices are in $\mathbb{Z}_3$ (thus, $-1 = 2$). The operation is matrix multiplication.

(a) Verify that G satisfies the group properties.

(b) Is G commutative or non-commutative? How do you know?

(c) Find the order of each element in G. Any conjectures?
Note: The *order* of $\alpha \in G$ refers to the smallest positive exponent n such that

$$\alpha^n = \underbrace{\alpha \cdot \alpha \cdot \cdots \cdot \alpha}_{n \text{ copies}} = \varepsilon.$$

8. Consider the set

$$G = \left\{ \begin{bmatrix} a & 0 \\ b & a \end{bmatrix} \,\middle|\, a, b \in \mathbb{R},\ a \neq 0 \right\}.$$

Prove: G is a group under matrix multiplication.

Note: You may assume that matrix multiplication is associative. So, you must show the following:

- G is closed under matrix multiplication.
- The identity matrix ε is in G.
- For each $\alpha \in G$, its inverse α^{-1} is also in G.

9. Consider the following subsets of $U_7 = \{1, 2, 3, 4, 5, 6\}$: $E = \{1, 6\}$, $S = \{2, 5\}$, and $T = \{3, 4\}$. Define the *set product* of, say, S and T by

$$S \cdot T = \{s \cdot t \mid s \in S,\ t \in T\}$$

where the multiplication $s \cdot t$ is done in $\mathbb{Z}_7$. Thus, we have

$$S \cdot T = \{2, 5\} \cdot \{3, 4\} = \{2 \cdot 3,\ 2 \cdot 4,\ 5 \cdot 3,\ 5 \cdot 4\} = \{6, 1, 1, 6\} = \{1, 6\};$$

i.e., $S \cdot T$ contains all possible products of an element in S with an element in T. Note that sets do *not* have repetition, and hence $\{6, 1, 1, 6\} = \{1, 6\}$.

(a) Let $G = \{E, S, T\}$. Compute the table for G, using set multiplication.

(b) Use the table created to verify the group properties for G.

(c) Is G commutative or non-commutative?

(This exercise is referenced in Section 21.1 and Chapter 21, Exercise #3.)

10. Let g be an element of a group. Use the associative law to show that these expressions are all equal:

$$(gg)(gg),\ (g(gg))g,\ ((gg)g)g,\ g((gg)g),\ g(g(gg)).$$

11. **Prove:** Let g be an element of a group, and let $m, n \in \mathbb{Z}$ with $m, n > 0$. Then $(g^m)^n = g^{mn}$.

12. Let g be an element of a group, and let $m, n \in \mathbb{Z}$.
    (a) Suppose $m = 0$. Explain why $g^{m+n} = g^m \cdot g^n$.
    (b) Suppose $m = 0$. Explain why $(g^m)^n = g^{mn}$.
    (c) Suppose $n = 0$. Explain why $(g^m)^n = g^{mn}$.

13. Let g be an element of a group. Verify each of the following:
    (a) $(g^{-2})^3 = g^{-6}$.
    (b) $(g^{-2})^{-3} = g^6$.
    (c) $g^{-2} \cdot g^{-3} = g^{-5}$.

14. In S_3, compute σ^{-5} if
$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

15. **(Right cancellation) Prove:** Let a, b, c be elements of a group. If $ba = ca$, then $b = c$.

16. Prove Theorem 8.10.
    Note: Avoid using the notation g^{-1} in your proof, since g^{-1} refers to the unique inverse of g; and we're trying to *prove* that g has a unique inverse.

17. Complete the proof of Theorem 8.24.

18. **Prove:** Let G be a group. If $g^{-1} = g$ for all $g \in G$, then G is commutative.

19. **Prove:** Let G be a group. If $g^2 = \varepsilon$ for all $g \in G$, then G is commutative.

20. Create a counterexample to show that division in $\mathbb{R}^*$ is not associative.

21. Prove Theorem 8.28.

22. Consider the direct product $U_{10} \times U_{12}$.
    (a) How many elements does $U_{10} \times U_{12}$ contain? Explain how you know.
    (b) Compute the products $(3, 5) \cdot (9, 7)$ and $(7, 1) \cdot (3, 11)$.
    (c) Find the multiplicative identity element of $U_{10} \times U_{12}$.
    (d) Find the multiplicative inverse of each of these elements: $(3, 5)$, $(9, 7)$, $(7, 1)$, $(3, 11)$.
    (e) Find all self-inverses in $U_{10} \times U_{12}$.

23. Consider the direct product $\mathbb{Z}_{15} \times \mathbb{Z}_8$.
    (a) How many elements does $\mathbb{Z}_{15} \times \mathbb{Z}_8$ contain? Explain how you know.
    (b) Compute the sums $(4, 6) + (13, 5)$ and $(6, 6) + (7, 7)$.
    (c) Find the additive identity element of $\mathbb{Z}_{15} \times \mathbb{Z}_8$.
    (d) Find the additive inverse of each of these elements: $(4, 6)$, $(13, 5)$, $(6, 6)$, $(7, 7)$.
    (e) Find all self-inverses in $\mathbb{Z}_{15} \times \mathbb{Z}_8$.

24. Determine if each of these direct products is commutative or non-commutative.

    (a) $D_4 \times \mathbb{Z}_{10}$. (b) $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$. (c) $U_{10} \times U_{12}$. (d) $S_3 \times S_3$.

25. **Prove:** Let G and H be groups. Then G and H are commutative if and only if $G \times H$ is commutative.

9

Groups of Small Size

In Chapter 8, we formally began our work with *abstraction*, the powerful mathematical process of extracting structural similarities that arise in different scenarios. In this chapter, we will continue with this work, deriving and proving statements that apply to *all* groups. (Of course, we will still rely on concrete examples to motivate and ground our work.)

In particular, we will explore groups of small size, ranging from one to four elements. For instance, we will show that all groups with three elements are essentially the same. Along the way, we will preview a couple of concepts that we will study in depth later on, namely, *group isomorphism* and *cyclic groups*.

9.1 Smallest group

We begin with the question: What is the smallest possible group, with the fewest number of elements? By definition, a group must contain at least one element, namely the identity. Is it possible for a group to contain *only* the identity element? Let's verify with an example.

Example 9.1. Consider the subset $\{0\}$ of the set of integers $\mathbb{Z}$. Here is the addition table for $\{0\}$:

$+$	0
0	0

Let's verify the group properties for $\{0\}$ under addition:

(1) $\{0\}$ is closed under addition. The only possible sum is $0 + 0 = 0$, and this sum is an element of $\{0\}$.

(2) The associative law holds, i.e., $(0 + 0) + 0 = 0 + (0 + 0)$, since both sides are equal to 0.

(3) $\{0\}$ has an identity element 0 that keeps all elements in $\{0\}$ unchanged. Observe that $0 + a = a$ and $a + 0 = a$ for all $a \in \{0\}$ (since $a \in \{0\}$ means $a = 0$ in this scenario).

(4) Every element in $\{0\}$ has an additive inverse. Specifically, $0 + 0 = 0$ implies that 0 is a self-inverse.

Example 9.2. The set $\{1\}$ is a group under multiplication, which you will verify in an exercise at the end of the chapter. Here, we view 1 as a real number and $\{1\}$ as a subset of

$$\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}.$$

It is possible to view 1 as an integer and $\{1\}$ as a subset of $\mathbb{Z}$. However, because the operation of $\{1\}$ is multiplication and $\mathbb{R}^*$ is a group under multiplication (see Chapter 8, Exercise #5a), we prefer to keep the operation the same between the two sets $\{1\}$ and $\mathbb{R}^*$. We will say more about this when we study *subgroups* in depth in Chapter 11.

Example 9.3. Let $\varepsilon \in S_3$; i.e.,

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

Then the set $\{\varepsilon\}$ is a group under composition. The same is true when $\varepsilon \in D_4$; i.e., ε is the "motion" that does not move the square.

Example 9.4. The subset $\left\{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right\}$ of the set of matrices $M(\mathbb{Z}_{10})$ is a group under matrix addition.

The above examples demonstrate how, while the elements and the operations may differ, all one-element groups are essentially the same. Here is the generalization. The smallest possible group has just the identity element. Below is the group table for any one-element group $G = \{\varepsilon\}$:

$\cdot$	ε
ε	ε

In this group, the only possible product is $\varepsilon \cdot \varepsilon = \varepsilon$. (Recall from Chapter 8 that we use the multiplicative notation as a default.) Thus the group is certainly closed. The associative property holds, i.e., $(\varepsilon \cdot \varepsilon) \cdot \varepsilon = \varepsilon \cdot (\varepsilon \cdot \varepsilon)$, since both sides are equal to ε. The identity element exists, since $\varepsilon \cdot a = a$ and $a \cdot \varepsilon = a$ for all $a \in G$. And ε is a self-inverse, so every element in G has an inverse.

9.2 Groups with two elements

Next, we consider groups with two elements. As we'll soon see, all of these groups are essentially the same.

Example 9.5. Consider the following groups: $\{1, -1\}$ under multiplication (where we view 1 and -1 as real numbers), $\mathbb{Z}_2 = \{0, 1\}$ under addition, and $\{\varepsilon, r_{180}\}$ under $\circ$ in D_4.

Here are their group tables:

Table for $\{1, -1\}$:

$\cdot$	1	-1
1	1	-1
-1	-1	1

Table for $\mathbb{Z}_2$:

$+$	0	1
0	0	1
1	1	0

Table for $\{\varepsilon, r_{180}\}$:

$\circ$	ε	r_{180}
ε	ε	r_{180}
r_{180}	r_{180}	ε

The three tables look essentially the same, each involving the identity and the non-identity element. In each table, the main diagonal entries (i.e., the top left and the bottom right) contain the identity element, and the other two entries (i.e., the top right and the bottom left) contain the non-identity element.

Next, we will explain why every two-element group must have a table that looks like the ones that we saw in Example 9.5. Let $G = \{\varepsilon, g\}$ be a two-element group, where ε and g are the identity and non-identity elements, respectively. Then its table must have the following form:

	ε	g
ε	ε	g
g	g	ε

Since ε is the identity element, we have $\varepsilon a = a$ and $a\varepsilon = a$ for all $a \in G$. Thus we have $\varepsilon\varepsilon = \varepsilon$, $\varepsilon g = g$, and $g\varepsilon = g$, and so we can immediately complete three entries in the table. The following theorem addresses the remaining entry gg. (**Note:** We present it as a theorem, not necessarily because it's such an important result, but because it utilizes a helpful proof technique that we highlight afterwards.)

Theorem 9.6. *Let $G = \{\varepsilon, g\}$ be a two-element group. Then $gg = \varepsilon$.*

PROOF. The product gg must be an element of G, because the group is closed. Thus, either $gg = \varepsilon$ or $gg = g$. Suppose for contradiction that $gg = g$. We also have $g = g\varepsilon$, since ε is the identity element. Combining $gg = g$ and $g = g\varepsilon$, we obtain $gg = g\varepsilon$. Then left cancellation yields $g = \varepsilon$. However, $g = \varepsilon$ is a contradiction, since g is the non-identity element of the group. Therefore, we must have $gg = \varepsilon$. ■

Proof know-how. A key step in the above proof was to rewrite g as $g\varepsilon$, which allowed us to apply left cancellation. This "inserting the identity" method comes in various forms (as we will see in future proofs) and can be a helpful proof-writing technique.

Thus, we come to the same conclusion that we made about one-element groups: While the elements and operations may differ, all two-element groups are essentially the same.

9.3 Groups with three elements

In this section, we will show that all three-element groups are essentially the same.

Example 9.7. Consider the following groups:

- $\mathbb{Z}_3 = \{0, 1, 2\}$ under addition.
- $\{\varepsilon, r_{120}, r_{240}\}$ under the operation $\circ$ in D_3.
- $\{\varepsilon, \sigma, \tau\}$ under the operation $\circ$ in S_3, where $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ and $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

Their tables are shown below. Again, notice how they look essentially the same. Specifically, the identity elements (0, ε, and ε), the first non-identity elements (1, r_{120}, and σ), and the second non-identity elements (2, r_{240}, and τ) can be found in the same locations in the three tables.

Table for $\mathbb{Z}_3$:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Table for $\{\varepsilon, r_{120}, r_{240}\}$:

$\circ$	ε	r_{120}	r_{240}
ε	ε	r_{120}	r_{240}
r_{120}	r_{120}	r_{240}	ε
r_{240}	r_{240}	ε	r_{120}

Table for $\{\varepsilon, \sigma, \tau\}$:

$\circ$	ε	σ	τ
ε	ε	σ	τ
σ	σ	τ	ε
τ	τ	ε	σ

We will now explain why every three-element group must have a table like the ones in Example 9.7. Thus, let $G = \{\varepsilon, a, b\}$ be a three-element group, with identity ε and two non-identity elements a and b. We will build its table in phases. In the first phase, we can complete the first row and column of the table, since $\varepsilon x = x$ and $x\varepsilon = x$ for all $x \in G$.

In the second phase, we note that $ab = \varepsilon$ and $ba = \varepsilon$, which you will prove in an exercise at the end of the chapter. We state them as a theorem here, since we will use those results when completing the table.

Theorem 9.8. *Let $G = \{\varepsilon, a, b\}$ be a three-element group. Then $ab = \varepsilon$ and $ba = \varepsilon$.*

To complete the table, we must show that $aa = b$ and $bb = a$. The following theorem addresses $aa = b$. (Notice the similarity to the proof of Theorem 9.6, including the use of the "inserting the identity" method.) The proof of $bb = a$ is left for you as an exercise.

Theorem 9.9. *Let $G = \{\varepsilon, a, b\}$ be a three-element group. Then $aa = b$.*

PROOF. The product aa must be an element of G, because the group is closed. Thus, either $aa = \varepsilon$, $aa = a$, or $aa = b$. Suppose for contradiction that $aa = \varepsilon$. By Theorem 9.8, we have $ab = \varepsilon$. Hence, $aa = ab$, since both sides are equal to ε. Then left cancellation yields $a = b$, which is a contradiction. Thus, $aa \neq \varepsilon$.

Again for contradiction, suppose $aa = a$. We also have $a = a\varepsilon$, since ε is the identity element. Combining $aa = a$ and $a = a\varepsilon$, we obtain $aa = a\varepsilon$. Then left cancellation yields $a = \varepsilon$, which is also a contradiction. Thus, $aa \neq a$. Therefore, the only remaining option must be true, namely $aa = b$. ∎

Phase 1:

	ε	a	b
ε	ε	a	b
a	a		
b	b		

Phase 2:

	ε	a	b
ε	ε	a	b
a	a		ε
b	b	ε	

Complete table:

	ε	a	b
ε	ε	a	b
a	a	b	ε
b	b	ε	a

Notice how the table for a general three-element group $G = \{\varepsilon, a, b\}$ is just like the tables we saw in Example 9.7. Once again, while the elements and operations may differ, all three-element groups are essentially the same.

9.4 Sudoku property

As we create and study many group tables, we notice a useful property that is shared by *all* tables, regardless of the number of elements in the group: In each row or column of the table, every element of G shows up exactly once. We will call this the *Sudoku property*, named after a popular number-placement game that is typically played on a 9×9 grid. One of the rules of Sudoku (the game) is that each row and each column of the grid must contain all the digits from 1 to 9 exactly once.

Theorem 9.10 (Sudoku property)**.** *Let G be a group. In each row or column of its group table, every element of G shows up exactly once.*

To prove this theorem, we will fix an *arbitrary* row of the table. Showing that the property holds for this row will imply that the property holds for *all* rows. The argument for columns is left for you as an exercise.

PROOF. Let $G = \{\varepsilon, a, b, \ldots, g, \ldots\}$, possibly infinite, and consider the group table for G:

	ε	a	b	$\cdots$	g	$\cdots$	y	$\cdots$
ε								
a								
$\vdots$								
g	g	ga	gb	$\cdots$	gg	$\cdots$	x	$\cdots$
$\vdots$								

We will fix an arbitrary row g and show the following:

(1) The elements in this row are all different.

(2) Every $x \in G$ appears in this row.

Note that claim (1) says every element of G shows up *at most* once in row g. Meanwhile, claim (2) says that every element of G shows up *at least* once in row g. Taken together, the two claims will imply that every element of G shows up *exactly* once in row g.

For claim (1), we must show the following: If $a \neq b$, then $ga \neq gb$. We will show its contrapositive: If $ga = gb$, then $a = b$. This follows immediately from left cancellation.

For claim (2), we must show that any $x \in G$ appears in row g. In other words, we must find an element $y \in G$ such that $x = gy$. Let $y = g^{-1}x$, which is in G because g^{-1}, $x \in G$ and G is closed. Then $gy = g(g^{-1}x) = (gg^{-1})x = \varepsilon x = x$ so that $x = gy$.

Similar arguments show that claims (1) and (2) hold for any *column* of the table as well. ■

Proof know-how. In the above proof, we had to find an element $y \in G$ such that $x = gy$. We showed that $y = g^{-1}x$ is the desired element. How did we come up with an expression for y? Our goal was $x = gy$, so we worked backwards and solved this equation for y by left-multiplying both sides by g^{-1}. This "working backwards" process is part of the scratch work that is done *before* we write the proof. It must *not* be included in the proof itself for a couple of reasons. First, our goal was to prove that $x = gy$, and we must not assume what we're trying to prove. Second, the logical arguments in the proof must flow forwards, not backwards.

9.5 Groups with four elements

In Section 9.3, we saw that all three-element groups are essentially the same (and likewise for all one-element groups and all two-element groups). More formally, we say that all three-element groups are *isomorphic* to each other. Later in the textbook, we will study isomorphic groups in much more depth. But in this section, we will see that not all four-element groups are the same.

Example 9.11. Consider the following groups:

- $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ under addition.
- $C = \{1, i, -1, -i\}$ under multiplication. (Here, i is the complex number $i = \sqrt{-1}$, where $i^2 = -1$.)
- $U_8 = \{1, 3, 5, 7\}$ under multiplication modulo 8.

Here are their group tables:

Table for $\mathbb{Z}_4$:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Table for C:

$*$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

Table for U_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

Note how $\mathbb{Z}_4$ and C both have exactly two self-inverses. In group $\mathbb{Z}_4$, elements 0 and 2 are self-inverses; in group C, elements 1 and -1 are self-inverses. Moreover, their group tables look the same. Meanwhile, *every* element of U_8 is a self-inverse and its table looks different from the tables of the other two groups.

Recall from Chapter 8, Exercise #6: We say that 1 is a *generator* of the additive group $\mathbb{Z}_{12}$, because its sums give all elements in the group, as shown:

$$1 = 1,\ 1 + 1 = 2,\ 1 + 1 + 1 = 3,\ 1 + 1 + 1 + 1 = 4,\ \ldots,\ \underbrace{1 + 1 + \cdots + 1}_{12 \text{ terms}} = 0.$$

Let's see if we can find a generator for each four-element group in Example 9.11 above:

- $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ under addition. We have $1 = 1$, $1 + 1 = 2$, $1 + 1 + 1 = 3$, $1 + 1 + 1 + 1 = 0$. Thus, 1 is a generator of $\mathbb{Z}_4$. You should verify on your own that 3 is also a generator of $\mathbb{Z}_4$.
- $C = \{1, i, -1, -i\}$ under multiplication. Since the operation is multiplication, a generator of C is an element whose *products* give all the elements in the group. We have $i = i$, $i \cdot i = -1$, $i \cdot i \cdot i = -i$, $i \cdot i \cdot i \cdot i = 1$. Thus, i is a generator of C. You should verify on your own that $-i$ is also a generator.
- $U_8 = \{1, 3, 5, 7\}$ under multiplication modulo 8. Again, the operation is multiplication. To determine if an element $g \in U_8$ is a generator, we multiply g by itself (or compute powers of g) and see if we obtain all the elements in U_8. Let's see:
  - $1^k = 1$ for all exponents k.
  - $3^1 = 3$, $3^2 = 1$, $3^3 = 3$, $3^4 = 1, \ldots$. The powers of 3 alternate between 3 and 1.
  - $5^1 = 5$, $5^2 = 1$, $5^3 = 5$, $5^4 = 1, \ldots$. The powers of 5 alternate between 5 and 1.
  - $7^1 = 7$, $7^2 = 1$, $7^3 = 7$, $7^4 = 1, \ldots$. The powers of 7 alternate between 7 and 1.

  Therefore, none of the elements in U_8 is a generator.

Below, we give a definition and examples of a *cyclic group*, a topic that we will explore in much more depth later in the textbook.

Definition 9.12 (Cyclic group)**.** A group is said to be *cyclic* if it has a generator.

Example 9.13. $\mathbb{Z}_4$ and $C = \{1, i, -1, -i\}$ are cyclic groups, but U_8 is not cyclic.

You will prove in an exercise that every four-element group resembles either of the two types given in this section: the cyclic group like $\mathbb{Z}_4$ (or C) or the non-cyclic group like U_8. In other words, while the elements and operations may differ, every four-element group is isomorphic to either $\mathbb{Z}_4$ or U_8.

Exercises

1. Let $1 \in \mathbb{R}^*$. Create a multiplication table for $\{1\}$ and verify that it is a group. (See Example 9.2.)
2. Let $\varepsilon \in S_3$. Create a composition table for $\{\varepsilon\}$ and verify that it is a group.
3. In the problems below, view 0 and 1 as integers.
   (a) Create a multiplication table for $\{0\}$ and verify that it is a group under multiplication.
   (b) Is $\{1\}$ a group under addition? Why or why not?

4. (a) Consider $6 \in \mathbb{Z}_{10}$. Create a multiplication table for $\{6\}$ (the multiplication is done modulo 10) and verify that it is a group under multiplication.
   (b) Find other pairs α and m where $\alpha \in \mathbb{Z}_m$ and $\{\alpha\}$ is a group under multiplication modulo m.

5. In Example 9.5, we saw that the subset $\{\varepsilon, r_{180}\} \subseteq D_4$ is a group under composition. For each non-identity element $g \in D_4$, determine if $\{\varepsilon, g\}$ is a group.

6. Find all two-element subsets of S_3 that are a group under composition.

7. Prove Theorem 9.8 *without* using the Sudoku property (i.e., Theorem 9.10).

8. **Prove:** Let $G = \{\varepsilon, a, b\}$ be a three-element group. Then $bb = a$.

9. In the proof of the Sudoku property (i.e., Theorem 9.10), prove that claims (1) and (2) are true for an arbitrary *column* of the group table.

10. Let $G = \{\varepsilon, a, b\}$ be a three-element group. We can complete the first row and column of the table, since $\varepsilon x = x$ and $x\varepsilon = x$ for all $x \in G$. Complete the rest of the table using the Sudoku property.

	ε	a	b
ε	ε	a	b
a	a		
b	b		

11. For each three-element group below, determine if it's cyclic or non-cyclic. If it's cyclic, find a generator.
   (a) $\mathbb{Z}_3 = \{0, 1, 2\}$ under addition.
   (b) $\{\varepsilon, r_{120}, r_{240}\}$ under the operation $\circ$ in D_3.
   (c) $\{\varepsilon, \sigma, \tau\}$ under the operation $\circ$ in S_3, where

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

12. Prove that every three-element group is cyclic.

13. This exercise refers to the groups described in Example 9.11.
   (a) Verify that 3 is a generator for $\mathbb{Z}_4$ under addition.
   (b) Verify that $-i$ is a generator for $C = \{1, i, -1, -i\}$ under multiplication.

14. For each four-element group below, determine if it's cyclic or non-cyclic. If it's cyclic, find a generator.
   (a) $\{\varepsilon, r_{180}, h, v\} \subseteq D_4$.
   (b) $U_5 = \{1, 2, 3, 4\}$ with multiplication in $\mathbb{Z}_5$.
   (c) $U_{12} = \{1, 5, 7, 11\}$ with multiplication in $\mathbb{Z}_{12}$.
   (d) $\{0, 3, 6, 9\}$ with addition in $\mathbb{Z}_{12}$.
   (e) $\{4, 8, 12, 16\}$ with multiplication in $\mathbb{Z}_{20}$. (See Chapter 8, Exercise #3e.)

15. Consider the direct product $\mathbb{Z}_2 \times \mathbb{Z}_2$.
    (a) Write down the elements of $\mathbb{Z}_2 \times \mathbb{Z}_2$. (**Note:** There should be four elements.)
    (b) Determine if $\mathbb{Z}_2 \times \mathbb{Z}_2$ is cyclic or non-cyclic. If it's cyclic, find a generator.
16. Repeat Exercise #15 with the direct product $U_3 \times U_6$.
17. Find your own four-element group. (Be sure to specify both the set and the operation.) Determine if the group is cyclic or non-cyclic. If it's cyclic, find all generators.
18. Find all four-element subsets of D_4 that are a group under composition. Determine if each group is cyclic or non-cyclic. If it's cyclic, find all generators. (Group table for D_4 is in Appendix B.)
19. Consider the groups U_7, U_9, U_{14}, U_{18}, and $\mathbb{Z}_2 \times \mathbb{Z}_3$.
    (a) Verify that each of these groups has six elements.
    (b) Determine if each group is cyclic or non-cyclic. If it's cyclic, find a generator.
    (c) Is every six-element group cyclic? Why or why not?
20. Explain why the group $D_3 = \{\varepsilon, r_{120}, r_{240}, v, d, d'\}$ is non-cyclic. (Group table for D_3 is in Chapter 5, Exercise #8a.)
21. Is the group D_4 cyclic or non-cyclic? Explain how you know. (Group table for D_4 is in Appendix B.)
22. Prove that every four-element group resembles either of the two types given in Section 9.5.
23. **(Challenge)** Categorize all groups with 5 elements; with 6 elements; and with 7 elements.

10

Matrix Groups

Consider the following word analogy: Kitten is to cat as puppy is to _____. The answer is "dog," since kitten and puppy are young versions of cat and dog, respectively. As in this example, analogies can help us better understand the *relationships* between the words involved. Analogies can also aid in making sense of *new ideas* (or words, in this case). For instance, to teach the meaning of the word "Lilliputian" to someone, we might use the following analogy: Big is to enormous as little is to Lilliputian.

Analogies are a useful tool when learning mathematics, too. In this chapter, we will consider the following mathematical analogy: $\mathbb{Z}_{10}$ is to U_{10} as $M(\mathbb{Z}_{10})$ is to _____. This will help us deepen our understanding of the relationship between the additive group $\mathbb{Z}_{10}$ and the multiplicative group U_{10}. The analogy will also lead us to a new group involving 2×2 matrices.

10.1 Groups $\mathbb{Z}_{10}$ and U_{10}

Consider $\mathbb{Z}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. When working with a particular group, we use only one operation. However, $\mathbb{Z}_{10}$ admits both addition *and* multiplication; i.e., $\mathbb{Z}_{10}$ is closed under both operations. In fact, $\mathbb{Z}_{10}$ is an example of a structure called a *ring*, which we will study much later in the textbook.

We've seen that $\mathbb{Z}_{10}$ is a group under addition. (See Example 8.3 for a justification.) But is $\mathbb{Z}_{10}$ also a group under multiplication? Let's check the group properties:

(1) $\mathbb{Z}_{10}$ is closed under multiplication.

(2) Multiplication in $\mathbb{Z}_{10}$ is associative.

(3) The multiplicative identity element is $1 \in \mathbb{Z}_{10}$, as $1 \cdot a = a$ and $a \cdot 1 = a$ for all $a \in \mathbb{Z}_{10}$.

(4) **Not** every $a \in \mathbb{Z}_{10}$ has a multiplicative inverse $a^{-1} \in \mathbb{Z}_{10}$ such that $a \cdot a^{-1} = 1$ and $a^{-1} \cdot a = 1$.

The last group property about inverses fails for $\mathbb{Z}_{10}$ with multiplication, as shown in the example below.

Example 10.1. We will determine which elements of $\mathbb{Z}_{10}$ do *not* have multiplicative inverses. For instance, $0 \cdot x = 1$ is not possible in $\mathbb{Z}_{10}$, since $0 \cdot x = 0$ for all $x \in \mathbb{Z}_{10}$. Thus, 0^{-1} does not exist in $\mathbb{Z}_{10}$. Likewise, $5 \cdot x = 1$ is not possible in $\mathbb{Z}_{10}$, since $5 \cdot x = 0$ or 5 for all $x \in \mathbb{Z}_{10}$. Hence 5^{-1} does not exist in $\mathbb{Z}_{10}$. Similar arguments show that 2^{-1}, 4^{-1}, 6^{-1}, and 8^{-1} also do not exist in $\mathbb{Z}_{10}$. We'll leave these arguments for you to make in an exercise at the end of the chapter.

Note that remaining elements of $\mathbb{Z}_{10}$, namely 1, 3, 7, and 9, do have multiplicative inverses: $1 \cdot 1 = 1$, $3 \cdot 7 = 1$, and $9 \cdot 9 = 1$ in $\mathbb{Z}_{10}$.

We salvage this situation by considering the subset $U_{10} = \{a \in \mathbb{Z}_{10} \mid a \text{ has a multiplicative inverse}\}$. In other words, we remove from $\mathbb{Z}_{10}$ the elements without multiplicative inverses. From Example 10.1, we have $U_{10} = \{1, 3, 7, 9\}$. We can verify that U_{10} is a multiplicative group by building its multiplication table, which you will do in an exercise at the end of the chapter. More generally, we saw in Example 8.6 that U_m is a group under multiplication. We will repeat those explanations here, since we will be referring back to them when we discuss matrix groups in the next section.

(1) U_m is closed under multiplication. This is proved (by you!) in Chapter 4, Exercise #17.

(2) Multiplication in U_m is associative.

(3) The element 1 is the multiplicative identity. Moreover, 1 is in U_m, because it has a multiplicative inverse, namely itself.

(4) By the definition of U_m, each element $a \in U_m$ has a multiplicative inverse a^{-1} such that $a \cdot a^{-1} = 1$ and $a^{-1} \cdot a = 1$. We must also verify that $a^{-1} \in U_m$, i.e., that a^{-1} has a multiplicative inverse; it does, since a is a multiplicative inverse of a^{-1}.

Remark. Both $\mathbb{Z}_{10}$ under addition and U_{10} under multiplication are *commutative* groups.

10.2 Groups $M(\mathbb{Z}_{10})$ and $G(\mathbb{Z}_{10})$

Recall that $M(\mathbb{Z}_{10})$ is the set of 2×2 matrices with entries in $\mathbb{Z}_{10}$. Moreover, $M(\mathbb{Z}_{10})$ admits both addition *and* multiplication; i.e., $M(\mathbb{Z}_{10})$ is closed under both operations. We'll later see that $M(\mathbb{Z}_{10})$ is also a ring.

Example 10.2. We'll briefly review how to add and multiply in $M(\mathbb{Z}_{10})$. For more details, see Section 7.1. Let $\alpha, \beta \in M(\mathbb{Z}_{10})$, where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]$. Then

$$\alpha + \beta = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 0 & 2 \end{bmatrix}$$

and

$$\alpha \cdot \beta = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{bmatrix} = \begin{bmatrix} 9 & 2 \\ 3 & 0 \end{bmatrix}.$$

Similar to $\mathbb{Z}_{10}$, we've seen that $M(\mathbb{Z}_{10})$ is a group under addition. (See Section 7.2 for a justification.) We now ask the same question that we posed earlier about $\mathbb{Z}_{10}$;

namely, is $M(\mathbb{Z}_{10})$ also a group under multiplication? We check the group properties:

(1) $M(\mathbb{Z}_{10})$ is closed under multiplication.

(2) Multiplication in $M(\mathbb{Z}_{10})$ is associative. (See Chapter 7, Exercise #6.)

(3) The multiplicative identity element is $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, as $\varepsilon \cdot \alpha = \alpha$ and $\alpha \cdot \varepsilon = \alpha$ for all $\alpha \in M(\mathbb{Z}_{10})$.

(4) **Not** every $\alpha \in M(\mathbb{Z}_{10})$ has a multiplicative inverse $\alpha^{-1} \in M(\mathbb{Z}_{10})$ such that $\alpha \cdot \alpha^{-1} = \varepsilon$ and $\alpha^{-1} \cdot \alpha = \varepsilon$.

Similar to $\mathbb{Z}_{10}$, the last group property about inverses fails for $M(\mathbb{Z}_{10})$ with multiplication. For instance, $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right] \cdot \beta = \varepsilon$ is not possible in $M(\mathbb{Z}_{10})$, since $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right] \cdot \beta = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ for all $\beta \in M(\mathbb{Z}_{10})$. Thus, $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]^{-1}$ does not exist in $M(\mathbb{Z}_{10})$. We also saw in Examples 7.13 and 7.14 that the matrices $\left[\begin{smallmatrix} 2 & 3 \\ 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & 2 \\ 3 & 6 \end{smallmatrix}\right]$, respectively, do not have multiplicative inverses in $M(\mathbb{Z}_{10})$.

To salvage this situation, we define the subset

$$G(\mathbb{Z}_{10}) = \{\alpha \in M(\mathbb{Z}_{10}) \mid \alpha \text{ has a multiplicative inverse}\}.$$

In other words, we remove from $M(\mathbb{Z}_{10})$ the matrices without multiplicative inverses. The set $G(\mathbb{Z}_{10})$ is often called the *general linear group*, hence the use of the letter G in its name. To verify that $G(\mathbb{Z}_{10})$ is a group under multiplication, we begin by proving closure.

Theorem 10.3. *$G(\mathbb{Z}_{10})$ is closed under multiplication.*

PROOF. Assume $\alpha, \beta \in G(\mathbb{Z}_{10})$. Thus, α and β have multiplicative inverses α^{-1} and β^{-1}, respectively. We must show that the product $\alpha \cdot \beta$ has a multiplicative inverse. In other words, we must find a matrix γ such that $(\alpha \cdot \beta) \cdot \gamma = \varepsilon$ and $\gamma \cdot (\alpha \cdot \beta) = \varepsilon$. Noting that α^{-1} and β^{-1} exist, let $\gamma = \beta^{-1} \cdot \alpha^{-1}$. Then

$$(\alpha \cdot \beta) \cdot \gamma = (\alpha \cdot \beta) \cdot (\beta^{-1} \cdot \alpha^{-1}) = \alpha \cdot (\beta \cdot \beta^{-1}) \cdot \alpha^{-1} = \alpha \cdot \varepsilon \cdot \alpha^{-1} = \alpha \cdot \alpha^{-1} = \varepsilon,$$

and thus $(\alpha \cdot \beta) \cdot \gamma = \varepsilon$. Similar computation shows that $\gamma \cdot (\alpha \cdot \beta) = \varepsilon$. Therefore, $\alpha \cdot \beta$ has a multiplicative inverse, namely γ. Hence, $\alpha \cdot \beta \in G(\mathbb{Z}_{10})$. ∎

Proof know-how. In the above proof, we needed a matrix γ such that $(\alpha \cdot \beta) \cdot \gamma = \varepsilon$ and $\gamma \cdot (\alpha \cdot \beta) = \varepsilon$. We claimed and verified that $\gamma = \beta^{-1} \cdot \alpha^{-1}$ is the desired matrix. How did we come up with an expression for γ? Our goal was $(\alpha \cdot \beta) \cdot \gamma = \varepsilon$, so we worked backwards and solved this equation for γ by first left-multiplying both sides by α^{-1}, which yields $\beta \cdot \gamma = \alpha^{-1}$ and then left-multiply both sides by β^{-1} to obtain $\gamma = \beta^{-1} \cdot \alpha^{-1}$. This "working backwards" process of solving for γ is scratch work and must *not* be included in the proof itself. (Compare this with the Proof know-how after Theorem 9.10.)

Now we are ready to verify that $G(\mathbb{Z}_{10})$ is a group under multiplication:

(1) $G(\mathbb{Z}_{10})$ is closed under multiplication. (See Theorem 10.3 above.)

(2) Multiplication in $G(\mathbb{Z}_{10})$ is associative. After all, multiplication in $G(\mathbb{Z}_{10})$ is matrix multiplication, which we showed is associative in Chapter 7, Exercise #6.

(3) The element $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is the multiplicative identity. Moreover, ε is in $G(\mathbb{Z}_{10})$, because it has a multiplicative inverse, namely itself.

(4) By the definition of $G(\mathbb{Z}_{10})$, each element $\alpha \in G(\mathbb{Z}_{10})$ has a multiplicative inverse α^{-1} such that $\alpha \cdot \alpha^{-1} = \varepsilon$ and $\alpha^{-1} \cdot \alpha = \varepsilon$. We must also verify that $\alpha^{-1} \in G(\mathbb{Z}_{10})$, i.e., that α^{-1} has a multiplicative inverse; it does, since α is a multiplicative inverse of α^{-1}.

Remark. $M(\mathbb{Z}_{10})$ under addition is a *commutative* group, but $G(\mathbb{Z}_{10})$ under multiplication is not.

We make one more analogy between U_{10} and $G(\mathbb{Z}_{10})$, and more generally between U_m and $G(\mathbb{Z}_m)$. Both groups have a "trick" that allows us to easily determine whether or not an element is in the group. By definition, U_m is the set of elements in $\mathbb{Z}_m$ with multiplicative inverses. In practice, however, we rely on Theorem 4.19: *Let $a \in \mathbb{Z}_m$. Then $a \in U_m$ if and only if* $\gcd(a, m) = 1$. For example, we find that $8 \in U_{35}$ because $\gcd(8, 35) = 1$. (**Note:** $8 \cdot 22 = 1$ in $\mathbb{Z}_{35}$.) And $10 \notin U_{35}$ since $\gcd(10, 35) \neq 1$.

Similarly, $G(\mathbb{Z}_m)$ is the set of elements in $M(\mathbb{Z}_m)$ with multiplicative inverses. In practice, we use the following theorem, derived by combining Theorems 7.20 and 7.21: *Let $\alpha \in M(\mathbb{Z}_m)$. Then $\alpha \in G(\mathbb{Z}_m)$ if and only if* $\det \alpha \in U_m$. For example, $\alpha = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$, because $\det \alpha = 3 \in U_{10}$. And $\beta = \left[\begin{smallmatrix} 2 & 1 \\ 3 & 4 \end{smallmatrix}\right] \notin G(\mathbb{Z}_{10})$, since $\det \beta = 5 \notin U_{10}$.

10.3 Group $S(\mathbb{Z}_{10})$

As a way of previewing our work with *subgroups* in the next chapter, define a new set

$$S(\mathbb{Z}_{10}) = \{\alpha \in M(\mathbb{Z}_{10}) \mid \det \alpha = 1\}.$$

The set $S(\mathbb{Z}_{10})$ is often called the *special linear group*, hence the use of the letter S in its name.

Example 10.4. Let $\alpha, \beta, \gamma \in M(\mathbb{Z}_{10})$, where $\alpha = \left[\begin{smallmatrix} 3 & 2 \\ 4 & 3 \end{smallmatrix}\right]$, $\beta = \left[\begin{smallmatrix} 7 & 2 \\ 5 & 3 \end{smallmatrix}\right]$, and $\gamma = \left[\begin{smallmatrix} 8 & 9 \\ 5 & 6 \end{smallmatrix}\right]$. Then $\det \alpha = 3 \cdot 3 - 2 \cdot 4 = 1$, $\det \beta = 7 \cdot 3 - 2 \cdot 5 = 1$, and $\det \gamma = 8 \cdot 6 - 9 \cdot 5 = 3$. Therefore, $\alpha, \beta \in S(\mathbb{Z}_{10})$, but $\gamma \notin S(\mathbb{Z}_{10})$.

We remark that $S(\mathbb{Z}_{10})$ is a subset of $G(\mathbb{Z}_{10})$, which you will prove in an exercise at the end of the chapter. Below, we will verify that $S(\mathbb{Z}_{10})$ is a group under multiplication, just like $G(\mathbb{Z}_{10})$. In fact, there is nothing special about $\mathbb{Z}_{10}$ here. Thus, we will generalize our argument to $S(\mathbb{Z}_m)$. Let's start with closure.

Theorem 10.5. *$S(\mathbb{Z}_m)$ is closed under multiplication.*

PROOF. Assume $\alpha, \beta \in S(\mathbb{Z}_m)$. Then $\det \alpha = 1$ and $\det \beta = 1$. Thus, $\det(\alpha \cdot \beta) = \det \alpha \cdot \det \beta = 1 \cdot 1 = 1$. Hence, $\det(\alpha \cdot \beta) = 1$ so that $\alpha \cdot \beta \in S(\mathbb{Z}_m)$. ■

Now we are ready to verify that $S(\mathbb{Z}_m)$ is a group under multiplication:

(1) $S(\mathbb{Z}_m)$ is closed under multiplication. (See Theorem 10.5 above.)

(2) Multiplication in $S(\mathbb{Z}_m)$ is associative. After all, multiplication in $S(\mathbb{Z}_m)$ is matrix multiplication, which we showed is associative in Chapter 7, Exercise #6.

(3) The element $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is the multiplicative identity. And ε itself is in $S(\mathbb{Z}_m)$, because $\det \varepsilon = 1$.

(4) Let $\alpha \in S(\mathbb{Z}_m)$. Since $S(\mathbb{Z}_m) \subseteq G(\mathbb{Z}_m)$, we have $\alpha \in G(\mathbb{Z}_m)$. We already know that $G(\mathbb{Z}_m)$ is a multiplicative group, and thus α^{-1} exists in $G(\mathbb{Z}_m)$. We must show that α^{-1} is in $S(\mathbb{Z}_m)$, which you will do in an exercise at the end of the chapter.

Therefore, $S(\mathbb{Z}_m)$ is a group under multiplication. Since $S(\mathbb{Z}_m)$ is a subset of the group $G(\mathbb{Z}_m)$ and they share the same operation (i.e., multiplication), we say that $S(\mathbb{Z}_m)$ is a *subgroup* of $G(\mathbb{Z}_m)$. We'll say much more about subgroups in the next chapter.

Exercises

1. **(Review)** Consider the set $U_{10} = \{1, 3, 7, 9\}$.
   (a) Complete its multiplication table below.

$\cdot$	1	3	7	9
1				
3				
7				
9				

   (b) Use the table created to verify the group properties for U_{10}.

2. Proceed as in Example 10.1 to show that 2^{-1}, 4^{-1}, 6^{-1}, and 8^{-1} do *not* exist in $\mathbb{Z}_{10}$.

3. (a) Determine the number of elements in U_5; in U_7; in U_{13}; in U_{29}; in U_{101}.
   (b) Determine the number of elements in U_p where p is prime. Explain your reasoning.

4. (a) Without listing them, find the number of elements in U_{35}.
   (b) Repeat part (a), but with U_{39}; with U_{55}; with U_{91}.
   (c) Repeat part (a), but with U_{pq} where p and q are distinct odd primes.

5. (a) Without listing them, find the number of elements in U_{27}.
   (b) Repeat part (a), but with U_{25}; with U_{81}; with U_{125}; with U_{343}.
   (c) Repeat part (a), but with U_m, where $m = p^n$ for a prime number p.

6. For each matrix in $M(\mathbb{Z}_{10})$ shown below, determine whether or not it is in $G(\mathbb{Z}_{10})$.

   (a) $\begin{bmatrix} 9 & 5 \\ 8 & 3 \end{bmatrix}$. (b) $\begin{bmatrix} 8 & 7 \\ 6 & 4 \end{bmatrix}$. (c) $\begin{bmatrix} 6 & 6 \\ 7 & 4 \end{bmatrix}$. (d) $\begin{bmatrix} 5 & 8 \\ 7 & 9 \end{bmatrix}$.

7. For each matrix α from Exercise #6 that is in $G(\mathbb{Z}_{10})$, find α^{-1} and verify that $\alpha \cdot \alpha^{-1} = \varepsilon$ and $\alpha^{-1} \cdot \alpha = \varepsilon$, where $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.

8. Find five matrices in $M(\mathbb{Z}_{10})$ that are *not* in $G(\mathbb{Z}_{10})$.

9. In the proof of Theorem 10.3, show that $\gamma \cdot (\alpha \cdot \beta) = \varepsilon$.

10. Use Theorem 7.24 (i.e., $\det(\alpha \cdot \beta) = \det \alpha \cdot \det \beta$) to prove that $G(\mathbb{Z}_{10})$ is closed under multiplication.

11. Find the values of m for which $G(\mathbb{Z}_m)$ could be defined as $G(\mathbb{Z}_m) = \{\alpha \in M(\mathbb{Z}_m) \mid \det \alpha \neq 0\}$.

12. **(Socks-shoes revisited)** Let $\alpha, \beta \in M(\mathbb{Z}_{10})$ where $\alpha = \begin{bmatrix} 3 & 4 \\ 5 & 1 \end{bmatrix}$ and $\beta = \begin{bmatrix} 7 & 1 \\ 0 & 3 \end{bmatrix}$.

    (a) Verify that α and β are in $G(\mathbb{Z}_{10})$.

    (b) Verify that $(\alpha \cdot \beta)^{-1} \neq \alpha^{-1} \cdot \beta^{-1}$, but $(\alpha \cdot \beta)^{-1} = \beta^{-1} \cdot \alpha^{-1}$.

13. Let $\tau \in G(\mathbb{Z}_{10})$ where $\tau = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

    (a) Explain why τ is in $G(\mathbb{Z}_{10})$.

    (b) With α as in Exercise #12 above, compute the products $\alpha \cdot \tau$ and $\tau \cdot \alpha$.

    (c) With β as in Exercise #12 above, compute the products $\beta \cdot \tau$ and $\tau \cdot \beta$.

    (d) Describe what happens when we multiply any matrix by τ.

14. Again, let $\tau \in G(\mathbb{Z}_{10})$ where $\tau = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$. Recall that the set $C(\tau) = \{\sigma \in G(\mathbb{Z}_{10}) \mid \sigma \cdot \tau = \tau \cdot \sigma\}$ is called the *centralizer* of τ in $G(\mathbb{Z}_{10})$. In other words, $C(\tau)$ is the set of elements in $G(\mathbb{Z}_{10})$ that commute with τ. Prove that $C(\tau) = G(\mathbb{Z}_{10})$.

15. (a) Find five matrices in $S(\mathbb{Z}_{10})$ other than $\varepsilon = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

    (b) Find the multiplicative inverse of each matrix in part (a).

    (c) Verify that the multiplicative inverses you found in part (b) are also in $S(\mathbb{Z}_{10})$.

16. **Prove:** $S(\mathbb{Z}_m) \subseteq G(\mathbb{Z}_m)$. (This exercise is referenced in Section 11.2.)

17. Let $\alpha \in S(\mathbb{Z}_m)$. Since $S(\mathbb{Z}_m) \subseteq G(\mathbb{Z}_m)$, we have $\alpha \in G(\mathbb{Z}_m)$. We already know that $G(\mathbb{Z}_m)$ is a multiplicative group, and thus α^{-1} exists in $G(\mathbb{Z}_m)$. Prove that α^{-1} is in $S(\mathbb{Z}_m)$.

18. Define the set $H = \{\alpha \in M(\mathbb{Z}_{10}) \mid \det \alpha = 3\}$.

    (a) **Prove:** $H \subseteq G(\mathbb{Z}_{10})$.

    (b) Is H a group under matrix multiplication? Explain why or why not.

19. **(Thought experiment)** U_{10} is a subset of $\mathbb{Z}_{10}$, and both are groups. Should U_{10} be called a *subgroup* of $\mathbb{Z}_{10}$? Analogously, should $G(\mathbb{Z}_{10})$ be called a subgroup of $M(\mathbb{Z}_{10})$? Why or why not?

In Exercises #20 through #24 below, let H be a subset of $G(\mathbb{Z}_m)$ defined by $H = \{\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \mid a \in U_m\}$.

20. Explain why H is a subset of $G(\mathbb{Z}_m)$.

21. Suppose $m = 35$ so that H is a subset of $G(\mathbb{Z}_{35})$.

    (a) Find two elements $\alpha, \beta \in H$.

    (b) Compute the product $\alpha \cdot \beta$. Verify that $\alpha \cdot \beta$ is in H.

    (c) Compute the inverses α^{-1} and β^{-1}. Verify that these inverses are in H.

22. **Prove:** H is a group under matrix multiplication.

    Note: You may assume that matrix multiplication is associative. So, you must show the following:

    - H is closed under matrix multiplication.
    - The identity matrix ε is in H.
    - For each $\alpha \in H$, its inverse α^{-1} is also in H.

23. For simplicity, let $G = G(\mathbb{Z}_m)$. The *center* of G is defined by $Z(G) = \{\sigma \in G \mid \sigma\tau = \tau\sigma \text{ for all } \tau \in G\}$.

    Prove: $H \subseteq Z(G)$.

24. With the notations as in Exercise #23, prove that $Z(G) \subseteq H$ so that $H = Z(G)$.

11

Subgroups

In Section 10.3, we saw that $S(\mathbb{Z}_{10})$ is a subset of $G(\mathbb{Z}_{10})$ and that $S(\mathbb{Z}_{10})$ is a group under multiplication, just like $G(\mathbb{Z}_{10})$. Thus, we say that $S(\mathbb{Z}_{10})$ is a *subgroup* of $G(\mathbb{Z}_{10})$. Subgroups play an important role in our study of groups. For instance, we computed the *order* of a group element in several earlier exercises (Chapter 4, Exercise #11; Chapter 5, Exercise #10; Chapter 6, Exercise #7). You may have conjectured that the order of a group element is a divisor of the number of elements in the group. The proof of this conjecture involves the use of a certain subgroup. This is just one of many ways in which a subgroup of a group provides useful information about the group itself. Furthermore, subgroups can also be used to create new groups, as we'll see later when we study *quotient groups*.

11.1 Examples of subgroups

Example 11.1. Recall that $\mathbb{Z}_8 = \{0, 1, 2, 3, 4, 5, 6, 7\}$ is a group under addition. Let $H = \{0, 2, 4, 6\}$ be a subset of $\mathbb{Z}_8$. The addition table for H is shown below:

$+$	0	2	4	6
0	0	2	4	6
2	2	4	6	0
4	4	6	0	2
6	6	0	2	4

We verify that H is also a group under addition:

(1) H is closed under addition, as every entry in the table (i.e., all possible sums) is an element of H.

(2) Addition in H is associative. After all, H uses addition in $\mathbb{Z}_8$, which is known to be associative.

(3) The element $0 \in H$ is the additive identity of H, because $0 + h = h$ and $h + 0 = h$ for all $h \in H$.

(4) Each element $h \in H$ has an additive inverse, which is also in H. Note that 2 and 6 are inverse pairs, while 0 and 4 are self-inverses.

Therefore, H is also a group, with the same operation as $\mathbb{Z}_8$.

In the example above, we say that H is a *subgroup* of $\mathbb{Z}_8$. Here's the definition.

Definition 11.2 (Subgroup)**.** Let G be a group. A subset $H \subseteq G$ is called a *subgroup* of G if H is also a group using the operation of G.

Example 11.3. In Example 9.7, we considered the subset $H = \{\varepsilon, r_{120}, r_{240}\}$ of the group D_3. Here is the table for H, using the operation $\circ$ of D_3:

$\circ$	ε	r_{120}	r_{240}
ε	ε	r_{120}	r_{240}
r_{120}	r_{120}	r_{240}	ε
r_{240}	r_{240}	ε	r_{120}

We see that H is also a group, with the same operation as D_3. Hence, H is a subgroup of D_3.

Example 11.4 (Non-example)**.** Consider the subset $H = \{0, 2, 6\}$ of the group $\mathbb{Z}_8$. Then H is not closed under addition, since $2 + 2 = 4$ and $4 \notin H$. Thus, H is not a subgroup of $\mathbb{Z}_8$. Interestingly, H does satisfy the other three group properties.

Example 11.5 (Non-example)**.** Consider $U_8 = \{1, 3, 5, 7\}$, which is a group under multiplication. Although U_8 is a subset of $\mathbb{Z}_8$, it is *not* a subgroup of $\mathbb{Z}_8$, since the operations of U_8 and $\mathbb{Z}_8$ are different.

Example 11.6. We've seen that $H = \{0, 2, 4, 6\}$ is a subgroup of $\mathbb{Z}_8$ under addition. Let's find all other subgroups of $\mathbb{Z}_8$. We start with the two extremes: $\{0\}$, i.e., the subset containing just the additive identity element 0, and the group $\mathbb{Z}_8$ itself. Both of these are subgroups of $\mathbb{Z}_8$.

Let H be a subgroup of $\mathbb{Z}_8$. Then H must contain the additive identity element 0. Now suppose $1 \in H$. Then since H is closed under addition, we must have

$$1 + 1 \in H,\ 1 + 1 + 1 \in H,\ 1 + 1 + 1 + 1 \in H,\ \ldots,$$

so that every element of $\mathbb{Z}_8$ is in H; i.e., $H = \mathbb{Z}_8$. But we've already counted $\mathbb{Z}_8$ as one of the subgroups, and thus we can assume that $1 \notin H$. Similarly, if 3, 5, or 7 is in H, then $H = \mathbb{Z}_8$. (You'll show this in an exercise at the end of the chapter.) Thus, we may also assume that $3, 5, 7 \notin H$.

Next, let's see what happens if $2 \in H$. Then by closure, $2 + 2 = 4$ and $2 + 2 + 2 = 6$ are also in H, which implies $H = \{0, 2, 4, 6\}$, which we've already seen is a subgroup of $\mathbb{Z}_8$. Thus, we can now assume that $2 \notin H$. Similarly, if $6 \in H$, then $6 + 6 = 4$ and $6 + 6 + 6 = 2$ are also in H, and so $H = \{0, 2, 4, 6\}$ again. Hence, let's assume that $6 \notin H$. The only remaining element is 4, from which we obtain $H = \{0, 4\}$.

In summary, here is the complete list of *all* subgroups of $\mathbb{Z}_8$: $\{0\}$, $\{0, 4\}$, $\{0, 2, 4, 6\}$, $\mathbb{Z}_8$ itself.

In Example 11.6, where we found the subgroups of $\mathbb{Z}_8$, we started with the two extremes, namely $\{0\}$ and $\mathbb{Z}_8$ itself. Here is a generalization.

Example 11.7. Let G be a group. Then $\{\varepsilon\}$, i.e., the subset containing just the identity element, is a subgroup of G. It is often called the *trivial* subgroup. Moreover, G is also its own subgroup.

Example 11.8 (Diagonal subgroup)**.** Consider the direct product $U_{10} \times U_{10} = \{(a, b) \mid a, b \in U_{10}\}$, and define its subset $\triangle U_{10} = \{(a, a) \mid a \in U_{10}\}$. For instance, $(3, 3) \in \triangle U_{10}$, but $(3, 7) \notin \triangle U_{10}$. We'll leave it to you as an exercise at the end of the chapter to verify that $\triangle U_{10}$ is a subgroup of $U_{10} \times U_{10}$. Note that we call $\triangle U_{10}$ the *diagonal subgroup* of $U_{10} \times U_{10}$.

11.2 Subgroup proofs

In Chapter 10, we considered the following sets of matrices:

- $G(\mathbb{Z}_m) = \{\alpha \in M(\mathbb{Z}_m) \mid \alpha \text{ has a multiplicative inverse}\}$.
- $S(\mathbb{Z}_m) = \{\alpha \in M(\mathbb{Z}_m) \mid \det \alpha = 1\}$.

When $S(\mathbb{Z}_m)$ was introduced, we'd already known that $G(\mathbb{Z}_m)$ is a group under matrix multiplication. We also found that $S(\mathbb{Z}_m)$ is a subset of $G(\mathbb{Z}_m)$, which you showed in Chapter 10, Exercise #16. Then we verified that $S(\mathbb{Z}_m)$ is a group under the same operation as $G(\mathbb{Z}_m)$. Thus, $S(\mathbb{Z}_m)$ is a subgroup of $G(\mathbb{Z}_m)$.

We will revisit the argument that $S(\mathbb{Z}_m)$ is a group, but this time using the framework of a subgroup. Using this example as a model, we will then describe a general format for writing a "subgroup proof." To show that $S(\mathbb{Z}_m)$ is a subgroup of $G(\mathbb{Z}_m)$, we must show the following:

(1) $S(\mathbb{Z}_m)$ is closed under multiplication: If $\alpha, \beta \in S(\mathbb{Z}_m)$, then $\alpha \cdot \beta \in S(\mathbb{Z}_m)$.

(2) Since $G(\mathbb{Z}_m)$ is already known to be a group, its operation is associative. Since $S(\mathbb{Z}_m)$ inherits the same operation from $G(\mathbb{Z}_m)$, we're ensured that the operation of $S(\mathbb{Z}_m)$ is also associative. *For this reason, we don't even have to mention associativity in a subgroup proof.*

(3) Since $G(\mathbb{Z}_m)$ is a group, it has an identity element ε where $\varepsilon \cdot \alpha = \alpha$ and $\alpha \cdot \varepsilon = \alpha$ for all $\alpha \in G(\mathbb{Z}_m)$. Because $S(\mathbb{Z}_m) \subseteq G(\mathbb{Z}_m)$, this identity element will also keep all elements in $S(\mathbb{Z}_m)$ unchanged. Thus, it remains to show that ε is contained in $S(\mathbb{Z}_m)$.

(4) We must show the following: If $\gamma \in S(\mathbb{Z}_m)$, then $\gamma^{-1} \in S(\mathbb{Z}_m)$. Given $\gamma \in S(\mathbb{Z}_m)$, we know that $\gamma \in G(\mathbb{Z}_m)$, since $S(\mathbb{Z}_m)$ is a subset of $G(\mathbb{Z}_m)$. Also, since $G(\mathbb{Z}_m)$ is a group, we know that γ^{-1} exists in $G(\mathbb{Z}_m)$. But we have to show that γ^{-1} is contained in $S(\mathbb{Z}_m)$.

Theorem 11.9. *$S(\mathbb{Z}_m)$ is a subgroup of $G(\mathbb{Z}_m)$.*

PROOF. We will show that $S(\mathbb{Z}_m)$ is closed under matrix multiplication. Assume α, $\beta \in S(\mathbb{Z}_m)$. Then $\det \alpha = 1$ and $\det \beta = 1$. Thus, $\det(\alpha \cdot \beta) = \det \alpha \cdot \det \beta = 1 \cdot 1 = 1$. Hence, $\det(\alpha \cdot \beta) = 1$ and so $\alpha \cdot \beta \in S(\mathbb{Z}_m)$. Next, note that $\det \varepsilon = \det \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] = 1$ and thus $\varepsilon \in S(\mathbb{Z}_m)$. Lastly, suppose $\gamma \in S(\mathbb{Z}_m)$. Then $\det \gamma = 1$. Note that $\det(\gamma^{-1}) = (\det \gamma)^{-1} = 1^{-1} = 1$. Thus $\gamma^{-1} \in S(\mathbb{Z}_m)$. ∎

Here is a general format of a subgroup proof:

Proof know-how. In a *subgroup proof*, we are typically given a group G and a subset $H \subseteq G$. The key here is that G is already known to be a group. Thus the operation of G, which is inherited by H, is known to be associative. And G has an identity element ε, which would serve as an identity element for H, once we can show that $\varepsilon \in H$. Lastly, suppose $h \in H$. Then since G is a group, we know that h^{-1} exists in G. It remains to show that h^{-1} is in H.

To summarize, here are the verifications needed to show that H is a subgroup of G. Thus, a subgroup proof has three parts: closure, identity, and inverses.

(1) Closure: If $a, b \in H$, then $ab \in H$.

(2) (No need to even mention associativity, since H inherits the associative operation from G.)

(3) Identity: $\varepsilon \in H$, where ε is the identity element of G.

(4) Inverses: If $h \in H$, then $h^{-1} \in H$.

Example 11.10. Let $H = \left\{ \left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right] \mid a \in U_{13} \right\}$. Examples of matrices in H include $\alpha = \left[\begin{smallmatrix} 4 & 0 \\ 0 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 5 & 0 \\ 0 & 5 \end{smallmatrix}\right]$, since $4, 5 \in U_{13}$. Notice that $\det \alpha = 4 \cdot 4 - 0 \cdot 0 = 3 \in U_{13}$, and thus $\alpha \in G(\mathbb{Z}_{13})$. Similarly, $\det \beta = 12 \in U_{13}$ so that $\beta \in G(\mathbb{Z}_{13})$. More generally, suppose $\gamma = \left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right] \in H$ with $a \in U_{13}$. Then $\det \gamma = a \cdot a \in U_{13}$, since U_{13} is closed under multiplication. Thus, $\gamma \in G(\mathbb{Z}_{13})$, which implies that H is a subset of $G(\mathbb{Z}_{13})$.

In Theorem 11.11, we will prove that H is a subgroup of $G(\mathbb{Z}_{13})$. The following calculations foreshadow the proof. We have

$$\alpha \cdot \beta = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 4 \cdot 5 + 0 \cdot 0 & 4 \cdot 0 + 0 \cdot 5 \\ 0 \cdot 5 + 4 \cdot 0 & 0 \cdot 0 + 4 \cdot 5 \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}.$$

Therefore $\alpha \cdot \beta = \left[\begin{smallmatrix} 7 & 0 \\ 0 & 7 \end{smallmatrix}\right]$, which is in H, since $7 \in U_{13}$. To find α^{-1}, let $\Delta = \det \alpha = 3$ so that $\Delta^{-1} = 9$ (since $3 \cdot 9 = 1$ modulo 13). Then,

$$\alpha^{-1} = \Delta^{-1} \cdot \begin{bmatrix} 4 & -0 \\ -0 & 4 \end{bmatrix} = 9 \cdot \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}.$$

Thus $\alpha^{-1} = \left[\begin{smallmatrix} 10 & 0 \\ 0 & 10 \end{smallmatrix}\right]$, which is in H, since $10 \in U_{13}$.

Theorem 11.11. *Let* $H = \left\{ \left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right] \in G(\mathbb{Z}_{13}) \mid a \in U_{13} \right\}$. *Then* H *is a subgroup of* $G(\mathbb{Z}_{13})$.

PROOF. We will show that H is closed under matrix multiplication. Assume $\alpha, \beta \in H$. Then $\alpha = \left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} b & 0 \\ 0 & b \end{smallmatrix}\right]$, where $a, b \in U_{13}$. Thus $\alpha \cdot \beta = \left[\begin{smallmatrix} a \cdot b & 0 \\ 0 & a \cdot b \end{smallmatrix}\right]$, which is in H, since $a \cdot b \in U_{13}$. Next, note that $\varepsilon = \left[\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right]$ with $a = 1 \in U_{13}$, and thus $\varepsilon \in H$. Lastly, suppose $\gamma = \left[\begin{smallmatrix} c & 0 \\ 0 & c \end{smallmatrix}\right] \in H$, where $c \in U_{13}$. To find γ^{-1}, let $\Delta = \det \gamma = c^2$ so that $\Delta^{-1} = c^{-2}$, which exists because $c \in U_{13}$. Then

$$\gamma^{-1} = \Delta^{-1} \cdot \begin{bmatrix} c & -0 \\ -0 & c \end{bmatrix} = c^{-2} \cdot \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = \begin{bmatrix} c^{-1} & 0 \\ 0 & c^{-1} \end{bmatrix},$$

where $c^{-1} \in U_{13}$. Thus, $\gamma^{-1} \in H$. ∎

Remark. In the above proof, we used the fact that U_{13} is a multiplicative group. We stated that $a \cdot b \in U_{13}$, because $a, b \in U_{13}$ (i.e., closure). We also claimed that c^{-2} exists and $c^{-1} \in U_{13}$, since $c \in U_{13}$ (i.e., inverses).

Example 11.12. Given a group G under multiplication, define its subset $H = \{g \in G \mid g^2 = \varepsilon\}$. Let's start by looking at an example where $G = U_{16} = \{1, 3, 5, 7, 9, 11, 13, 15\}$. In this case, the identity element is $\varepsilon = 1$. We have $1^2 = 1$, $7^2 = 1$, $9^2 = 1$, $15^2 = 1$. No other element of U_{16}, when squared, equals 1. Therefore, $H = \{1, 7, 9, 15\}$.

The multiplication table of H is given by

$\cdot$	1	7	9	15
1	1	7	9	15
7	7	1	15	9
9	9	15	1	7
15	15	9	7	1

Note how (1) the set H is closed, (2) multiplication is associative (or observe that H inherits the associative operation from U_{16}), (3) $\varepsilon = 1$ is contained in H, and (4) every element of H has an inverse in H; in this case, the definition of H implies that every element of H is a self-inverse. Thus, H is a subgroup of U_{16}. Below, we will prove that such a subset H is always a subgroup, provided that G is commutative.

Theorem 11.13. *If G is a commutative group, then $H = \{g \in G \mid g^2 = \varepsilon\}$ is a subgroup of G.*

PROOF. We will first show that H is closed. Assume $a, b \in H$. Then $a^2 = \varepsilon$ and $b^2 = \varepsilon$. Since G is commutative, we have $(ab)^2 = a^2b^2 = \varepsilon \cdot \varepsilon = \varepsilon$. Hence, $(ab)^2 = \varepsilon$ and so $ab \in H$.

Next, note that $\varepsilon^2 = \varepsilon$ and thus $\varepsilon \in H$. Lastly, suppose $c \in H$. Then $c^2 = \varepsilon$. Note that $(c^{-1})^2 = (c^2)^{-1} = \varepsilon^{-1} = \varepsilon$. Thus $(c^{-1})^2 = \varepsilon$ so that $c^{-1} \in H$. ∎

Remark. Note how the proofs of Theorems 11.9, 11.11, and 11.13 are quite similar. In fact, most subgroup proofs will have the same structure as these three proofs.

11.3 Center and centralizer revisited

In this section, we will re-introduce a couple of important subgroups that we've already seen. We may view our first subgroup as the "commutative part" of a given group G.

Definition 11.14 (Center)**.** Given a group G, define the *center* of G by

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

In other words, $Z(G)$ contains the elements of G that commute with *all* elements of G. As we noted in Chapter 5, Exercise #7, the use of the letter Z originates from the German word Zentrum ("center").

Example 11.15 (Chapter 5, Exercise #7 revisited)**.** Let's find the elements of $Z(D_4)$, where D_4 is the group of symmetries of a square. We have $\varepsilon \circ \sigma = \sigma \circ \varepsilon$ (both are equal to σ) for all $\sigma \in D_4$, and thus $\varepsilon \in Z(D_4)$. We also claim that $r_{180} \in Z(D_4)$. In the composition table for D_4 (see Appendix B), the *row* r_{180}, which contains the elements of the form $r_{180} \circ \sigma$ (where $\sigma \in D_4$), and the *column* r_{180}, containing elements of the form $\sigma \circ r_{180}$, have the same orderings of their elements. This implies $r_{180} \circ \sigma = \sigma \circ r_{180}$

for all $\sigma \in D_4$, and thus $r_{180} \in Z(D_4)$. The table also tells us that ε and r_{180} are the only elements in $Z(D_4)$. For instance, we have $r_{90} \circ d \neq d \circ r_{90}$. Thus neither r_{90} nor d is in $Z(D_4)$; i.e., if they don't commute with each other, they certainly don't commute with *all* elements of D_4. Therefore, we conclude that $Z(D_4) = \{\varepsilon, r_{180}\}$.

Example 11.16. If G is a commutative group, then $Z(G) = G$. The converse is also true: If $Z(G) = G$, then G is commutative. You will prove these in an exercise at the end of the chapter.

To define our second subgroup, we begin by fixing a group element, say h. Then we consider all elements of the group that commute with h.

Definition 11.17 (Centralizer)**.** Given a *fixed* element h in a group G, define the *centralizer* of h in G by

$$C(h) = \{g \in G \mid gh = hg\}.$$

Example 11.18. In Section 5.3, we fixed $h \in D_4$, i.e., the reflection of a square about its horizontal axis. Then we found that $C(h) = \{\varepsilon, r_{180}, h, v\}$. For example, $v \in C(h)$, because $v \circ h = h \circ v$, i.e., v commutes with h. Likewise, $r_{180} \in C(h)$, because $r_{180} \circ h = h \circ r_{180}$. In fact, we saw in Example 11.15 above that r_{180} commutes with *all* elements of D_4. Thus, it is expected that r_{180} commutes with h. On the other hand, $r_{90} \notin C(h)$, because $r_{90} \circ h \neq h \circ r_{90}$.

As implied above, both $Z(G)$ and $C(h)$ are subgroups of the group G. We'll state those facts as theorems below. You'll be asked to prove them in an exercise at the end of the chapter.

Theorem 11.19. *Let G be a group. Then $Z(G) = \{z \in G \mid zg = gz \text{ for all } g \in G\}$ is a subgroup of G.*

Theorem 11.20. *Let h be a fixed element in a group G. Then $C(h) = \{g \in G \mid gh = hg\}$ is a subgroup of G.*

Exercises

1. In each case, explain why H is *not* a subgroup of G.

   (a) $G = \mathbb{Z}$ under addition, $H = \{n \in \mathbb{Z} \mid n \geq 0\}$.

   (b) $G = \mathbb{Z}_{10}$ under addition, $H = \{0, 3, 7\}$.

   (c) $G = \mathbb{Z}_{12}$ under addition, $H = U_{12}$.

   (d) $G = S_3$ under composition, $H = \{\varepsilon, \sigma\}$ where $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$.

2. Let H be a subgroup of $\mathbb{Z}_8$. Explain why if $3 \in H$, then $H = \mathbb{Z}_8$. Do likewise for $5 \in H$ and $7 \in H$.

   Note: See Example 11.6 for the context surrounding this problem.

3. Let H be a subgroup of G. For each statement, if it's true, prove it; if it's false, give a counterexample.
   (a) If G is non-commutative, then H is non-commutative.
   (b) If H is non-commutative, then G is non-commutative.
4. Suppose K is a subgroup of H and H is a subgroup of G. Is K a subgroup of G? Why or why not?
5. Recall the following:
   - $M(\mathbb{Z}_{10})$ = the set of 2×2 matrices with entries in $\mathbb{Z}_{10}$.
   - $G(\mathbb{Z}_{10}) = \{\alpha \in M(\mathbb{Z}_{10}) \mid \alpha \text{ has a multiplicative inverse}\}$.

   (a) $M(\mathbb{Z}_{10})$ is a group under what operation?
   (b) $G(\mathbb{Z}_{10})$ is a group under what operation?
   (c) Is $G(\mathbb{Z}_{10})$ a subgroup of $M(\mathbb{Z}_{10})$? Why or why not?
6. In this problem, you'll work with the group D_4. See Appendix B for its group table.
   (a) Verify that $\{\varepsilon, r_{90}, r_{180}, r_{270}\}$ is a subgroup of D_4.
   (b) Verify that $\{\varepsilon, v\}$ is a subgroup of D_4.
   (c) **(Challenge)** Find all subgroups of D_4.
7. Find all subgroups of the additive group $\mathbb{Z}_{12}$. (This exercise is referenced in Example 14.10.)
8. (a) Find all subgroups of the additive group $\mathbb{Z}_7$.
   (b) Find all subgroups of the additive group $\mathbb{Z}_{11}$.
   (c) Find all subgroups of the additive group $\mathbb{Z}_{101}$.
   (d) Any conjectures?
9. Find all subgroups of the additive group $\mathbb{Z}$.
10. Let $H = \left\{ \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \mid a \in U_m,\ b \in \mathbb{Z}_m \right\}$.
    (a) With $m = 10$, find several matrices that are in H. Verify that they're also in $G(\mathbb{Z}_{10})$.
    (b) Explain why $H \subseteq G(\mathbb{Z}_m)$.
    (c) Prove that H is a subgroup of $G(\mathbb{Z}_m)$.
11. Let $H = \left\{ \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \mid b \in \mathbb{Z}_m \right\}$.
    (a) With $m = 10$, find several matrices that are in H. Verify that they're also in $G(\mathbb{Z}_{10})$.
    (b) Explain why $H \subseteq G(\mathbb{Z}_m)$.
    (c) Prove that H is a subgroup of $G(\mathbb{Z}_m)$.

    (This exercise is referenced in Chapter 14, Exercise #11 and Chapter 16, Exercise #10.)
12. In the proof of Theorem 11.13, we say: Since G is commutative, we have $(ab)^2 = a^2b^2 = \varepsilon \cdot \varepsilon = \varepsilon$. Elaborate on how the commutativity of G was used in this sentence.

13. Let G be a group and let H and K be its subgroups. Define $M = H \cap K = \{g \in G \mid g \in H \text{ and } g \in K\}$; i.e., M is the *intersection* of H and K. Prove that M is a subgroup of G. (This exercise is referenced in Chapter 20, Exercise #4 and Chapter 26, Exercise #20.)

14. Using a counterexample, show that Theorem 11.13 is false when G is non-commutative.

15. Define the set $H = \{\alpha \in U_{20} \mid \alpha = g^2 \text{ where } g \in U_{20}\}$.

    (a) Find the elements of H.

    (b) Verify that H is a subgroup of U_{20}.

16. **Prove:** If G is a commutative group, then $H = \{\alpha \in G \mid \alpha = g^2 \text{ where } g \in G\}$ is a subgroup of G.

17. **Prove:** $5\mathbb{Z}$ is a subgroup of the additive group $\mathbb{Z}$.

18. Consider the additive group $\mathbb{Q}$, i.e., the set of rational numbers. Define a subset $H \subseteq \mathbb{Q}$ given by $H = \left\{\frac{3}{4}m + \frac{2}{7}n \mid m, n \in \mathbb{Z}\right\}$. Prove that H is a subgroup of $\mathbb{Q}$. (This exercise is referenced in Chapter 14, Exercise #12.)

19. Consider the set $H = \{5^k \mid k \in \mathbb{Z}\}$.

    (a) Explain why $H \subseteq \mathbb{R}^*$ where $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$.

    (b) Find three elements $\alpha, \beta, \gamma \in H$. Verify that $\alpha \cdot \beta \in H$ and that γ^{-1} exists and is in H.

    (c) **Prove:** H is a subgroup of $\mathbb{R}^*$.

20. Let G be a group and fix an element $h \in G$. For each statement, if it's true, prove it; if it's false, give a counterexample.

    (a) $C(h)$ is a subset of $Z(G)$.

    (b) $Z(G)$ is a subset of $C(h)$.

21. **Prove:** Let G be a group. Then G is commutative if and only if $Z(G) = G$.

22. (a) Prove Theorem 11.19.

    (b) Prove Theorem 11.20.

23. Consider the subsets $H = \{1, 12\}$ and $K = \{1, 3, 9\}$ of the multiplicative group U_{13}.

    (a) Verify that H is a subgroup of U_{13}. Do likewise for K.

    (b) With H and K as above, compute the set $HK = \{hk \mid h \in H, k \in K\}$ by multiplying every element of H by every element of K.

    (c) Verify that the set HK you found in part (b) is a subgroup of U_{13}.

24. Let H and K be subgroups of a commutative group G. Define $HK = \{hk \mid h \in H, k \in K\}$. Prove that HK is a subgroup of G. (This exercise is referenced in Chapter 17, Exercise #21.)

25. Consider the direct product $U_{10} \times U_{10} = \{(a, b) \mid a, b \in U_{10}\}$ and its subset $\triangle U_{10} = \{(a, a) \mid a \in U_{10}\}$.
    (a) Write down the elements of $\triangle U_{10}$.
    (b) Create a group table for $\triangle U_{10}$.
    (c) Use the table in part (b) to verify that $\triangle U_{10}$ is a subgroup of $U_{10} \times U_{10}$. (See Example 11.8.)

26. Let G be a group. Consider the direct product $G \times G$, and define its subset $\triangle G = \{(g, g) \mid g \in G\}$. Prove that $\triangle G$ is a subgroup of $G \times G$. (This exercise is referenced in Chapter 24, Exercise #22.)

27. **(Challenge) Prove:** For $n \geq 3$:
    - $Z(D_n) = \{\varepsilon, r_{180}\}$, if n is even.
    - $Z(D_n) = \{\varepsilon\}$, if n is odd.

12

Order of an Element

In previous exercises, we computed the *order* of a group element in various settings (in U_7, D_4, S_3, and $M(\mathbb{Z}_{10})$, just to name a few). The order is a useful property that tells us not only about the element itself, but also about the group containing that element.

In this chapter, we will formalize the notion of a *remainder* in integer division. For instance, when attempting to divide 263 by 6, we obtain a remainder of 5, because $263 = 6 \cdot 43 + 5$. The remainder will play a prominent role in various proofs, both in this chapter and beyond.

12.1 Motivating example

Example 12.1. Consider $3 \in U_7$. Recall that the order of 3 is the smallest positive exponent n such that $3^n = 1$. To find it, we must compute the powers of 3, noting that the calculation is done using the $\mathbb{Z}_7$ clock. We have $3^1 = 3$, $3^2 = 9 = 2$, $3^3 = 27 = 6$, $3^4 = 81 = 4$. Rather than computing 3^5 and reducing the result modulo 7, here's an alternate approach. We've found $3^4 = 4$. Multiplying both sides by 3 gives $3^5 = 4{\cdot}3 = 5$. Likewise, to find 3^6, we multiply both sides of $3^5 = 5$ by 3 to get $3^6 = 5 \cdot 3 = 1$. We've finally found $3^6 = 1$, so that the order of 3 is 6. We may denote this by $|3| = 6$ or $\text{ord}(3) = 6$.

Example 12.2. For $2 \in U_7$, we have $2^1 = 2$, $2^2 = 4$, $2^3 = 1$, $2^4 = 2$, $2^5 = 4$, $2^6 = 1$, … on the $\mathbb{Z}_7$ clock. Thus, $\text{ord}(2) = 3$. Although $2^6 = 1$, 6 is not the *smallest* positive exponent for which $2^n = 1$. So, $\text{ord}(2) \neq 6$.

Definition 12.3 (Order)**.** Let g be an element of a group. The *order* of g is the *smallest* positive exponent n such that $g^n = \varepsilon$. We often write $|g| = n$ or $\text{ord}(g) = n$.

Example 12.4. In any group, we have $\varepsilon^1 = \varepsilon$. Thus, the order of ε is 1. Moreover, if g is an element of a group with $\text{ord}(g) = 1$, then $g^1 = \varepsilon$ or simply $g = \varepsilon$. Thus, ε is the *only* element of any group with order 1.

Example 12.5. Consider $13 \in U_{15}$. To compute the order of 13, we must compute the powers of 13. For easier calculation, we observe that $-2 = 13$ in $\mathbb{Z}_{15}$. Thus, we compute the powers of -2 instead. We have $(-2)^1 = -2 = 13$, $(-2)^2 = 4$, $(-2)^3 = -8 = 7$, $(-2)^4 = 16 = 1$. Hence, $\text{ord}(13) = 4$.

Example 12.6. Let g be an element of a group with $\text{ord}(g) = 18$. Now let $h = g^3$. We claim that $\text{ord}(h) = 6$. To verify this, we must show that $h^6 = \varepsilon$ and that $n = 6$ is the smallest positive exponent such that $h^n = \varepsilon$. First, note that $h^6 = (g^3)^6 = g^{18}$. We have $g^{18} = \varepsilon$, since 18 is the order of g. Therefore, $h^6 = \varepsilon$.

Next, can there be a positive exponent $n < 6$ such that $h^n = \varepsilon$? Let's try $n = 4$, for instance. Then $h^n = (g^3)^4 = g^{12}$. But $g^{12} \neq \varepsilon$, since $\text{ord}(g) = 18$; i.e., $m = 18$ is the smallest positive exponent such that $g^m = \varepsilon$. We can similarly argue that $h^n \neq \varepsilon$ when $n = 1, 2, 3, 4$, or 5. Hence, $\text{ord}(h) = 6$, as desired.

Example 12.6 motivates the following theorem and proof.

Theorem 12.7. *Suppose g is a group element with $\text{ord}(g) = n$, and let d be a positive integer such that $d \mid n$. Then $\text{ord}(g^d) = \frac{n}{d}$.*

Remark. In the theorem statement above, the condition $d \mid n$ implies that $\frac{n}{d}$ is a (positive) integer.

PROOF. Since $d \mid n$, we have $n = dk$ for some integer k. Moreover, k is positive, since n and d are positive. We have $\frac{n}{d} = k$, so we must show that $\text{ord}(g^d) = k$. First, note that $(g^d)^k = g^{dk} = g^n$, as $n = dk$. We have $g^n = \varepsilon$, since n is the order of g. Therefore, $(g^d)^k = \varepsilon$.

Next, we will show that k is the smallest positive exponent such that $(g^d)^k = \varepsilon$. For contradiction, assume there is a smaller positive exponent $m < k$ such that $(g^d)^m = \varepsilon$. This equation can be written as $g^{dm} = \varepsilon$. As $0 < m < k$ and d is positive, we have $0 < dm < dk$. Substituting $n = dk$ implies $0 < dm < n$. But then $g^{dm} = \varepsilon$ contradicts the fact that $\text{ord}(g) = n$ is the smallest positive exponent such that $g^n = \varepsilon$.

Therefore, $\text{ord}(g^d) = \frac{n}{d}$, as desired. ■

Proof know-how. Let g be a group element. To prove $\text{ord}(g) = n$, we must show not only that $g^n = \varepsilon$, but also that n is the smallest such exponent. The latter can be shown using proof by contradiction; i.e., suppose for contradiction that there exists a smaller positive exponent $m < n$ such that $g^m = \varepsilon$.

Conversely, if we're given that $\text{ord}(g) = n$, then we may assume not only $g^n = \varepsilon$, but also that there does *not* exist a positive exponent $m < n$ such that $g^m = \varepsilon$, or equivalently that $g^1, g^2, g^3, \ldots, g^{n-1} \neq \varepsilon$.

12.2 When does $g^k = \varepsilon$?

Example 12.8. Consider $3 \in U_7$. We saw in Example 12.1 that $\text{ord}(3) = 6$ so that $3^6 = 1$ modulo 7. To find 3^{48}, we note that $6 \mid 48$ where $48 = 6 \cdot 8$. Then, $3^{48} = 3^{6 \cdot 8} = (3^6)^8 = 1^8 = 1$.

The above example motivates the following theorem.

Theorem 12.9. *Let g be an element of a group with $\operatorname{ord}(g) = n$. If $n \mid k$, then $g^k = \varepsilon$.*

PROOF. Assume $n \mid k$, so that $k = n \cdot q$ for some integer q. Since $\operatorname{ord}(g) = n$, we have $g^n = \varepsilon$. Thus, $g^k = g^{n \cdot q} = (g^n)^q = \varepsilon^q = \varepsilon$, as desired. ■

Example 12.10. Next, we will compute 3^{263} where $3 \in U_7$. This time, we note that $6 \nmid 263$; i.e., 6 is *not* a divisor of 263. When trying to divide 263 by 6, we get a remainder of 5: $263 = 6 \cdot 43 + 5$. Using laws of exponents, we obtain $3^{263} = 3^{6 \cdot 43 + 5} = (3^6)^{43} \cdot 3^5 = 1^{43} \cdot 3^5 = 3^5 = 5$. In particular, $3^{263} \neq 1$.

Example 12.11. Here's a (slight) generalization of Example 12.10 above. Let g be an element of a group with $\operatorname{ord}(g) = 6$. Recall that 6 is *not* a divisor of 263. Specifically, we have $263 = 6 \cdot 43 + 5$, i.e, a remainder of 5. Thus, $g^{263} = g^{6 \cdot 43 + 5} = (g^6)^{43} \cdot g^5 = \varepsilon^{43} \cdot g^5 = g^5$. Thus, $g^{263} = g^5$. Since $\operatorname{ord}(g) = 6$, there does *not* exist a positive exponent $m < 6$ such that $g^m = \varepsilon$. Since $5 < 6$, this implies that $g^5 \neq \varepsilon$, and therefore $g^{263} \neq \varepsilon$.

Before proceeding, let's dig deeper into the notion of a *remainder*. When dividing 263 by 6, we get a remainder of 5, because $263 = 6 \cdot 43 + 5$. It's also true that $263 = 6 \cdot 42 + 11$ and $263 = 6 \cdot 44 + (-1)$. But we don't say that the remainder is 11 or -1. Indeed, the remainder r must be less than the divisor (i.e., $r < 6$) and also non-negative (i.e., $r \geq 0$). Thus, $r = 5$ is the only possible remainder.

Example 12.12. When dividing 264 by 6, we obtain $264 = 6 \cdot 44 + 0$, or more simply $264 = 6 \cdot 44$. The remainder is 0 and thus 6 is a divisor of 264.

Example 12.13. When dividing -220 by 6, we obtain $-220 = 6 \cdot (-37) + 2$, so that the remainder is 2. Notice how the remainder is still less than the divisor and non-negative; i.e., $0 \leq 2 < 6$.

Example 12.14. Let's compute 3^{-220} where $3 \in U_7$. One approach is to use $-220 = 6 \cdot (-37) + 2$ from Example 12.13. Thus, $3^{-220} = 3^{6 \cdot (-37) + 2} = (3^6)^{-37} \cdot 3^2 = 1^{-37} \cdot 3^2 = 3^2 = 2$. Here, 1^{-37} can be viewed as $(1^{37})^{-1} = 1^{-1} = 1$, where $1^{-1} = 1$, because the multiplicative inverse of 1 is 1 in $\mathbb{Z}_7$ (or in any $\mathbb{Z}_m$).

Example 12.15. Here's a different approach to computing 3^{-220} where $3 \in U_7$. We write 3^{-220} as $(3^{220})^{-1}$. Since $220 = 6 \cdot 36 + 4$, we have $3^{220} = 3^{6 \cdot 36 + 4} = (3^6)^{36} \cdot 3^4 = 1^{36} \cdot 3^4 = 3^4 = 4$. Thus, $3^{-220} = (3^{220})^{-1} = 4^{-1} = 2$. Here, $4^{-1} = 2$ (i.e., the multiplicative inverse of 4 is 2 in $\mathbb{Z}_7$), because $4 \cdot 2 = 1$ in $\mathbb{Z}_7$.

The fact that we can always find the remainder that's "just right" may be something that you've taken for granted. It's a theorem from number theory, which we'll assume (without proof) in this textbook.

Theorem 12.16 (Division algorithm in $\mathbb{Z}$)**.** *Let a and b be integers, with $b > 0$. Then there exist $q, r \in \mathbb{Z}$ such that $a = b \cdot q + r$ with $0 \leq r < b$.*

We are now ready to fully generalize Examples 12.10 and 12.11. In the statement of the theorem below, note that $n \nmid k$ is a shorthand for "n is *not* a divisor of k."

Theorem 12.17. *Let g be an element of a group with $\operatorname{ord}(g) = n$. If $n \nmid k$, then $g^k \neq \varepsilon$.*

PROOF. Assume $n \nmid k$. Then $k = n \cdot q + r$ for some $q, r \in \mathbb{Z}$ with $0 < r < n$ (i.e., r is a non-zero remainder). We have $g^k = g^{n \cdot q + r} = (g^n)^q \cdot g^r = \varepsilon^q \cdot g^r = g^r$, so that $g^k = g^r$. Since r is positive and less than $n = \text{ord}(g)$, we know that $g^r \neq \varepsilon$. Thus $g^k \neq \varepsilon$, as desired. ■

Another way to prove Theorem 12.17 is to prove its contrapositive. The proof will look slightly different, which we present below.

PROOF. We will prove the contrapositive, namely: If $g^k = \varepsilon$, then $n \mid k$. Assume $g^k = \varepsilon$. By the division algorithm, we can write $k = n \cdot q + r$ for some $q, r \in \mathbb{Z}$ with $0 \leq r < n$. We will show that $r = 0$, so that $k = n \cdot q$. Solving for r in $k = n \cdot q + r$, we get $r = k - n \cdot q$. And since $\text{ord}(g) = n$, we have $g^n = \varepsilon$. Thus,

$$g^r = g^{k - n \cdot q} = g^k \cdot (g^n)^{-q} = \varepsilon \cdot \varepsilon^{-q} = \varepsilon.$$

Therefore, $g^r = \varepsilon$. But $r < n$ and n is the smallest positive integer such that $g^n = \varepsilon$. Hence, r cannot be positive. But $r \geq 0$, so it follows that $r = 0$. Thus $k = n \cdot q$, so that $n \mid k$ as desired. ■

The second proof of Theorem 12.17 highlights a couple of useful proof techniques:

Proof know-how. To prove that $n \mid k$, try to divide k by n and show that the remainder is 0. More precisely, we use the division algorithm to write $k = n \cdot q + r$ for some q, $r \in \mathbb{Z}$ with $0 \leq r < n$. Then show that $r = 0$, so that we get $k = n \cdot q$. In the proof, we showed that $g^r = \varepsilon$ where r is less than $n = \text{ord}(g)$. But n is the smallest *positive* exponent such that $g^n = \varepsilon$, and thus r cannot be positive. This led to the conclusion that $r = 0$ (since $r \geq 0$).

Combining Theorem 12.9 and the contrapositive of Theorem 12.17, we obtain the following "if and only if" theorem. In other words, only $\text{ord}(g)$ or its multiples satisfy $g^k = \varepsilon$.

Theorem 12.18. *Let g be an element of a group with $\text{ord}(g) = n$. Then $n \mid k$ if and only if $g^k = \varepsilon$.*

12.3 Conjugates

In this section, we will examine orders of elements in the matrix group $G(\mathbb{Z}_{10})$. Also embedded in these order calculations is the notion of a *conjugate* of an element, a concept which we will revisit when studying *normal subgroups* in Chapter 24.

Example 12.19. Let $\alpha \in G(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right]$. We have

$$\alpha = \begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix}, \ \alpha^2 = \begin{bmatrix} 9 & 6 \\ 0 & 1 \end{bmatrix}, \ \alpha^3 = \begin{bmatrix} 8 & 3 \\ 5 & 4 \end{bmatrix}, \ \alpha^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

so that $\text{ord}(\alpha) = 4$. Here, we employed the technique from Example 12.1 when computing the powers of α. For instance, after finding $\alpha^2 = \alpha \cdot \alpha$, we found α^3 by multiplying α^2 by α.

Example 12.20. Let $\alpha = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$ as in Example 12.19 above. Also let $\gamma = \left[\begin{smallmatrix} 3 & 5 \\ 1 & 2 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$ and note that $\gamma^{-1} = \left[\begin{smallmatrix} 2 & 5 \\ 9 & 3 \end{smallmatrix}\right]$. Define a new matrix in $G(\mathbb{Z}_{10})$ by $\beta = \gamma \cdot \alpha \cdot \gamma^{-1}$; i.e.,

$$\beta = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} 9 & 4 \\ 5 & 7 \end{bmatrix}.$$

Now, let's compute the order of β. We have

$$\beta = \begin{bmatrix} 9 & 4 \\ 5 & 7 \end{bmatrix},\ \beta^2 = \begin{bmatrix} 1 & 4 \\ 0 & 9 \end{bmatrix},\ \beta^3 = \begin{bmatrix} 9 & 2 \\ 5 & 3 \end{bmatrix},\ \beta^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

so that $\operatorname{ord}(\beta) = 4$, which is the same as $\operatorname{ord}(\alpha)$.

In Example 12.20, we defined $\beta = \gamma \cdot \alpha \cdot \gamma^{-1}$. We call β a *conjugate* of α. Here is a general definition.

Definition 12.21 (Conjugate element)**.** Let a and b be elements of a group G. We say that b is a *conjugate* of a if $b = gag^{-1}$ for some $g \in G$.

Remark. In the above definition, we have $b = gag^{-1}$ for some $g \in G$. Left-multiplying by g^{-1} and right-multiplying by g on both sides of the equation yield $g^{-1}bg = a$. Letting $h = g^{-1} \in G$, we have $h^{-1} = (g^{-1})^{-1} = g$. Thus, we can rewrite $g^{-1}bg = a$ as $a = hbh^{-1}$. Hence, if b is a conjugate of a, then a is a conjugate of b. Therefore, we can unambiguously say that a and b are conjugates of each other.

Returning to Example 12.20, here is another way to compute β^4, where $\beta = \gamma \cdot \alpha \cdot \gamma^{-1}$. Notice how the regrouping cancels all the γ and γ^{-1} in the interior of the product, leaving us with just $\gamma \cdot \alpha^4 \cdot \gamma^{-1}$:

$$\begin{aligned} \beta^4 &= (\gamma \cdot \alpha \cdot \gamma^{-1})^4 \\ &= (\gamma \cdot \alpha \cdot \gamma^{-1}) \cdot (\gamma \cdot \alpha \cdot \gamma^{-1}) \cdot (\gamma \cdot \alpha \cdot \gamma^{-1}) \cdot (\gamma \cdot \alpha \cdot \gamma^{-1}) \\ &= \gamma \cdot \alpha \cdot (\gamma^{-1} \cdot \gamma) \cdot \alpha \cdot (\gamma^{-1} \cdot \gamma) \cdot \alpha \cdot (\gamma^{-1} \cdot \gamma) \cdot \alpha \cdot \gamma^{-1} \\ &= \gamma \cdot \alpha \cdot (\varepsilon) \cdot \alpha \cdot (\varepsilon) \cdot \alpha \cdot (\varepsilon) \cdot \alpha \cdot \gamma^{-1} \\ &= \gamma \cdot \alpha^4 \cdot \gamma^{-1} \end{aligned}$$

You'll be asked to prove the following generalization in an exercise at the end of the chapter. Note that the exponent n can be either positive or negative (or zero).

Theorem 12.22. *Let a and b be conjugate elements in a group G, where $b = gag^{-1}$ for some $g \in G$. Then $b^n = ga^ng^{-1}$ for any integer n.*

In Example 12.20, the conjugate elements α and β had the same order. In fact, this is always true.

Theorem 12.23. *Let a and b be conjugate elements in a group G, where $b = gag^{-1}$ for some $g \in G$. Then* $\operatorname{ord}(a) = \operatorname{ord}(b)$.

PROOF. Let $m = \text{ord}(a)$ and $n = \text{ord}(b)$. We must show that $m = n$. Since m is the order of a, we have $a^m = \varepsilon$. Then Theorem 12.22 implies $b^m = (gag^{-1})^m = g \cdot a^m \cdot g^{-1} = g \cdot \varepsilon \cdot g^{-1} = \varepsilon$. Therefore $b^m = \varepsilon$. And since $n = \text{ord}(b)$, we conclude from Theorem 12.18 that $n \mid m$.

Next, let $h = g^{-1}$ so that $a = hbh^{-1}$ as shown in the remark after Definition 12.21. Since n is the order of b, we have $b^n = \varepsilon$. Thus, we obtain $a^n = (hbh^{-1})^n = h \cdot b^n \cdot h^{-1} = h \cdot \varepsilon \cdot h^{-1} = \varepsilon$, so that $a^n = \varepsilon$. And since $m = \text{ord}(a)$, we conclude that $m \mid n$.

Hence, $n \mid m$ and $m \mid n$, where m and n are positive integers (since they're orders). Thus, $m = n$. ■

Proof know-how. In the above proof, we let $m = \text{ord}(a)$ and $n = \text{ord}(b)$. This means $a^m = \varepsilon$ and $b^n = \varepsilon$. But the key to the proof is to swap the exponents and show that $a^n = \varepsilon$ and $b^m = \varepsilon$. Now, $a^n = \varepsilon$ allows us to conclude that $m = \text{ord}(a)$ is a divisor of n. Similarly, $b^m = \varepsilon$ implies that $n = \text{ord}(b)$ is a divisor of m. This "swap the exponents" technique can be useful when proving that the orders of two elements are equal.

12.4 Order in an additive group

Let g be an element of a group. We defined the *order* of g as the smallest positive exponent n such that $g^n = \varepsilon$. Note that

$$g^n = \underbrace{g \cdot g \cdot g \cdot \cdots \cdot g}_{n \text{ terms}}$$

entails multiplying g by itself n times. But what if we have a group whose operation is addition?

Example 12.24. Consider the *additive* group $\mathbb{Z}_8$. Let's find the order of $2 \in \mathbb{Z}_8$. Since the operation is addition, we seek the smallest positive number of times we add 2 to itself to get the additive identity 0. We have $2+2+2+2 = 0$, so that $\text{ord}(2) = 4$. Below, we compute the order of each element in $\mathbb{Z}_8$:

- $\text{ord}(0) = 1$, because $0 = 0$.
- $\text{ord}(1) = 8$, because $1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 = 0$.
- $\text{ord}(2) = 4$, because $2 + 2 + 2 + 2 = 0$.
- $\text{ord}(3) = 8$, because $3 + 3 + 3 + 3 + 3 + 3 + 3 + 3 = 0$.
- $\text{ord}(4) = 2$, because $4 + 4 = 0$.
- $\text{ord}(5) = 8$, because $5 + 5 + 5 + 5 + 5 + 5 + 5 + 5 = 0$.
- $\text{ord}(6) = 4$, because $6 + 6 + 6 + 6 = 0$.
- $\text{ord}(7) = 8$, because $7 + 7 + 7 + 7 + 7 + 7 + 7 + 7 = 0$.

Consistent with our past conjectures (see, for example, Chapter 4, Exercise #11), for each $g \in \mathbb{Z}_8$, $\text{ord}(g)$ is a divisor of 8, where 8 is the number of elements in $\mathbb{Z}_8$.

In the computation of $\text{ord}(2)$ where $2 \in \mathbb{Z}_8$, we found that $2 + 2 + 2 + 2 = 0$ and thus $\text{ord}(2) = 4$. We can rewrite $2 + 2 + 2 + 2$ as $4 \cdot 2$, which motivates the following definition.

Definition 12.25 (Order in an additive group)**.** Let g be an element of a group whose operation is addition. The *order* of g is the *smallest* positive integer n such that

$$n \cdot g = \underbrace{g + g + g + \cdots + g}_{n \text{ terms}} = 0.$$

Example 12.26. Consider $14 \in \mathbb{Z}_{16}$. To compute the order of 14, we must compute the sums of 14. For easier calculation, we observe that $-2 = 14$ in $\mathbb{Z}_{16}$. Thus, we compute the sums of -2 instead. We have

$$8 \cdot (-2) = \underbrace{(-2) + (-2) + (-2) + \cdots + (-2)}_{8 \text{ terms}} = 0,$$

where 8 is the smallest positive integer with this property. Thus, $\text{ord}(14) = 8$. How is this related to the fact that $\text{ord}(2) = 8$? You'll find out in an exercise at the end of the chapter!

12.5 Elements with infinite order

Let g be an element of a group. If there is no positive integer n such that $g^n = \varepsilon$ (in a multiplicative group) or $n \cdot g = 0$ (in an additive group), then we say that g has *infinite order*. Below are some examples.

Example 12.27. In Chapter 8, we saw that $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$ is a group under multiplication and contains all non-zero elements of $\mathbb{R}$. (Note that $\mathbb{R}$ denotes the set of all real numbers.) Consider $3 \in \mathbb{R}^*$. Then there is no positive integer n such that $3^n = 1$. Thus, $\text{ord}(3)$ is infinite.

Example 12.28. Consider $1 \in \mathbb{Z}$, where $\mathbb{Z}$ is a group under addition. Then there is no positive integer n such that $\underbrace{1 + 1 + 1 + \cdots + 1}_{n \text{ terms}} = 0$. Thus, $\text{ord}(1)$ is infinite.

In Example 12.27, we saw that in the group $\mathbb{R}^*$, the order of 3 is infinite. This means 3^1, 3^2, 3^3, 3^4, … will never equal 1. Moreover, all of these powers of 3 are different from each other. They are real numbers after all, so the value of 3^n gets larger as n increases. Therefore, the only way that $3^m = 3^n$ in $\mathbb{R}^*$ is when the exponents m and n are equal. (Contrast this to, say, $3 \in U_7$ where $\text{ord}(3) = 6$. We saw in Example 12.10 that $3^{263} = 3^5$, even though the exponents 263 and 5 are unequal.) Here is a generalization.

Theorem 12.29. *Let g be an element of a group with infinite order. Then $g^k = g^\ell$ if and only if $k = \ell$.*

PROOF. We must prove two implications:

- If $g^k = g^\ell$, then $k = \ell$.
- If $k = \ell$, then $g^k = g^\ell$.

The second implication is immediate, so the proof will focus on the first implication.

Assume $g^k = g^\ell$ where $k, \ell \in \mathbb{Z}$. Suppose for contradiction that $k \neq \ell$, so that $k > \ell$ or $\ell > k$. We will proceed with the case $k > \ell$. (The argument for the case $\ell > k$

follows similarly.) We right-multiply both sides of $g^k = g^\ell$ by $g^{-\ell}$. Then $g^k \cdot g^{-\ell} = g^\ell \cdot g^{-\ell}$, so that $g^{k-\ell} = g^{\ell-\ell}$. But $g^{\ell-\ell} = g^0 = \varepsilon$, so that $g^{k-\ell} = \varepsilon$. Since $k > \ell$, we have $k - \ell > 0$. Thus, we have found a positive exponent $k - \ell$ such that $g^{k-\ell} = \varepsilon$. This contradicts the fact that $\text{ord}(g)$ is infinite. Therefore, we must have $k = \ell$. ■

Exercises

1. Find the order of each element in the multiplicative group U_{20}.

2. In U_{26}, find $\text{ord}(25)$ and $\text{ord}(23)$, by writing 25 and 23 as negatives in $\mathbb{Z}_{26}$. (See Example 12.5.)

3. Find the order of each element in the additive group $\mathbb{Z}_{20}$.

4. In $\mathbb{Z}_{26}$, find $\text{ord}(25)$ and $\text{ord}(24)$, by writing 25 and 24 as negatives in $\mathbb{Z}_{26}$. (See Example 12.26.)

5. Let g be a non-identity group element. Prove that g is a self-inverse if and only if $\text{ord}(g) = 2$.

6. Let g be an element of a group with $\text{ord}(g) = 18$. Find each of the following:

   (a) $\text{ord}(g^2)$. (b) $\text{ord}(g^6)$. (c) $\text{ord}(g^4)$. (d) $\text{ord}(g^5)$.

7. Consider $2 \in U_{23}$. It turns out that $2^{22} = 1$, but $\text{ord}(2) \neq 22$. Find the order of 2 in U_{23}.

8. (a) Consider $3 \in U_{11}$. Verify that $\text{ord}(3) = 5$.
   (b) Find integers q and r where $312 = 5 \cdot q + r$ with $0 \leq r < 5$.
   (c) Find integers q and r where $-312 = 5 \cdot q + r$ with $0 \leq r < 5$.
   (d) Find the value of 3^{-312} by writing it as $(3^{312})^{-1}$ and using the result from part (b).
   (e) Find the value of 3^{-312} again, this time by using the result from part (c).

9. Let g be an element of a group with $\text{ord}(g) = 5$. Find the smallest non-negative integer k such that $g^{-312} = g^k$. Explain your reasoning.

10. Let g be an element of a group with $\text{ord}(g) = 6$. (Or just use $g = 3$ in U_7.)

    (a) Are g^{20} and g^{32} equal? Why or why not?
    (b) What about $g^{123,405}$ and $g^{123,465}$? How do you know?
    (c) What about g^{800} and g^{862}? How do you know?
    (d) What about g^{-241} and g^{359}? How do you know?
    (e) What's going on here? Can you generalize and justify?

    (This exercise is referenced in Example 13.11.)

11. In the multiplicative group U_7:
    (a) Verify that 2 and 4 are multiplicative inverses of each other. Compute $\text{ord}(2)$ and $\text{ord}(4)$.
    (b) Verify that 3 and 5 are multiplicative inverses of each other. Compute $\text{ord}(3)$ and $\text{ord}(5)$.
    (c) What conjecture do you have?

12. In the multiplicative group U_{13}:
    (a) Find all inverse pairs g and g^{-1} such that $g \neq g^{-1}$ (i.e., g is not a self-inverse).
    (b) For each pair g and g^{-1}, compute $\text{ord}(g)$ and $\text{ord}(g^{-1})$.
    (c) What conjecture do you have?

13. Let g be a group element with finite order. Prove each of these statements:
    (a) $\text{ord}(g^{-1})$ is finite.
    (b) $\text{ord}(g^{-1}) = \text{ord}(g)$.

14. **Prove:** Let g be a group element. If $\text{ord}(g)$ is infinite, then $\text{ord}(g^{-1})$ is infinite.

15. Let a, b be elements of a commutative group.
    (a) Suppose $\text{ord}(a) = 3$ and $\text{ord}(b) = 5$. Explain why $(ab)^{15} = \varepsilon$.
    (b) Suppose $\text{ord}(a) = 4$ and $\text{ord}(b) = 9$. Explain why $(ab)^{36} = \varepsilon$.
    (c) Suppose $\text{ord}(a) = 4$ and $\text{ord}(b) = 6$. Explain why $(ab)^{24} = \varepsilon$.
    (d) Elizabeth says, "In part (c), I showed $(ab)^{24} = \varepsilon$. That means $\text{ord}(ab) = 24$." Do you agree or disagree with her? Explain your reasoning.

16. Let a, b be elements of a commutative group, each with finite order. Using a counterexample, show that the following statement is false: $\text{ord}(ab) = \text{ord}(a) \cdot \text{ord}(b)$.

17. **Prove:** Let a, b be elements of a commutative group. If $\text{ord}(a)$ and $\text{ord}(b)$ are finite, then $\text{ord}(ab)$ is finite.

18. **Prove:** Let a, b be elements of a commutative group with $m = \text{ord}(a)$ and $n = \text{ord}(b)$. If $\gcd(m, n) = 1$, then $\text{ord}(ab) = mn$.

19. (a) Find all elements $n \in \mathbb{Z}$ whose order is finite.
    (b) Find all elements $x \in \mathbb{R}^*$ whose order is finite.

20. **Prove:** Let g be an element of a group with infinite order. If $g^n = \varepsilon$ where $n \in \mathbb{Z}$, then $n = 0$.

21. Let a and b be conjugate elements in a group G, where $b = gag^{-1}$ for some $g \in G$. Verify that $b^{-4} = ga^{-4}g^{-1}$ in the following two ways:
    (a) By writing $b^{-4} = (gag^{-1})^{-4}$ as $((gag^{-1})^4)^{-1}$.
    (b) By writing $b^{-4} = (gag^{-1})^{-4}$ as $((gag^{-1})^{-1})^4$.

22. Prove Theorem 12.22. (**Note:** Be sure to consider the cases $n > 0$, $n = 0$, and $n < 0$.)

23. Let G be a group and let H be a subgroup of G. Fix an element $g \in G$, and define the set
$$gHg^{-1} = \{ghg^{-1} \mid h \in H\}.$$
Prove that gHg^{-1} is a subgroup of G. (**Note:** The subgroup gHg^{-1} is called a *conjugate* of H.)

    (This exercise is referenced in Chapter 14, Exercise #21 and in Section 24.4.)

24. Let G be a group and suppose $z \in G$ is the only element of order 2. Prove that $z \in Z(G)$.

25. Consider the additive groups $\mathbb{Z}_{12}$ and $\mathbb{Z}_{20}$ and their direct product $\mathbb{Z}_{12} \times \mathbb{Z}_{20}$.
    (a) Find the orders of $3 \in \mathbb{Z}_{12}$, $2 \in \mathbb{Z}_{20}$, and $(3, 2) \in \mathbb{Z}_{12} \times \mathbb{Z}_{20}$.
    (b) Find the orders of $9 \in \mathbb{Z}_{12}$, $4 \in \mathbb{Z}_{20}$, and $(9, 4) \in \mathbb{Z}_{12} \times \mathbb{Z}_{20}$.
    (c) Find the orders of $1 \in \mathbb{Z}_{12}$, $1 \in \mathbb{Z}_{20}$, and $(1, 1) \in \mathbb{Z}_{12} \times \mathbb{Z}_{20}$.
    (d) Find the orders of $2 \in \mathbb{Z}_{12}$, $15 \in \mathbb{Z}_{20}$, and $(2, 15) \in \mathbb{Z}_{12} \times \mathbb{Z}_{20}$.
    (e) What conjecture do you have?

26. **Prove:** Let G and H be groups and consider the direct product $G \times H$. Let $(g, h) \in G \times H$ where $\text{ord}(g) = m$ and $\text{ord}(h) = n$. Then the order of (g, h) is the least common multiple of m and n.

13

Cyclic Groups, Part I

Cyclic groups are an important type of groups that is generated by a single element. Recall from Chapter 8, Exercise #6 that 1 is a *generator* of the additive group $\mathbb{Z}_{12}$, because its sums give all elements in the group, as shown below:

$$1 = 1,\ \ 1+1 = 2,\ \ 1+1+1 = 3,\ \ 1+1+1+1 = 4,\ \ldots,\ \underbrace{1+1+\cdots+1}_{12 \text{ terms}} = 0.$$

Thus, we say that $\mathbb{Z}_{12}$ is *cyclic*, because it has a generator.

In the next two chapters, we'll explore the various properties of cyclic groups. In this chapter, for instance, we'll find all the generators of $\mathbb{Z}_{12}$, and more generally of $\mathbb{Z}_m$. We'll also learn that the multiplicative group U_{13} is not only cyclic but behaves just like $\mathbb{Z}_{12}$. Understanding the relationship between these two groups will prepare us for our excursion into *isomorphism* in the next unit.

13.1 Generators of the additive group $\mathbb{Z}_{12}$

Example 13.1. We will find all the generators of $\mathbb{Z}_{12} = \{0,\ 1,\ 2,\ 3,\ \ldots,\ 11\}$.

- 0 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 0's only yield 0, i.e., $0,\ 0+0,\ 0+0+0$, and so on.
- We've seen that 1 is a generator of $\mathbb{Z}_{12}$.
- 2 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 2's only yield 0, 2, 4, 6, 8, 10.
- 3 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 3's only yield 0, 3, 6, 9.
- 4 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 4's only yield 0, 4, 8.
- 5 is a generator of $\mathbb{Z}_{12}$. Here are the first few sums of 5's in $\mathbb{Z}_{12}$:

  $$5 = 5,\ \ 5+5 = 10,\ \ 5+5+5 = 3,\ \ 5+5+5+5 = 8,\ \ldots.$$

  You'll complete this list in an exercise at the end of the chapter to show that every element of $\mathbb{Z}_{12}$ can be expressed as a sum of 5's.

- 6 is *not* a generator, since the sums of 6's only yield 0, 6.
- 7 is a generator of $\mathbb{Z}_{12}$. We can verify this by computing the sums of 7's, but here's another approach. Observe that $7 = -5$ in $\mathbb{Z}_{12}$. Thus, if 5 generates all elements of $\mathbb{Z}_{12}$, then 7 would generate the *negatives* of all elements of $\mathbb{Z}_{12}$, which is equal to the set $\mathbb{Z}_{12}$ itself.
- 8 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 8's only yield 0, 8, 4.
- 9 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 9's only yield 0, 9, 6, 3.
- 10 is *not* a generator of $\mathbb{Z}_{12}$, since the sums of 10's only yield 0, 10, 8, 6, 4, 2.
- 11 is a generator of $\mathbb{Z}_{12}$, because $11 = -1$ where 1 is a generator.

Therefore, the generators of $\mathbb{Z}_{12}$ are 1, 5, 7, and 11. Notice how these are precisely the elements of $\mathbb{Z}_{12}$ that are relatively prime to 12; i.e., $\gcd(1, 12) = \gcd(5, 12) = \gcd(7, 12) = \gcd(11, 12) = 1$.

Example 13.2. Next, consider the additive group $\mathbb{Z}_{15} = \{0, 1, 2, 3, \ldots, 14\}$. Based on our observation from Example 13.1, we suspect that the generators of $\mathbb{Z}_{15}$ are 1, 2, 4, 7, 8, 11, 13, 14. Again, these are the elements of $\mathbb{Z}_{15}$ that are relatively prime to 15.

For instance, let's verify that 7 is indeed a generator of $\mathbb{Z}_{15}$. If so, then we should be able to obtain any element of $\mathbb{Z}_{15}$, say 3, as a sum of 7's:

$$\underbrace{7 + 7 + \cdots + 7}_{k \text{ terms}} = 3,$$

so that $k \cdot 7 = 3$ in $\mathbb{Z}_{15}$. We can solve for k by multiplying both sides of the equation by 7^{-1}, which is 13 since $7 \cdot 13 = 1$ in $\mathbb{Z}_{15}$. This gives $k = 3 \cdot 13 = 9$, so that

$$\underbrace{7 + 7 + \cdots + 7}_{9 \text{ terms}} = 3.$$

Let's also verify that 6 is *not* a generator of $\mathbb{Z}_{15}$. Suppose for contradiction that it is. Then we should be able to obtain $1 \in \mathbb{Z}_{15}$ as a sum of 6's:

$$\underbrace{6 + 6 + \cdots + 6}_{k \text{ terms}} = 1,$$

so that $k \cdot 6 = 1$ in $\mathbb{Z}_{15}$. But this implies that 6 has a multiplicative inverse, namely k, and hence $6 \in U_{15}$. Then Theorem 4.19 would imply that $\gcd(6, 15) = 1$, which is a contradiction since $\gcd(6, 15) = 3$.

Example 13.2 above motivates the following theorem and proof.

Theorem 13.3. *Let $a \in \mathbb{Z}_m$. Then a is a generator of $\mathbb{Z}_m$ if and only if* $\gcd(a, m) = 1$.

PROOF. We must prove two implications:

- If a is a generator of $\mathbb{Z}_m$, then $\gcd(a, m) = 1$.
- If $\gcd(a, m) = 1$, then a is a generator of $\mathbb{Z}_m$.

We will prove the second implication. The proof of the first implication is left for you as an exercise.

Assume that $\gcd(a, m) = 1$. We must show that a is a generator of $\mathbb{Z}_m$. Thus, given an element $b \in \mathbb{Z}_m$, we must be able to write it as a sum of a's. Since $\gcd(a, m) = 1$, Theorem 4.19 implies that $a \in U_m$. Hence, a has a multiplicative inverse a^{-1} such that $a \cdot a^{-1} = 1$ and $a^{-1} \cdot a = 1$. Let $k = b \cdot a^{-1}$ (which is an element of $\mathbb{Z}_m$) so that multiplying both sides by a yields $k \cdot a = b$ or, equivalently,

$$\underbrace{a + a + \cdots + a}_{k \text{ terms}} = b.$$

Therefore, a is a generator of $\mathbb{Z}_m$, as desired. ■

Proof know-how. Once again, the "working backwards" technique was utilized when preparing this proof. (See, for comparison, the proof of Theorem 10.3.) Here, we needed an integer k such that $k \cdot a = b$. Thus, we worked backwards and solved this equation for k by multiplying both sides by a^{-1}, which yields $k = b \cdot a^{-1}$. This process of solving for k is scratch work and must *not* be included in the proof itself.

13.2 Generators of the multiplicative group U_{13}

Example 13.4. Consider the group $U_{13} = \{1, 2, 3, 4, \ldots, 12\}$. Since its operation is multiplication, a generator of U_{13} is an element whose *products* give all the elements in the group. We claim that 2 is a generator. To verify, we multiply 2 by itself (or compute powers of 2) and obtain all the elements of U_{13}:

$$2^1 = 2,\ 2^2 = 4,\ 2^3 = 8,\ 2^4 = 3,\ 2^5 = 6,\ 2^6 = 12$$
$$2^7 = 11,\ 2^8 = 9,\ 2^9 = 5,\ 2^{10} = 10,\ 2^{11} = 7,\ 2^{12} = 1.$$

We also claim that 7 is a generator of U_{13}. We can verify this by computing the powers of 7, but here's another approach. Observe that $7 = 2^{-1}$, i.e., 7 is the multiplicative inverse of 2, since $2 \cdot 7 = 1$ modulo 13. Thus, if 2 generates all elements of U_{13}, then 7 would generate the *multiplicative inverses* of all elements of U_{13}, which is equal to the set U_{13} itself.

However, 3 is *not* a generator of U_{13}. Taking powers of 3, we find

$$3^1 = 3,\ 3^2 = 9,\ 3^3 = 1,\ 3^4 = 3,\ 3^5 = 9,\ 3^6 = 1, \ldots,$$

so that the powers of 3 only yield 1, 3, 9.

Example 13.5. Consider the multiplicative group $U_7 = \{1, 2, 3, 4, 5, 6\}$. We have

$$3^1 = 3,\ 3^2 = 2,\ 3^3 = 6,\ 3^4 = 4,\ 3^5 = 5,\ 3^6 = 1,$$

so that 3 is a generator of U_7. Note that $5 = 3^{-1}$, since $3 \cdot 5 = 1$ modulo 7. Thus, 5 is also a generator of U_7, which you will verify in an exercise at the end of the chapter.

Example 13.6. Consider the multiplicative group $U_{12} = \{1, 5, 7, 11\}$. Let's see if it has a generator.

- Powers of 1 only yield 1.
- Powers of 5 only yield 1 and 5.

- Powers of 7 only yield 1 and 7.
- Powers of 11 only yield 1 and 11.

Thus, U_{12} does *not* have an element whose products (or powers) give all the elements in the group. Therefore, U_{12} does *not* have a generator.

Example 13.7. Here is a non-commutative group $D_4 = \{\varepsilon, r_{90}, r_{180}, r_{270}, h, v, d, d'\}$.

- Powers of ε only yield ε.
- Powers of r_{90} only yield $\varepsilon, r_{90}, r_{180}, r_{270}$. The same is true for powers of r_{270}.
- Powers of r_{180} only yield ε, r_{180}.
- Powers of each reflection only yield ε and the reflection itself (e.g., powers of h only yield ε and h).

Therefore, D_4 does not have a generator and thus is not cyclic. In an exercise at the end of the chapter, you will prove the following: If a group is cyclic, then it is commutative. The contrapositive of that statement is: If a group is not commutative, then it is not cyclic. The group D_4 serves as an example of the contrapositive.

13.3 Matching $\mathbb{Z}_{12}$ and U_{13}

In Example 13.4, we found that 2 and 7 are generators of U_{13}. Are those the only generators of U_{13}, or are there more generators? To answer this question, we will *not* compute the powers of each element in U_{13}. While that's a valid approach, it can also get quite tedious! Instead, we will form a *correspondence* between $\mathbb{Z}_{12}$ and U_{13}. Then we will apply our knowledge of the additive group $\mathbb{Z}_{12}$ to help us better understand the multiplicative group U_{13}.

Once again, here are the powers of the generator 2 in U_{13}:

$$2^1 = 2,\ 2^2 = 4,\ 2^3 = 8,\ 2^4 = 3,\ 2^5 = 6,\ 2^6 = 12,$$
$$2^7 = 11,\ 2^8 = 9,\ 2^9 = 5,\ 2^{10} = 10,\ 2^{11} = 7,\ 2^{12} = 1.$$

Now, rather than writing $U_{13} = \{1, 2, 3, 4, \ldots, 12\}$, we write it as follows:

$$\begin{aligned} U_{13} &= \{\ 1,\ 2,\ 4,\ 8,\ 3,\ 6,\ 12,\ 11,\ 9,\ 5,\ 10,\ 7\ \} \\ &= \{\, 2^0,\, 2^1,\, 2^2,\, 2^3,\, 2^4,\, 2^5,\, 2^6,\, 2^7,\, 2^8,\, 2^9,\, 2^{10},\, 2^{11}\} \end{aligned}$$

where $2^0 = 2^{12} = 1$ is the multiplicative identity. When we express the elements of U_{13} as powers of 2 (its generator), we highlight the correspondence between $\mathbb{Z}_{12}$ and U_{13}. For example, $5 \in \mathbb{Z}_{12}$ corresponds to $2^5 \in U_{13}$, which we denote as $5 \leftrightarrow 2^5$. More generally, $k \in \mathbb{Z}_{12}$ corresponds to $2^k \in U_{13}$, denoted $k \leftrightarrow 2^k$.

We just saw that the *elements* of $\mathbb{Z}_{12}$ and U_{13} match up via the correspondence $k \leftrightarrow 2^k$. In the example below, we see that this correspondence respects their *operations* as well. Thus, in essence, the two groups are essentially the same. (We will formalize this notion of sameness soon.)

Example 13.8. In the calculations below, note that $2^{12} = 1$ in U_{13}.

- In $\mathbb{Z}_{12}$: $3 + 5 = 8$.
  In U_{13}: $2^3 \cdot 2^5 = 2^8$.

- In $\mathbb{Z}_{12}$: $9 + 7 = 16 = 12 + 4 = 0 + 4 = 4$.
  In U_{13}: $2^9 \cdot 2^7 = 2^{16} = 2^{12} \cdot 2^4 = 1 \cdot 2^4 = 2^4$.
- In $\mathbb{Z}_{12}$: The additive inverse of 4 is 8. In other words, $-4 = 8$, because $4 + 8 = 0$.
  In U_{13}: The multiplicative inverse of 2^4 is 2^8. In other words, $(2^4)^{-1} = 2^8$, because $2^4 \cdot 2^8 = 1$.

To generalize, we have $a + b = c$ in $\mathbb{Z}_{12}$ if and only if $2^a \cdot 2^b = 2^c$ in U_{13}.

Example 13.9. Let's return to our earlier question: What are all the generators of U_{13}? In Example 13.1, we found that the generators of $\mathbb{Z}_{12}$ are 1, 5, 7, 11. Using the correspondence $k \leftrightarrow 2^k$ between $\mathbb{Z}_{12}$ and U_{13}, the generators of U_{13} *should* be $2^1 = 2$, $2^5 = 6$, $2^7 = 11$, $2^{11} = 7$. We've already seen that 2 and 7 are generators of U_{13}. In an exercise at the end of the chapter, you will verify that 6 is a generator of U_{13}. Then, since $11 = 6^{-1}$ (as $6 \cdot 11 = 1$ modulo 13), we conclude that 11 is a generator of U_{13} as well.

13.4 Taking positive and *negative* powers of g

Example 13.10. Consider the element 4 in $U_{13} = \{1, 2, 3, 4, \ldots, 12\}$. The positive integer powers of 4 are

$$4^1 = 4,\ 4^2 = 3,\ 4^3 = 12,\ 4^4 = 9,\ 4^5 = 10,\ 4^6 = 1,\ 4^7 = 4,\ 4^8 = 3,\ 4^9 = 12,\ 4^{10} = 9, \ldots.$$

This sequence of positive powers repeats after $4^6 = 1$. For instance, $4^{11} = 4^{6+5} = 4^6 \cdot 4^5 = 1 \cdot 4^5 = 4^5$, and hence $4^{11} = 4^5$. We also have $4^0 = 1$ by definition. To compute the *negative* powers of 4, observe that $4 \cdot 10 = 1$ modulo 13 so that $4^{-1} = 10$ (i.e., the multiplicative inverse of 4 is 10). Thus, we have

$$\begin{aligned}
4^{-1} &= 10,\\
4^{-2} &= (4^{-1})^2 = 10^2 = 9,\\
4^{-3} &= (4^{-1})^3 = 10^3 = 12,\\
4^{-4} &= (4^{-1})^4 = 10^4 = 3,\\
4^{-5} &= (4^{-1})^5 = 10^5 = 4,\\
4^{-6} &= (4^{-1})^6 = 10^6 = 1.\\
&\vdots
\end{aligned}$$

But these have already been accounted for by the positive powers of 4. Therefore, the distinct integer powers of 4 are $4^0, 4^1, 4^2, 4^3, 4^4, 4^5$, or equivalently, 1, 4, 3, 12, 9, 10.

Example 13.11 (Chapter 12, Exercise #10 revisited)**.** Let g be an element of a group with $\text{ord}(g) = 6$, so that $g^6 = \varepsilon$. For instance, we can use $g = 4$ in U_{13} from Example 13.10. Then we have the following equalities/inequalities between the powers of g:

- $g^{32} = g^{20}$, because $g^{32} = g^{6\cdot 2+20} = (g^6)^2 \cdot g^{20} = \varepsilon^2 \cdot g^{20} = g^{20}$.
- $g^{862} \neq g^{800}$, because the difference of the exponents (i.e., $862 - 800 = 62$) is not divisible by 6. More rigorously, suppose for contradiction that $g^{862} = g^{800}$. Right-multiplying both sides by g^{-800} yields $g^{862} \cdot g^{-800} = g^{800} \cdot g^{-800}$, so that $g^{62} = \varepsilon$. Then Theorem 12.18 implies $6 \mid 62$, where $\text{ord}(g) = 6$. This is a contradiction, as 6 is not a divisor of 62. Thus, $g^{862} \neq g^{800}$.

- $g^{359} = g^{-241}$, because the difference of the exponents (i.e., $359 - (-241) = 600$) is divisible by 6. More rigorously, we have $g^{359} = g^{6\cdot 100-241} = (g^6)^{100} \cdot g^{-241} = \varepsilon^{100} \cdot g^{-241} = g^{-241}$.

The above example suggests the following theorem and proof.

Theorem 13.12. *Let g be an element of a group with $\operatorname{ord}(g) = n$. Then $g^k = g^\ell$ if and only if $n \mid (k - \ell)$.*

PROOF. We must prove two implications:

- If $g^k = g^\ell$, then $n \mid (k - \ell)$.
- If $n \mid (k - \ell)$, then $g^k = g^\ell$.

We will prove the first implication. The proof of the second implication is left for you as an exercise.

Assume $g^k = g^\ell$. Right-multiplying both sides by $g^{-\ell}$ yields $g^k \cdot g^{-\ell} = g^\ell \cdot g^{-\ell}$, so that $g^{k-\ell} = \varepsilon$. Then, Theorem 12.18 implies $n \mid (k - \ell)$ as desired. ∎

Remark. Recall that $n \mid (k - \ell)$ is equivalent to $k = \ell$ in $\mathbb{Z}_n$. Therefore, Theorem 13.12 above can be restated as follows: *Let g be an element of a group with $\operatorname{ord}(g) = n$. Then $g^k = g^\ell$ if and only if $k = \ell$ in $\mathbb{Z}_n$.*

In Example 13.10, we considered the integer powers of $4 \in U_{13}$, where $\operatorname{ord}(4) = 6$. Then Theorem 13.12 concludes: $4^k = 4^\ell$ if and only if $k = \ell$ in $\mathbb{Z}_6$. Thus 4^0, 4^1, 4^2, 4^3, 4^4, 4^5 are precisely the distinct integer powers of 4, since the exponents 0, 1, 2, 3, 4, 5 are precisely the distinct elements of $\mathbb{Z}_6$. The theorem below generalizes this example, and its proof is left for you as an exercise at the end of the chapter.

Theorem 13.13. *Let g be an element of a group with $\operatorname{ord}(g) = n$. Then the distinct integer powers of g are $\varepsilon, g^1, g^2, g^3, \ldots, g^{n-1}$, where $\varepsilon = g^0$.*

We've been talking about "the integer powers of g" throughout this chapter. Here is a notation and a definition that generalizes the concept.

Definition 13.14. Let g be an element of a multiplicative group. We define $\langle g \rangle$ to be the set of all integer powers of g. Thus, $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$, and we call the set $\langle g \rangle$ the *cyclic subgroup generated by g*.

Remark. As its name suggests, $\langle g \rangle$ is a subgroup of G, where G is the group containing the element g. You will prove this in an exercise at the end of the chapter.

Example 13.15. Let $4 \in U_{13}$ and recall from Example 13.10 that $\operatorname{ord}(4) = 6$. By definition, $\langle 4 \rangle = \{4^k \mid k \in \mathbb{Z}\}$ is the set of *all* integer powers of 4; i.e.,

$$\langle 4 \rangle = \{\ldots, 4^{-3}, 4^{-2}, 4^{-1}, 4^0, 4^1, 4^2, 4^3, \ldots\}.$$

Thus, it may seem that the set $\langle 4 \rangle$ contains infinitely many elements. But we saw in Example 13.10 that 4^0, 4^1, 4^2, 4^3, 4^4, 4^5 are precisely the distinct integer powers of 4. Therefore, we have

$$\langle 4 \rangle = \{4^0, 4^1, 4^2, 4^3, 4^4, 4^5\} = \{1, 4, 3, 12, 9, 10\}.$$

Example 13.16. Let g be an element of a multiplicative group with $\text{ord}(g) = 12$. By Theorem 13.13, the distinct integer powers of g are $\varepsilon, g^1, g^2, g^3, \ldots, g^{11}$, where $\varepsilon = g^0$. Thus $\langle g\rangle = \{\varepsilon, g^1, g^2, g^3, \ldots, g^{11}\}$. We have the correspondence $k \leftrightarrow g^k$ between $\mathbb{Z}_{12}$ and $\langle g\rangle$. Moreover, $\langle g\rangle$ behaves like $\mathbb{Z}_{12}$. For instance,

$$g^9 \cdot g^7 = g^{9+7} = g^{16} = g^{12+4} = g^{12} \cdot g^4 = \varepsilon \cdot g^4 = g^4,$$

so that $g^9 \cdot g^7 = g^4$, which is just like $9 + 7 = 4$ in $\mathbb{Z}_{12}$. This is the same calculation we did in Example 13.8, with $g \in \langle g\rangle$ instead of $2 \in U_{13}$, where $U_{13} = \langle 2\rangle$.

The following theorem is a generalization of Examples 13.15 and 13.16. Note how it's essentially a restatement of Theorem 13.13 using our new notation $\langle g\rangle$.

Theorem 13.17. *Let g be an element of a group with $\text{ord}(g) = n$. Then $\langle g\rangle = \{g^k \mid k \in \mathbb{Z}\}$ contains n distinct elements; namely $\langle g\rangle = \{\varepsilon, g^1, g^2, g^3, \ldots, g^{n-1}\}$, where $\varepsilon = g^0$.*

Example 13.18. In Example 13.4, we saw that 2 is a generator of U_{13}, because every element of U_{13} can be expressed as an integer power of 2. Using our new notation, we have $U_{13} = \langle 2\rangle$. Since 6, 7, 11 are also generators of U_{13} (see Example 13.9), we have $U_{13} = \langle 6\rangle = \langle 7\rangle = \langle 11\rangle$ as well.

Recall that a group is said to be *cyclic* if it has a generator. Then Example 13.18 above can be generalized to derive the following definition.

Definition 13.19 (Cyclic group)**.** A group G is said to be *cyclic* if there exists $g \in G$ such that $G = \langle g\rangle$. The element g is said to be a *generator* of G.

13.5 When the group operation is addition

How can we adapt the notation $\langle g\rangle = \{g^k \mid k \in \mathbb{Z}\}$ when the group operation is addition? At the beginning of this chapter, we saw that 1 is a generator of the *additive* group $\mathbb{Z}_{12}$, because

$$1 = 1,\ 1 + 1 = 2,\ 1 + 1 + 1 = 3,\ 1 + 1 + 1 + 1 = 4,\ \ldots,\ \underbrace{1 + 1 + \cdots + 1}_{12 \text{ terms}} = 0.$$

But we can also write this more succinctly as

$$1 \cdot 1 = 1,\ 2 \cdot 1 = 2,\ 3 \cdot 1 = 3,\ 4 \cdot 1,\ \ldots,\ 12 \cdot 1 = 0,$$

which gives us the definition $\langle 1\rangle = \{k \cdot 1 \mid k \in \mathbb{Z}\}$. The other generators of $\mathbb{Z}_{12}$ are 5, 7, and 11. Thus, the additive group $\mathbb{Z}_{12}$ is cyclic with $\mathbb{Z}_{12} = \langle 1\rangle = \langle 5\rangle = \langle 7\rangle = \langle 11\rangle$. Here is a generalization.

Definition 13.20. Let g be an element of an additive group. We define $\langle g\rangle$ to be the set of all integer multiples of g. Thus, $\langle g\rangle = \{k \cdot g \mid k \in \mathbb{Z}\}$, and we call the set $\langle g\rangle$ the *cyclic subgroup generated by g*.

Example 13.21. Consider the element $9 \in \mathbb{Z}_{12}$. We begin by finding the order of 9 in $\mathbb{Z}_{12}$:

$$9 = 9,\ 9 + 9 = 6,\ 9 + 9 + 9 = 3,\ \underbrace{9 + 9 + 9 + 9}_{4 \text{ terms}} = 0,$$

so that $\text{ord}(9) = 4$. Based on our work with multiplicative groups, we expect the cyclic subgroup $\langle 9 \rangle$ to contain 4 distinct elements; namely, $0 \cdot 9 = 0$, $1 \cdot 9 = 9$, $2 \cdot 9 = 6$, $3 \cdot 9 = 3$. Indeed, we have

$$\begin{aligned}\langle 9 \rangle &= \{k \cdot 9 \mid k \in \mathbb{Z}\} \\ &= \{\ldots, -4 \cdot 9, -3 \cdot 9, -2 \cdot 9, -1 \cdot 9, 0 \cdot 9, 1 \cdot 9, 2 \cdot 9, 3 \cdot 9, 4 \cdot 9, \ldots\} \\ &= \{\ldots, 0, 9, 6, 3, 0, 9, 6, 3, 0, \ldots\} \\ &= \{0, 9, 6, 3\}.\end{aligned}$$

Remark. By default, we will assume that the group operation is multiplication; and we'll employ the definition $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$. If we know that the operation is addition (when we're working with $\mathbb{Z}_m$, for example), then we'll use $\langle g \rangle = \{k \cdot g \mid k \in \mathbb{Z}\}$. The context should make it clear which definition to use.

Exercises

1. Verify that every element of $\mathbb{Z}_{12}$ can be expressed as a sum of 5's. (See Example 13.1.)

2. Recall from Example 13.2 that 8 is a generator of $\mathbb{Z}_{15}$.

   (a) Find a positive integer k such that
   $$\underbrace{8 + 8 + \cdots + 8}_{k \text{ terms}} = 11.$$
   (b) Repeat part (a), but this time, express 4 as a sum of 8's in $\mathbb{Z}_{15}$.
   (c) Repeat part (a), but this time, express 3 as a sum of 8's in $\mathbb{Z}_{15}$.
   (d) Repeat part (a), but this time, express 12 as a sum of 8's in $\mathbb{Z}_{15}$.
   (e) How are your answers in parts (a) and (b) related? What about your answers in parts (c) and (d)? Can you explain what's going on and why?

3. Complete the proof of Theorem 13.3 by proving its first implication.

4. Let G be a group and define the set $H = \{g^{-1} \mid g \in G\}$. Prove that $H = G$.

   Note: This theorem is used in Example 13.4. Do you see how?

5. (a) Verify that 5 is a generator of U_7. (See Example 13.5.)
   (b) Verify that 3 and 5 are the only generators of U_7.

6. Verify that 6 is a generator of U_{13}. (See Example 13.9.)

7. (a) Find all the generators of the additive group $\mathbb{Z}_{18}$.
   (b) Verify that 2 is a generator of U_{19}.
   (c) Using the results from parts (a) and (b), find all the generators of U_{19}.

   (This exercise is referenced in Chapter 14, Exercise #6 and Chapter 18, Exercise #19.)

8. Let g be an element of a group with $\text{ord}(g) = 18$.
   (a) How many distinct elements does $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$ contain? Explain how you know.
   (b) Find all the generators of $\langle g \rangle$.

9. Find the elements of U_{16} and verify that U_{16} does *not* have a generator.

10. (a) Find the elements of the group U_3 and determine whether or not U_3 is cyclic.
    (b) Repeat part (a) with U_5, U_{11}, U_{17}, U_{19}. (**Note:** We've already seen that U_7 and U_{13} are cyclic.)
    (c) Repeat part (a) with U_6, U_{10}, U_{14}, U_{22}, U_{26}, U_{34}, U_{38}.
    (d) What conjectures do you have?

11. Complete the proof of Theorem 13.12 by proving its second implication.

12. Consider the *additive* group $\mathbb{Z}_{12}$. We've seen that $\mathbb{Z}_{12} = \langle 1 \rangle$, where $\langle 1 \rangle = \{k \cdot 1 \mid k \in \mathbb{Z}\}$, i.e., the set of all *sums* we can make using 1. This means $\mathbb{Z}_{12}$ is cyclic with generator 1.
    (a) Compute $\langle m \rangle$ for all other $m \in \mathbb{Z}_{12}$.
    (b) Verify that each $\langle m \rangle$ is a subgroup of $\mathbb{Z}_{12}$.
    (c) How many *different* subgroups of $\mathbb{Z}_{12}$ did you obtain?
    (d) Are there other subgroups of $\mathbb{Z}_{12}$ that are *not* of the form $\langle m \rangle$? Why or why not?

13. Consider the *multiplicative* group U_{13}. We've seen that $U_{13} = \langle 2 \rangle$, where $\langle 2 \rangle = \{2^k \mid k \in \mathbb{Z}\}$, i.e., the set of all *products* we can make using 2. This means U_{13} is cyclic with generator 2.
    (a) Compute $\langle m \rangle$ for all other $m \in U_{13}$.
    (b) Verify that each $\langle m \rangle$ is a subgroup of U_{13}.
    (c) How many *different* subgroups of U_{13} did you obtain?
    (d) Are there other subgroups of U_{13} that are *not* of the form $\langle m \rangle$? Why or why not?

14. Consider the element 4 in the additive group $\mathbb{Z}_{30}$.
    (a) Find the elements of the cyclic subgroup $\langle 4 \rangle$.
    (b) Find all $m \in \mathbb{Z}_{30}$ such that $\langle m \rangle = \langle 4 \rangle$.

15. Let g be an element of a group with $\text{ord}(g) = 18$.
    (a) Find the smallest positive integer k such that $g^{-1} = g^k$.
    (b) Same as above, but with g^{-1} replaced by each of the following: g^{-3}, g^{532}, g^{-625}.
    (c) Is it possible that $g^7 = g^{12}$? Why or why not?

16. Let g be an element of a multiplicative group G. Prove that $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$ is a subgroup of G. (This exercise is referenced in Section 14.1.)

17. Prove Theorem 13.13.

18. **Prove:** If a group G is cyclic, then G is commutative.

    Hint: If G is cyclic, then it has a generator g. Thus, we can write $G = \langle g \rangle$.

    Note: The converse of this statement is false. See Exercise #19.

19. Give an example of a commutative group that is *not* cyclic. Explain your reasoning.

20. Fix an element h in a multiplicative group G. Recall that the *centralizer* of h in G is
$$C(h) = \{g \in G \mid gh = hg\}.$$
Prove that $\langle h \rangle \subseteq C(h)$.

21. Determine whether or not each direct product is cyclic. If it's cyclic, find a generator.

    (a) $\mathbb{Z}_{10} \times \mathbb{Z}_{12}$. (b) $\mathbb{Z}_6 \times \mathbb{Z}_7$. (c) $\mathbb{Z}_6 \times \mathbb{Z}_8$. (d) $\mathbb{Z}_{12} \times \mathbb{Z}_{20}$. (e) $\mathbb{Z}_{12} \times \mathbb{Z}_{35}$.

22. Consider the direct product $\mathbb{Z}_6 \times \mathbb{Z}_n$.

    (a) Find three values of n for which $\mathbb{Z}_6 \times \mathbb{Z}_n$ is cyclic.
    (b) Find three values of n for which $\mathbb{Z}_6 \times \mathbb{Z}_n$ is *not* cyclic.

23. Repeat Exercise #22, but with $\mathbb{Z}_7 \times \mathbb{Z}_n$.

24. **Prove:** The direct product $\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic if and only if $\gcd(m, n) = 1$. (This exercise is referenced in Chapter 14, Exercise #15.)

14

Cyclic Groups, Part II

In this chapter, we will continue our study of *cyclic groups*, i.e., groups that are generated by a single element. In particular, we will examine cyclic groups generated by an element with infinite order. We will compare one such group with the group of integers $\mathbb{Z}$ and conclude that the two groups are essentially the same. We will revisit this notion of sameness when we learn about *isomorphism* in the next unit.

The main focus of this chapter, however, is on the *subgroups* of cyclic groups. These subgroups have a beautiful structure to them. For example, $\mathbb{Z}_{12}$ (a cyclic group with 12 elements) has six subgroups, which are themselves cyclic, with the following number of elements: 1, 2, 3, 4, 6, 12. (Any conjectures?) By the end of this chapter, we will be able to precisely describe all the subgroups of any cyclic group.

14.1 Why negative powers are needed

In Example 13.15, we considered $4 \in U_{13}$ with $\text{ord}(4) = 6$. We found that $\langle 4 \rangle = \{4^k \mid k \in \mathbb{Z}\}$ contains just the elements 4^0, 4^1, 4^2, 4^3, 4^4, 4^5, even though the set $\langle 4 \rangle$, by definition, contains *all* integer powers of 4. In light of this and other similar examples that we studied in Chapter 13, it's natural to wonder why the *negative* powers of g are needed in $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$. The example below answers this question.

Example 14.1. Recall that $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$. We've seen that $\mathbb{R}^*$ contains all non-zero real numbers and is a group under multiplication. Let H be the smallest subgroup of $\mathbb{R}^*$ that contains the element 3. By closure, H must also contain $3 \cdot 3 = 9$, $3 \cdot 3 \cdot 3 = 27$, $3^4 = 81$, and all positive powers of 3. H must also contain the multiplicative identity 1. Moreover, H must contain the multiplicative inverses of

its elements, i.e., $\frac{1}{3}$, $\frac{1}{9}$, $\frac{1}{27}$, $\frac{1}{81}$, and so on. Therefore,

$$\begin{aligned} H &= \{\ldots,\ \tfrac{1}{81},\ \tfrac{1}{27},\ \tfrac{1}{9},\ \tfrac{1}{3},\ 1,\ 3,\ 9,\ 27, 81, \ldots\} \\ &= \{\ldots,\ 3^{-4},\ 3^{-3},\ 3^{-2},\ 3^{-1},\ 3^0,\ 3^1,\ 3^2,\ 3^3,\ 3^4,\ \ldots\} \\ &= \{3^k \mid k \in \mathbb{Z}\} \\ &= \langle 3 \rangle, \end{aligned}$$

so that $H = \langle 3 \rangle$, where $\langle 3 \rangle$ must contain both positive and *negative* powers of 3 (as well as $3^0 = 1$).

In Example 14.1 above, we notice that the integer powers of 3 are distinct from each other. This is because $\text{ord}(3)$ is infinite; i.e., there is no positive integer n such that $3^n = 1$ in $\mathbb{R}^*$. Then Theorem 12.29 allows us to conclude that $3^k = 3^\ell$ if and only if $k = \ell$ in $\mathbb{Z}$. Contrast this to the case of $4 \in U_{13}$. Since $\text{ord}(4) = 6$, Theorem 13.12 implies that $4^k = 4^\ell$ if and only if $k = \ell$ in $\mathbb{Z}_6$.

Before proceeding, we state the result of Chapter 13, Exercise #16 as a theorem. This theorem explains why we call $\langle g \rangle$ the cyclic *subgroup* generated by g.

Theorem 14.2. *Let g be an element of a group G. Then $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$ is a subgroup of G.*

In the cyclic subgroup $\langle 3 \rangle$ where $3 \in \mathbb{R}^*$, the negative powers of 3 are the multiplicative inverses of the positive powers of 3. While the examples below deal with elements of finite order, they will examine the role that multiplicative inverses play in cyclic subgroups.

Example 14.3. Consider the elements $4, 10 \in U_{13}$. We have $4 \cdot 10 = 1$ modulo 13, so that 4 and 10 are multiplicative inverses of each other. We have $\langle 4 \rangle = \{1, 4, 3, 12, 9, 10\}$ and $\langle 10 \rangle = \{1, 10, 9, 12, 3, 4\}$, so that $\langle 4 \rangle = \langle 10 \rangle$; i.e., the two sets contain the same elements.

Example 14.4. Consider the elements $3, 7 \in U_{20}$. We have $3 \cdot 7 = 1$ modulo 20, so that 3 and 7 are multiplicative inverses of each other. In an exercise at the end of the chapter, you will show that $\langle 3 \rangle = \langle 7 \rangle$.

Here is a generalization, which you will also prove in an exercise.

Theorem 14.5. *Let g be an element of a group. Then $\langle g \rangle = \langle g^{-1} \rangle$.*

14.2 Additive groups revisited

In Section 13.5, we adapted the notation $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$ when the group operation is addition. Consider again the additive group $\mathbb{Z}_{12}$ and recall that 1 is a generator, because:

$$1 = 1,\ \ 1+1 = 2,\ \ 1+1+1 = 3,\ \ 1+1+1+1 = 4,\ \ldots,\ \underbrace{1+1+\cdots+1}_{12 \text{ terms}} = 0.$$

Writing this more succinctly as

$$1 \cdot 1 = 1,\ \ 2 \cdot 1 = 2,\ \ 3 \cdot 1 = 3,\ \ 4 \cdot 1 = 4,\ \ldots,\ \ 12 \cdot 1 = 0,$$

we obtained the definition $\langle 1 \rangle = \{k \cdot 1 \mid k \in \mathbb{Z}\}$. In other words, $\langle 1 \rangle$ includes all integer *multiples* of 1, rather than all integer *powers* of 1. We conclude that $\mathbb{Z}_{12}$ is cyclic with $\mathbb{Z}_{12} = \langle 1 \rangle$.

For the finite group $\mathbb{Z}_{12} = \langle 1 \rangle$, we only need *positive* multiples of 1. But for the infinite group $\mathbb{Z}$, we do need positive and *negative* multiples of 1 (as well as $0 \cdot 1 = 0$). Indeed, we have

$$\begin{aligned}\mathbb{Z} &= \{\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\} \\ &= \{\ldots, -4 \cdot 1, -3 \cdot 1, -2 \cdot 1, -1 \cdot 1, 0 \cdot 1, 1 \cdot 1, 2 \cdot 1, 3 \cdot 1, 4 \cdot 1, \ldots\} \\ &= \{k \cdot 1 \mid k \in \mathbb{Z}\} \\ &= \langle 1 \rangle,\end{aligned}$$

so that $\mathbb{Z} = \langle 1 \rangle$. We also have $\mathbb{Z} = \langle -1 \rangle = \{k \cdot (-1) \mid k \in \mathbb{Z}\}$, so that 1 and -1 are both generators of the cyclic group $\mathbb{Z}$. Note that Theorem 14.5, when adapted for additive groups, says $\langle g \rangle = \langle -g \rangle$. Thus we would expect $\langle 1 \rangle = \langle -1 \rangle$. You should convince yourself that $\mathbb{Z}$ has no other generator.

14.3 $\langle 3 \rangle$ behaves "just like" $\mathbb{Z}$

In Example 14.1, we studied the subgroup $\langle 3 \rangle$ of the multiplicative group $\mathbb{R}^*$. Recall that

$$\begin{aligned}\langle 3 \rangle &= \{3^k \mid k \in \mathbb{Z}\} \\ &= \{\ldots, 3^{-4}, 3^{-3}, 3^{-2}, 3^{-1}, 3^0, 3^1, 3^2, 3^3, 3^4, \ldots\},\end{aligned}$$

where $3^0 = 1$ is the multiplicative identity. Since these integer powers of 3 are distinct, we can form a correspondence between $\mathbb{Z}$ and $\langle 3 \rangle$. For instance, $5 \in \mathbb{Z}$ corresponds to $3^5 \in \langle 3 \rangle$, which we denote $5 \leftrightarrow 3^5$. More generally, the *elements* of $\mathbb{Z}$ and $\langle 3 \rangle$ match up via the correspondence $k \leftrightarrow 3^k$. The example below shows how this correspondence also respects the *operations* of the two groups.

Example 14.6. We will illustrate how the multiplicative group $\langle 3 \rangle$ behaves just like the additive group $\mathbb{Z}$. To understand what that means, let's do some calculations in both groups.

- $3^{17} \cdot 3^{25} = 3^{17+25} = 3^{42}$ in $\langle 3 \rangle$, which is just like $17 + 25 = 42$ in $\mathbb{Z}$.
- The multiplicative identity of $\langle 3 \rangle$ is $3^0 = 1$, just like how the additive identity of $\mathbb{Z}$ is 0.
- The multiplicative inverse of 3^{17} is 3^{-17} in $\langle 3 \rangle$, just like how the additive inverse of 17 is -17 in $\mathbb{Z}$.

Soon, we'll formalize what it means for $\langle 3 \rangle$ to behave "just like" $\mathbb{Z}$.

Example 14.7. In Section 14.2, we saw that the generators of $\mathbb{Z}$ are 1 and -1. Using the correspondence $k \leftrightarrow 3^k$ between $\mathbb{Z}$ and $\langle 3 \rangle$, the generators of $\langle 3 \rangle$ should be $3^1 = 3$ and $3^{-1} = \frac{1}{3}$. In an exercise at the end of the chapter, you'll show that $\langle 3 \rangle = \left\langle \frac{1}{3} \right\rangle$, so that $\frac{1}{3}$ is indeed a generator of $\langle 3 \rangle$.

The next example is not about $\mathbb{Z}$ and $\langle 3\rangle$, but it highlights the notion of *sameness* between two groups.

Example 14.8. Consider the additive group $\mathbb{Z}_{40}$ and its cyclic subgroup $\langle 4\rangle = \{k \cdot 4 \mid k \in \mathbb{Z}\}$, containing the following elements:

$$\begin{aligned}\langle 4\rangle &= \{0\cdot 4,\ 1\cdot 4,\ 2\cdot 4,\ \ldots,\ 9\cdot 4\}\\ &= \{0,\ 4,\ 8,\ 12,\ 16,\ 20,\ 24,\ 28,\ 32,\ 36\}.\end{aligned}$$

We will find all the generators of $\langle 4\rangle$. Observe that $\operatorname{ord}(4) = 10$ in $\mathbb{Z}_{40}$. Thus, we form a correspondence between $\mathbb{Z}_{10}$ and $\langle 4\rangle$, where $k \in \mathbb{Z}_{10}$ corresponds to $k \cdot 4 \in \langle 4\rangle$. This correspondence also respects the *operations* of the two groups, i.e., addition modulo 10 and addition modulo 40. Given that $\operatorname{ord}(4) = 10$ in $\mathbb{Z}_{40}$ so that $10 \cdot 4 = 0$ in $\langle 4\rangle$, here are some calculations.

- In $\langle 4\rangle$, we have $5\cdot 4 + 7\cdot 4 = (5+7)\cdot 4 = (10+2)\cdot 4 = 10\cdot 4 + 2\cdot 4 = 0 + 2\cdot 4 = 2\cdot 4$ so that $\mathbf{5}\cdot 4 + \mathbf{7}\cdot 4 = \mathbf{2}\cdot 4$, which is just like $5 + 7 = 2$ in $\mathbb{Z}_{10}$.
- The additive identity of $\langle 4\rangle$ is $0 \cdot 4 = 0$, which is just like how the additive identity of $\mathbb{Z}_{10}$ is 0.
- The additive inverse of $3 \cdot 4$ is $7 \cdot 4$ in $\langle 4\rangle$, just like how the additive inverse of 3 is 7 in $\mathbb{Z}_{10}$.

Using Theorem 13.3, the generators of $\mathbb{Z}_{10}$ are 1, 3, 7, 9. Thus, the generators of $\langle 4\rangle$ should be $1\cdot 4 = 4,\ 3\cdot 4 = 12,\ 7\cdot 4 = 28,$ and $9\cdot 4 = 36$.

14.4 Subgroups of cyclic groups

Example 14.9. The additive group $\mathbb{Z}_8$ is cyclic with generator 1; i.e., $\mathbb{Z}_8 = \langle 1\rangle$. In Example 11.6, we found all subgroups of $\mathbb{Z}_8$, namely, $\{0\}$, $\{0, 4\}$, $\{0, 2, 4, 6\}$, and $\mathbb{Z}_8$. Moreover, each of these subgroups is cyclic:

- $\{0\} = \langle 0\rangle$.
- $\{0, 4\} = \langle 4\rangle$.
- $\{0, 2, 4, 6\} = \langle 2\rangle = \langle 6\rangle$.
- $\mathbb{Z}_8 = \langle 1\rangle = \langle 3\rangle = \langle 5\rangle = \langle 7\rangle$.

Example 14.10. Consider the cyclic group $\mathbb{Z}_{12} = \langle 1\rangle$. Proceeding as in Example 11.6, we can find all subgroups of $\mathbb{Z}_{12}$, namely, $\{0\}$, $\{0, 6\}$, $\{0, 4, 8\}$, $\{0, 3, 6, 9\}$, $\{0, 2, 4, 6, 8, 10\}$, and $\mathbb{Z}_{12}$. (See Chapter 11, Exercise #7.) Each of these subgroups is cyclic as well:

- $\{0\} = \langle 0\rangle$.
- $\{0, 6\} = \langle 6\rangle$.
- $\{0, 4, 8\} = \langle 4\rangle = \langle 8\rangle$.

- $\{0, 3, 6, 9\} = \langle 3\rangle = \langle 9\rangle$.
- $\{0, 2, 4, 6, 8, 10\} = \langle 2\rangle = \langle 10\rangle$.
- $\mathbb{Z}_{12} = \langle 1\rangle = \langle 5\rangle = \langle 7\rangle = \langle 11\rangle$.

Example 14.11. Consider the multiplicative group U_{13}. In Example 13.4, we saw that U_{13} is cyclic with generator 2. Moreover, we wrote U_{13} in the following manner:

$$\begin{aligned} U_{13} &= \langle 2\rangle \\ &= \{\, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}, 2^{11}\} \\ &= \{\, 1,\ \ 2,\ \ 4,\ \ 8,\ \ \ 3,\ \ \ 6,\ 12,\ 11,\ \ 9,\ \ 5,\ \ \ 10,\ \ \ 7\,\}. \end{aligned}$$

This highlights the correspondence between $\mathbb{Z}_{12}$ and U_{13}, where $k \in \mathbb{Z}_{12}$ corresponds to $2^k \in U_{13}$. Using the subgroups of $\mathbb{Z}_{12}$ from Example 14.10 and the correspondence $k \leftrightarrow 2^k$, we find the subgroups of U_{13}. Notice how each subgroup of U_{13} is cyclic as well.

$$\begin{aligned} \{0\} &= \langle 0\rangle, & \{2^0\} &= \langle 2^0\rangle = \{1\}, \\ \{0, 6\} &= \langle 6\rangle, & \{2^0, 2^6\} &= \langle 2^6\rangle = \{1, 12\}, \\ \{0, 4, 8\} &= \langle 4\rangle, & \{2^0, 2^4, 2^8\} &= \langle 2^4\rangle = \{1, 3, 9\}, \\ \{0, 3, 6, 9\} &= \langle 3\rangle, & \{2^0, 2^3, 2^6, 2^9\} &= \langle 2^3\rangle = \{1, 8, 12, 5\}, \\ \{0, 2, 4, 6, 8, 10\} &= \langle 2\rangle, & \{2^0, 2^2, 2^4, 2^6, 2^8, 2^{10}\} &= \langle 2^2\rangle = \{1, 4, 3, 12, 9, 10\}, \\ \mathbb{Z}_{12} &= \langle 1\rangle, & U_{13} &= \langle 2^1\rangle. \end{aligned}$$

The above examples suggest the following theorem.

Theorem 14.12 (Subgroups of cyclic groups)**.** *Let G be a cyclic group and let H be a subgroup of G. Then H is also cyclic.*

The proof of this theorem is rather complicated, so we will give some pre-proof remarks here. And afterwards, Example 14.13 will illustrate the technical details that are given in the proof.

- Since G is cyclic, it has a generator $g \in G$ such that $G = \langle g\rangle = \{g^k \mid k \in \mathbb{Z}\}$.
- If H contained only the identity element ε, then H is cyclic because $H = \langle \varepsilon\rangle$.
- Suppose H contains a non-identity element, say $h \in H$ where $h \neq \varepsilon$. Then $h \in G$ (since $H \subseteq G$), so that h can be written as an integer power of g. Let $h = g^k$ for some non-zero integer k. If k is negative, say $h = g^{-4}$ for instance, then $h^{-1} = (g^{-4})^{-1} = g^4$ is also in H, because H is a subgroup and hence contains inverses of its elements. Thus, H must contain a *positive* power of g.

PROOF. Suppose G is a cyclic group and let H be a subgroup of G. Since G is cyclic, it has a generator $g \in G$ such that $G = \langle g\rangle = \{g^k \mid k \in \mathbb{Z}\}$. If H contained only the identity element ε, then H is cyclic as $H = \langle \varepsilon\rangle$ and we're done with the proof. Thus we may assume that H contains a non-identity element. Then, as discussed in the pre-proof remarks above, H must contain a positive power of g. Let m be the *smallest*

positive integer such that $g^m \in H$. We show that $H = \langle g^m \rangle$ by proving the following inclusions:

(1) $\langle g^m \rangle \subseteq H$: Let $x \in \langle g^m \rangle$ so that $x = (g^m)^k$ for some integer k. We know $g^m \in H$. Thus, $(g^m)^k$ is also in H, because H is closed and contains inverses of its elements. Thus, $x \in H$ so that $\langle g^m \rangle \subseteq H$.

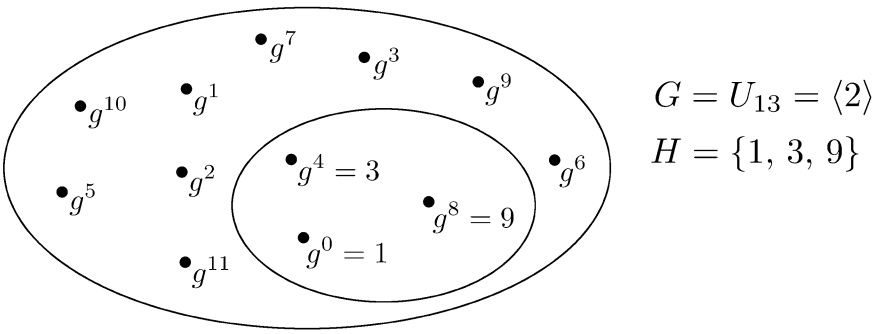
(2) $H \subseteq \langle g^m \rangle$: This is left for you to prove as an exercise at the end of the chapter.

Therefore, $H = \langle g^m \rangle$, implying that H is cyclic with generator g^m. ■

Example 14.13. Let's illustrate the technical details in the proof of Theorem 14.12 with a concrete example. Suppose $G = U_{13} = \langle 2 \rangle$, and let $H = \{1, 3, 9\}$ be its subgroup. With g = 2, we have

$$g^1 = 2 \notin H,\ g^2 = 4 \notin H,\ g^3 = 8 \notin H,\ g^4 = 3 \in H.$$

Thus, $m = 4$ is the smallest positive integer such that $g^m \in H$.



Since $g^4 \in H$, we know that $g^8 = (g^4)^2 \in H$ and $g^0 = g^{12} = (g^4)^3 \in H$; i.e., every integer power of g^4 is also contained in H. This is captured by the set inclusion $\langle g^m \rangle \subseteq H$.

The proof also shows that *only* the integer powers of g^4 are contained in H. In other words, if g^k is an element of H, then k must be a multiple of 4. This is captured by the set inclusion $H \subseteq \langle g^m \rangle$.

To conclude, we have $g^0 = 1, g^4 = 3, g^8 = 9$, so that $H = \{1, 3, 9\} = \{g^0, g^4, g^8\} = \langle g^4 \rangle$.

Example 14.14. We now know that subgroups of a cyclic group are also cyclic. And if a cyclic group G happens to be finite (like $\mathbb{Z}_{12}$, for instance), then we can say a bit more about its subgroups.

- Consider the cyclic group $\mathbb{Z}_8$ containing eight elements. Recall from Example 11.6 that the subgroups of $\mathbb{Z}_8$ are $\{0\}$, $\{0, 4\}$, $\{0, 2, 4, 6\}$, and $\mathbb{Z}_8$. The *sizes* (or numbers of elements) of these subgroups are 1, 2, 4, and 8, respectively, which are the divisors of 8.

- Consider the cyclic group $\mathbb{Z}_{12}$ containing 12 elements. Recall from Example 14.10 that the subgroups of $\mathbb{Z}_{12}$ are $\{0\}$, $\{0, 6\}$, $\{0, 4, 8\}$, $\{0, 3, 6, 9\}$, $\{0, 2, 4, 6, 8, 10\}$, and $\mathbb{Z}_{12}$. The sizes these subgroups are 1, 2, 3, 4, 6, and 12, respectively, which are the divisors of 12.

- Consider the cyclic group U_{13} containing 12 elements. Recall from Example 14.11 that U_{13} has six subgroups, one for each divisor of 12, whose sizes are 1, 2, 3, 4, 6, and 12.

Here is a generalization, whose proof is beyond the scope of this textbook.

Theorem 14.15 (Subgroups of *finite* cyclic groups)**.** *Suppose G is cyclic with n elements. Then G has a unique subgroup of size d for every divisor d of n, and those are the only subgroups of G.*

What makes the above theorem special is *not* that the size of a subgroup is a divisor of the size of the group. In fact, this is true for *any* group, even those that are not cyclic, which we will prove later in the textbook. Instead, what distinguishes cyclic groups is the fact that there is exactly one subgroup of size d for each divisor d of n (where n is the size of the group).

Example 14.16 (Non-example)**.** In Example 13.6, we saw that $U_{12} = \{1, 5, 7, 11\}$ is *not* cyclic. Observe that U_{12} has *three* subgroups of size 2, namely $\{1, 5\}$, $\{1, 7\}$, and $\{1, 11\}$. Thus, U_{12} does *not* have a unique subgroup of size d for every divisor d of 4 (where $n = 4$ is the size of U_{12}).

Exercises

1. Let $3, 7 \in U_{20}$. Compute $\langle 3 \rangle$ and $\langle 7 \rangle$ and show that they are equal. (See Example 14.4.)

2. In Example 14.1, we considered the cyclic subgroup $\langle 3 \rangle \subseteq \mathbb{R}^*$. Now consider the cyclic subgroup $\left\langle \frac{1}{3} \right\rangle \subseteq \mathbb{R}^*$. In particular, how do $\langle 3 \rangle$ and $\left\langle \frac{1}{3} \right\rangle$ compare?

3. Prove Theorem 14.5.

   Note: This is a set equality. So you must show $\langle g \rangle \subseteq \langle g^{-1} \rangle$ and $\langle g^{-1} \rangle \subseteq \langle g \rangle$.

4. (a) Explain why 1 and -1 are the only generators of the cyclic group $\mathbb{Z}$.

   (b) Explain why 3 and $\frac{1}{3}$ are the only generators of the cyclic subgroup $\langle 3 \rangle \subseteq \mathbb{R}^*$.

5. Find all subgroups of $\mathbb{Z}_{18}$.

6. It turns out that $U_{19} = \langle 2 \rangle$. (See Chapter 13, Exercise #7.) Find all subgroups of U_{19}.

7. Find all subgroups of $\mathbb{Z}$.

8. Explain why the additive group $\mathbb{Q}$ is *not* cyclic. (Here, $\mathbb{Q}$ denotes the set of rational numbers.)

9. Explain why the multiplicative group $\mathbb{R}^*$ is *not* cyclic.

10. Give an example of a group G satisfying the following conditions:

    - G is *not* cyclic.
    - Every proper subgroup of G is cyclic.

    Note: A *proper* subgroup of G is a subgroup that is not G itself.

11. Suppose $H = \left\{ \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \mid b \in \mathbb{Z}_m \right\}$. We know H is a subgroup of $G(\mathbb{Z}_m)$. (See Chapter 11, Exercise #11.) Now show that H is cyclic by finding a generator.

12. Consider the additive group $\mathbb{Q}$, i.e., the set of rational numbers. Define a subset $H \subseteq \mathbb{Q}$ given by $H = \left\{ \frac{3}{4}m + \frac{2}{7}n \mid m, n \in \mathbb{Z} \right\}$. We know that H is a subgroup of $\mathbb{Q}$. (See Chapter 11, Exercise #18.) Now show that H is cyclic by finding a generator.

13. Consider the cyclic subgroup $\langle 8 \rangle \subseteq \mathbb{Z}_{30}$.

    (a) Find the elements of $\langle 8 \rangle$.

    (b) Find all the generators of $\langle 8 \rangle$.

14. Consider the cyclic subgroup $\langle 8 \rangle \subseteq \mathbb{Z}$. (Notice how this is different from Exercise #13.)

    (a) Find the elements of $\langle 8 \rangle$.

    (b) Find all the generators of $\langle 8 \rangle$.

15. Find all the generators of the cyclic group $\mathbb{Z}_3 \times \mathbb{Z}_4$.

    Note: We know that $\mathbb{Z}_3 \times \mathbb{Z}_4$ is cyclic from Chapter 13, Exercise #24.

16. Repeat Exercise #15 with $\mathbb{Z}_6 \times \mathbb{Z}_7$.

17. Complete the proof of Theorem 14.12 by showing that $H \subseteq \langle g^m \rangle$.

    Hint: Let $h \in H$ so that $h = g^k$ for some $k \in \mathbb{Z}$. Now show that $m \mid k$.

18. Consider the group D_4, which contains eight elements.

    (a) Anita says, "D_4 is non-commutative. So it can't be cyclic." What might she mean?

    (b) Demonstrate how D_4 does *not* have a unique subgroup of size d for every divisor d of 8.

19. **Prove:** Let G be a group with finitely many elements. If $g \in G$, then $\text{ord}(g)$ is finite.

    Hint: What can you say about the sequence of elements $g^1, g^2, g^3, g^4, \ldots$?

20. **Prove:** Let G be a group, and let H be a subset of G that is nonempty and finite. If H is closed using the operation of G, then H is a subgroup of G.

21. Let G be a group, and let H be a subgroup of G. Fix an element $g \in G$, and define the set

$$gHg^{-1} = \{ghg^{-1} \mid h \in H\}.$$

    Prove: If H is cyclic, then gHg^{-1} is also cyclic.

    Recall: We showed in Chapter 12, Exercise #23 that gHg^{-1} is a subgroup of G.

22. Consider the group $\mathbb{Z}_{70}$.
    (a) Find the subgroup H of $\mathbb{Z}_{70}$ with size 7.
    (b) Find the subgroup K of $\mathbb{Z}_{70}$ with size 10.
    (c) Verify that $H \cap K = \{g \in \mathbb{Z}_{70} \mid g \in H \text{ and } g \in K\}$ contains just the additive identity 0.

We make two remarks about the proofs in Exercises #23 and #24:

- They must be proved *without* invoking Theorem 14.15.
- We will later see that these statements are true even when G is *not* cyclic.

23. **Prove:** Suppose G is cyclic with n elements, and let H be a subgroup of G with m elements. Then m is a divisor of n.

24. **Prove:** Suppose G is a cyclic group, and suppose H and K are subgroups of G containing m and n elements, respectively. If $\gcd(m, n) = 1$, then $H \cap K = \{\varepsilon\}$. (This exercise is referenced in Chapter 20, Exercise #5.)

Unit IV: Group Homomorphisms

The next four chapters are devoted to *functions* from a group to another group. Chapter 15 takes an in-depth look at the notion of functions, including what it means for a function to be *one-to-one* and *onto*. Then we delve into *isomorphisms*, which act as a mathematical bridge between a pair of groups and allow us to conclude that they're "essentially the same" by matching up their elements and operations. In Chapter 16, we learn a beautiful result (spoiler alert!) that every cyclic group is isomorphic to $\mathbb{Z}_n$ if it's finite with n elements or to $\mathbb{Z}$ if it's infinite.

Two chapters are devoted to *homomorphisms*, which are a generalization of isomorphisms. Chapter 18 ends with the following conjecture: A group homomorphism *partitions* the domain into equal-sized subsets. This conjecture foreshadows the study of *cosets*, which constitutes our final unit on group theory.

Here is a taste of what you'll be able to accomplish in this unit:

- Learn how to prove that a function is one-to-one and/or onto.
- Prove that an isomorphism preserves orders of group elements. In other words, if $\theta : G \to H$ is an isomorphism, then $\text{ord}(\theta(g)) = \text{ord}(g)$ for all $g \in G$.
- Recognize that homomorphisms provide a *unifying language* to describe familiar algebraic properties such as the exponent law $g^{a+b} = g^a \cdot g^b$ or the distributive law $6(a + b) = 6a + 6b$.

15

Functions

In Section 13.3, we formed a correspondence between the additive group $\mathbb{Z}_{12}$ and the multiplicative group U_{13}, where $k \in \mathbb{Z}_{12}$ corresponds to $2^k \in U_{13}$. (Note that 2 is one of the generators of U_{13}.) This correspondence not only matches up the elements of the two groups but also respects their operations. Thus, we said that the two groups are essentially the same.

In the next chapter, we'll study *isomorphisms*, which more precisely capture this notion of sameness. Then we'll explore *homomorphisms*, which are a generalization of isomorphisms. Both isomorphisms and homomorphisms are *functions*, which is the focus of this chapter. We will examine the various components of a function and what it means for a function to be *one-to-one* and *onto*.

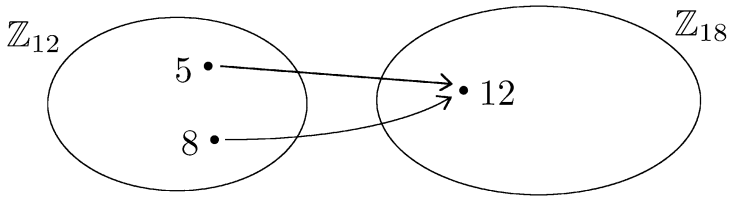
15.1 Domain and codomain

Example 15.1. Consider the function $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.

- The *domain* of the function is $\mathbb{Z}_{12}$. This is the set of all possible inputs into the function.
- The *codomain* of the function is $\mathbb{Z}_{18}$. This set contains all outputs (and possibly other elements).
- The *rule* of the function is $\gamma(a) = 6a$. Here, a is in the domain $\mathbb{Z}_{12}$ and $\gamma(a)$ is in the codomain $\mathbb{Z}_{18}$.

For instance, we have $\gamma(5) = 6 \cdot 5 = 30 = 12$, where the computation $30 = 12$ was done in $\mathbb{Z}_{18}$. Similarly, $\gamma(8) = 6 \cdot 8 = 48 = 12$, where the computation $48 = 12$ occurs in $\mathbb{Z}_{18}$. Thus, the elements 5 and 8 in the domain $\mathbb{Z}_{12}$ both map to the same element 12 in the

codomain $\mathbb{Z}_{18}$. This is depicted in the diagram below:



Example 15.2. Consider the function $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. For instance, we have $\varphi(43) = 3$, since $43 = 3 \pmod 5$. Similarly, $\varphi(-14) = 1$, because $-14 = 1 \pmod 5$. The domain of the function is $\mathbb{Z}$ and its codomain is $\mathbb{Z}_5$. The rule of the function is $\varphi(a) = a \pmod 5$.

The domain and codomain of a function can be the same set, as shown in Example 15.3 below.

Example 15.3. Consider the function $f : U_{35} \to U_{35}$ where $f(x) = 3x$ for all $x \in U_{35}$. Here, the domain and codomain are both U_{35}. For instance, we have $f(4) = 3 \cdot 4 = 12$, where 4 is in the domain and 12 is in the codomain. More generally, $x \in U_{35}$ is in the domain and $f(x) \in U_{35}$ is in the codomain.

15.2 One-to-one function

Example 15.4. Consider again the function $f : U_{35} \to U_{35}$ where $f(x) = 3x$ for all $x \in U_{35}$. Choose two different inputs from the domain U_{35}, say $a = 8$ and $b = 22$. Their corresponding outputs are $f(a) = 24$ and $f(b) = 31$, which are in the codomain U_{35}. These outputs are different from each other as well.

In fact, the following is true for all $a, b \in U_{35}$: If $a \neq b$, then $f(a) \neq f(b)$; i.e., *different inputs map to different outputs.* We can verify this by computing $f(x)$ for every $x \in U_{35}$ and seeing that all the outputs are different. Rather than taking this tedious approach (after all, there are 24 elements in U_{35}), we will prove this implication in a way that can be generalized to other scenarios.

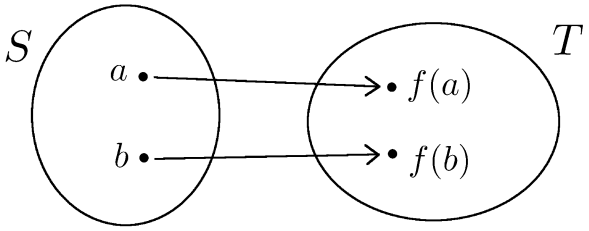
Specifically, we will prove its contrapositive; namely: If $f(a) = f(b)$, then $a = b$.

> Assume $f(a) = f(b)$, where $a, b \in U_{35}$. Then $3a = 3b$ in U_{35}. The multiplicative inverse of 3 is 12, where $3 \cdot 12 = 1$ and $12 \cdot 3 = 1$ modulo 35. Multiplying both sides of the equation $3a = 3b$ by 12, we obtain $12 \cdot (3a) = 12 \cdot (3b)$. Thus $(12 \cdot 3) \cdot a = (12 \cdot 3) \cdot b$, which implies $1 \cdot a = 1 \cdot b$. Therefore, $a = b$.

Example 15.4 illustrates the notion of a *one-to-one* function, which is defined below.

Definition 15.5 (One-to-one function)**.** Let $f : S \to T$ be a function from domain S to codomain T. We say f is *one-to-one* when it satisfies the following property for all a, $b \in S$: If $a \neq b$, then $f(a) \neq f(b)$.

Below is a picture that depicts a one-to-one function. Again, the key property is that different elements of the domain map to different elements of the codomain.



The definition of a one-to-one function contains the following implication: If $a \neq b$, then $f(a) \neq f(b)$. As we saw in Example 15.4, it is often effective to work with the contrapositive; namely: If $f(a) = f(b)$, then $a = b$.

Proof know-how. To show that f is one-to-one:

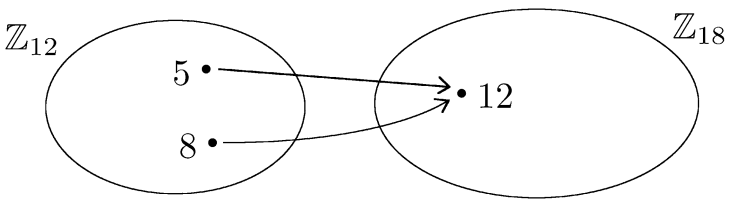
(1) Assume $f(a) = f(b)$, where $a, b \in S$.

(2) Show that $a = b$.

Example 15.6. Consider the function $\theta : \mathbb{R} \to \mathbb{R}$ where $\theta(x) = 5x - 2$ for all $x \in \mathbb{R}$. We will prove that this function is one-to-one:

> Suppose $\theta(a) = \theta(b)$, where $a, b \in \mathbb{R}$. Then $5a-2 = 5b-2$ in the codomain $\mathbb{R}$. Adding 2 to both sides of the equation yields $5a = 5b$. Multiplying both sides by $\frac{1}{5}$, we obtain $a = b$ as desired.

Example 15.7 (Non-example)**.** Consider the function $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$. In Example 15.1, we saw that $\gamma(5) = 12$ and $\gamma(8) = 12$. Thus, different inputs $5, 8 \in \mathbb{Z}_{12}$ both map to the same element 12 in the codomain $\mathbb{Z}_{18}$. Hence, γ is *not* one-to-one.



Example 15.8 (Non-example)**.** Consider the function $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. We have $\varphi(7) = 2$ and $\varphi(12) = 2$. Thus, different inputs $7, 12 \in \mathbb{Z}$ both map to the same element 2 in the codomain $\mathbb{Z}_5$. Hence, φ is *not* one-to-one.

Remark. In Examples 15.7 and 15.8, multiple elements of the domain are mapped to one element in the codomain. Such functions are often called *many-to-one* functions. This is in contrast to *one-to-one* functions, in which (at most) one element in the domain is mapped to one element in the codomain.

15.3 Onto function

In Example 15.1, the *codomain* of a function was described as follows: This set contains all outputs (and possibly other elements). The parenthetical remark "and possibly other elements" is key to understanding the notion of *onto* functions. We start with a non-example, i.e., a function that's not onto.

Example 15.9 (Non-example)**.** Consider the function $f : \mathbb{Z}_8 \to \mathbb{Z}_{20}$ where $f(a) = 5a$ for all $a \in \mathbb{Z}_8$. To find all possible outputs of f, let's evaluate the function at every element of the domain:

- $f(0) = f(4) = 0$.
- $f(1) = f(5) = 5$.
- $f(2) = f(6) = 10$.
- $f(3) = f(7) = 15$.

Not every element of the codomain $\mathbb{Z}_{20}$ is "hit" by the function f. For instance, there is no input element $a \in \mathbb{Z}_8$ such that $f(a) = 1$ (i.e., $1 \in \mathbb{Z}_{20}$ isn't "hit" by f). Thus, we say that f is *not* an onto function.

Example 15.10. Consider the function $f : U_{35} \to U_{35}$ where $f(x) = 3x$ for all $x \in U_{35}$. Choose any element from the codomain, say $y = 11 \in U_{35}$. Then, we find an element from the domain $x \in U_{35}$ such that $f(x) = y$. We have $x = 27$, since $f(27) = 3 \cdot 27 = 81 = 11$ modulo 35. As another example, let $y = 16$ be an element of the codomain U_{35}. Then $x = 17$ is the desired element from the domain U_{35}, since $f(17) = 3 \cdot 17 = 51 = 16$ modulo 35. We claim and will prove the following: Given any element y in the codomain U_{35}, we can find an element x in the domain U_{35} such that $f(x) = y$. As in Example 15.4, when we proved that f is one-to-one, the multiplicative inverse of 3 will play a role.

> Let $y \in U_{35}$ (the codomain). The multiplicative inverse of 3 is 12, where $3 \cdot 12 = 1$ and $12 \cdot 3 = 1$ modulo 35. Then let $x = 12y$, which is an element of U_{35} (the domain), since 12, $y \in U_{35}$ and U_{35} is closed under multiplication. We now verify that $f(x) = y$. We have $f(x) = f(12y) = 3 \cdot (12y) = (3 \cdot 12) \cdot y = 1 \cdot y = y$, so that $f(x) = y$ as desired.

Proof know-how. When developing the above proof, we used the "working backwards" technique that we've employed in the past (e.g., the proof of Theorem 13.3). Here, we wanted to find an element x in the domain U_{35} such that $f(x) = y$, or equivalently, $3x = y$. We worked backwards from this goal and solved the equation $3x = y$ for x by multiplying both sides by 3^{-1} (or 12), which yields $x = 12y$. As we've noted previously, this process of solving for x is scratch work and must *not* be included in the proof itself.

Example 15.10 illustrates the notion of an *onto* function, which is defined below.

Definition 15.11 (Onto function)**.** Let $f : S \to T$ be a function from domain S to codomain T. We say f is *onto* when for every $t \in T$ there exists $s \in S$ such that $f(s) = t$.

Informally, this definition says that every element of the codomain gets "hit" by an onto function f.

Proof know-how. To show that f is onto:

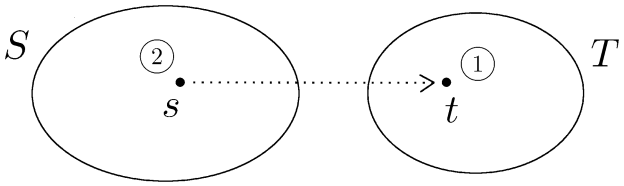
(1) Let $t \in T$.

(2) Find $s \in S$ such that $f(s) = t$. This s is usually expressed in terms of t.

(3) **(If necessary)** Verify that s is actually in set S.

(4) Show that s satisfies the desired property, namely $f(s) = t$.

The figure below depicts this Proof know-how. Note how the element t in the codomain T is chosen first, followed by the derivation and verification of the element s in the domain S.



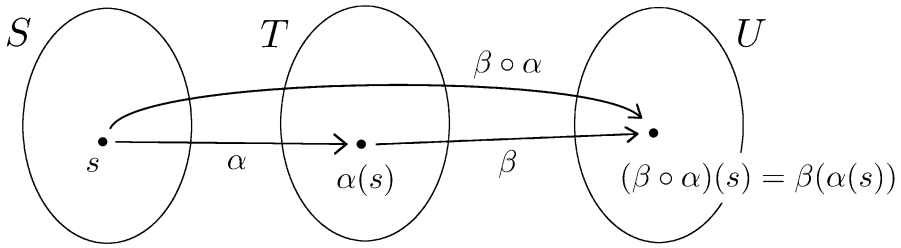
Example 15.12. Consider the function $\theta : \mathbb{R} \to \mathbb{R}$ where $\theta(x) = 5x - 2$ for all $x \in \mathbb{R}$. We will prove that this function is onto:

> Let $y \in \mathbb{R}$ (the codomain). Then let $x = \frac{1}{5}(y + 2)$, which is an element of $\mathbb{R}$ (the domain). We now verify that $\theta(x) = y$. We have
>
> $$\theta(x) = \theta\left(\tfrac{1}{5}(y + 2)\right) = 5 \cdot \tfrac{1}{5}(y + 2) - 2 = (y + 2) - 2 = y,$$
>
> so that $\theta(x) = y$ as desired.

We remark that the scratch work for this proof involved solving $\theta(x) = y$ or $5x - 2 = y$ for x.

Example 15.13 (Non-example)**.** Consider the function $\varphi : \mathbb{Z} \to \mathbb{Z}$ where $\varphi(n) = 5n-2$ for all $n \in \mathbb{Z}$. The rule of this function is the same as the one in Example 15.12, but the domain and codomain have been changed from $\mathbb{R}$ to $\mathbb{Z}$. We claim that there is no element $n \in \mathbb{Z}$ (the domain) such that $\varphi(n) = 0$. Suppose for contradiction that such an integer n exists. Then $\varphi(n) = 0$; i.e., $5n - 2 = 0$ and thus $5n = 2$. But this implies that 5 is a divisor of 2, which is a contradiction. Hence, $0 \in \mathbb{Z}$ (the codomain) isn't "hit" by φ, so that φ is *not* an onto function.

For the theorem below, consider the functions $\alpha : S \to T$, $\beta : T \to U$, and $\beta \circ \alpha : S \to U$. Note that $\beta \circ \alpha$ is a *composition* of α and β, where $(\beta \circ \alpha)(s) = \beta(\alpha(s))$ for all $s \in S$. (See the figure below for the visual depiction of $\beta \circ \alpha$.) This is just like composing symmetries or permutations, which are functions after all.



Theorem 15.14. *Consider the functions $\alpha : S \to T$, $\beta : T \to U$, and $\beta \circ \alpha : S \to U$. If $\beta \circ \alpha$ is one-to-one and α is onto, then β is one-to-one.*

PROOF. Assume that $\beta \circ \alpha$ is one-to-one and α is onto. To prove that β is one-to-one, assume $\beta(t_1) = \beta(t_2)$, where $t_1, t_2 \in T$. We must show that $t_1 = t_2$. Since α is onto, there exist $s_1, s_2 \in S$ such that $\alpha(s_1) = t_1$ and $\alpha(s_2) = t_2$. Thus, $\beta(t_1) = \beta(t_2)$ can be rewritten as $\beta(\alpha(s_1)) = \beta(\alpha(s_2))$, or equivalently, $(\beta \circ \alpha)(s_1) = (\beta \circ \alpha)(s_2)$. But since $\beta \circ \alpha$ is one-to-one, we obtain $s_1 = s_2$. Therefore, $t_1 = \alpha(s_1) = \alpha(s_2) = t_2$, so that $t_1 = t_2$ as desired. ■

Proof know-how. The above theorem and its proof highlight how to *use* the fact that a function is one-to-one or onto, rather than how to *prove* it. For instance, we were given that α is onto, and we had elements $t_1, t_2 \in T$, which is the codomain of α. Thus, we could deduce the existence of $s_1, s_2 \in S$, i.e., the domain of α, such that $\alpha(s_1) = t_1$ and $\alpha(s_2) = t_2$. Likewise, we had shown that $(\beta \circ \alpha)(s_1) = (\beta \circ \alpha)(s_2)$, where $s_1, s_2 \in S$, the domain of $\beta \circ \alpha$. Then the given fact that $\beta \circ \alpha$ is one-to-one allows us to conclude that $s_1 = s_2$.

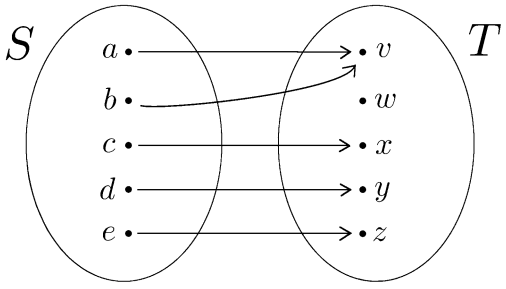
15.4 When domain and codomain have the same size

Given a function $f : S \to T$, it may be difficult to prove that f is onto, but easier to show that it's one-to-one (or vice versa). If the domain and codomain happen to have the same number of elements, however, then we only need to show one of the two and obtain the other "for free."

Example 15.15. Consider a function $f : S \to T$ where S and T each has 5 elements. Suppose f is one-to-one. We claim that f is onto. First, let $S = \{a, b, c, d, e\}$. Since f is one-to-one, $f(a)$, $f(b)$, $f(c)$, $f(d)$, $f(e)$ are 5 different elements in T. But T contains 5 elements, so the set $\{f(a), f(b), f(c), f(d), f(e)\}$ is equal to the set T. Thus, each element of T is equal to one of $f(a)$, $f(b)$, $f(c)$, $f(d)$, $f(e)$. Hence, f is onto.

Remark. In Example 15.15 above, the set $\{f(a), f(b), f(c), f(d), f(e)\}$, or more generally $\{f(s) \mid s \in S\}$, is called the *image* of f.

Example 15.16. Consider a function $f : S \to T$ where S and T each has 5 elements. Suppose f is onto. Then f must be one-to-one, for otherwise, we'd end up with a scenario like this, where (at most) 4 elements of T are "hit" by the function f, which contradicts the fact that f is onto.



Here is a generalization of the above examples.

Theorem 15.17. *Consider a function $f : S \to T$, where S and T are finite sets of the same size. Then f is one-to-one if and only if f is onto.*

PROOF. We must prove two implications:

- If f is one-to-one, then f is onto.
- If f is onto, then f is one-to-one.

We will prove the first implication. The proof of the second implication is left for you as an exercise.

Assume f is one-to-one. Let n be the size of both S and T. Consider the image of f, namely the set $I = \{f(s) \mid s \in S\}$. The set I contains n distinct elements, because f is one-to-one; i.e., the n different elements of S map to n different elements in T. Thus, $I \subseteq T$ and both I and T contain n elements, which implies that $I = T$. To show that f is onto, let $t \in T$. Then $t \in I$, because $I = T$. Thus, $t = f(s)$ for some $s \in S$. Therefore, f is onto as desired. ■

Remark. It's important that S and T are finite. Otherwise, even if S and T are the same set, the theorem does not hold. For example, consider $f : \mathbb{Z} \to \mathbb{Z}$ where $f(a) = 2a$ for all $a \in \mathbb{Z}$. In an exercise at the end of the chapter, you will explain why f is one-to-one, but not onto. In another exercise, you will find a function $f : \mathbb{Z} \to \mathbb{Z}$ which is onto, but not one-to-one.

Example 15.18. Consider the function $f : U_{35} \to U_{35}$ where $f(x) = 3x$ for all $x \in U_{35}$. In Example 15.4, we proved that f is one-to-one. Moreover, the domain and codomain are finite sets of the same size, since they're the same set U_{35}. Thus, we may use Theorem 15.17 to immediately conclude that f is onto.

Exercises

1. Consider the function $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.
   (a) In Example 15.1, we saw that $\gamma(5) = 12$ and $\gamma(8) = 12$. Find *all* $a \in \mathbb{Z}_{12}$ such that $\gamma(a) = 12$.
   (b) Find all $a \in \mathbb{Z}_{12}$ such that $\gamma(a) = 6$.
   (c) For each $b \in \mathbb{Z}_{18}$, find all $a \in \mathbb{Z}_{12}$ such that $\gamma(a) = b$.
   (d) Elizabeth says, "γ isn't one-to-one. It's actually four-to-one." What might she mean?

2. Consider the function $f : U_{35} \to U_{35}$ where $f(x) = 3x$ for all $x \in U_{35}$. Anita wonders, "How do we know that $f(x)$ actually is in the codomain U_{35} for all inputs $x \in U_{35}$?" How would you respond to her? (In other words, if $x \in U_{35}$, why must $f(x)$ also be in U_{35}?)

3. Consider the function $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$.

   (a) Find all $a \in \mathbb{Z}$ such that $\varphi(a) = 0$.

   (b) Find all $a \in \mathbb{Z}$ such that $\varphi(a) = 1$.

   (c) For each $b \in \mathbb{Z}_5$, find all $a \in \mathbb{Z}$ such that $\varphi(a) = b$.

4. Determine if each function is one-to-one. Explain your reasoning.

   (a) $f : \mathbb{Z} \to \mathbb{Z}$ where $f(n) = 4n + 1$ for all $n \in \mathbb{Z}$.

   (b) $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$ for all $x \in \mathbb{R}$. (Recall that $\mathbb{R}$ refers to the set of all real numbers.)

   (c) $f : U_7 \to U_7$ where $f(a) = a^3$ for all $a \in U_7$.

   (d) $f : U_{13} \to U_{13}$ where $f(a) = a^{-1}$ for all $a \in U_{13}$.

5. For each function in Exercise #4, determine if it is onto. Explain your reasoning.

6. Let $S = \{x \in \mathbb{R} \mid x \geq 0\}$. Determine if each function is one-to-one. Explain your reasoning.

   (a) $f : S \to \mathbb{R}$ where $f(x) = x^2$ for all $x \in S$.

   (b) $f : \mathbb{R} \to S$ where $f(x) = x^2$ for all $x \in \mathbb{R}$.

   (c) $f : S \to S$ where $f(x) = x^2$ for all $x \in S$.

7. For each function in Exercise #6, determine if it is onto. Explain your reasoning.

8. Consider $f : \mathbb{Z} \to \mathbb{Z}$ where $f(a) = 2a$ for all $a \in \mathbb{Z}$. Explain why f is one-to-one, but not onto.

9. Describe a function $f : \mathbb{Z} \to \mathbb{Z}$ that is onto, but not one-to-one.

10. **Prove:** Let G be a group and consider the function $\theta : G \to G$ where $\theta(g) = g^{-1}$ for all $g \in G$. Then θ is one-to-one and onto. (This exercise is referenced in Chapter 16, Exercise #14.)

11. Complete the proof of Theorem 15.17 by proving its second implication.

    Hint: Try proving its contrapositive.

12. Let $S = \{a, b, c\}$ and $T = \{v, w, x, y, z\}$.

    (a) How many different functions are there with domain S and codomain T? Explain your reasoning.

    (b) How many different functions are there with domain T and codomain S? Again, explain.

13. Define S and T as in Exercise #12. How many different *onto* functions are there with domain T and codomain S? Explain how you know.

14. Again, let S and T be as in Exercise #12. How many different *one-to-one* functions are there with domain S and codomain T? Explain.

15. Suppose S and T are sets with m and n elements, respectively.
    (a) If $m > n$, can there be a one-to-one function with domain S and codomain T? Why or why not?
    (b) Same as part (a), but with $m < n$.
    (c) If $m > n$, can there be an onto function with domain S and codomain T? Why or why not?
    (d) Same as part (c), but with $m < n$.

For Exercises #16 through #20, consider the functions $\alpha : S \to T$, $\beta : T \to U$, and $\beta \circ \alpha : S \to U$. (See the narrative surrounding Theorem 15.14 for details.)

16. **Prove:** If α and β are one-to-one, then $\beta \circ \alpha$ is one-to-one.

17. **Prove:** If α and β are onto, then $\beta \circ \alpha$ is onto.

18. (a) **Prove:** If $\beta \circ \alpha$ is one-to-one, then α is one-to-one.
    (b) Use a counterexample to show that this is false: If $\beta \circ \alpha$ is one-to-one, then β is one-to-one.

19. (a) **Prove:** If $\beta \circ \alpha$ is onto, then β is onto.
    (b) Use a counterexample to show that this is false: If $\beta \circ \alpha$ is onto, then α is onto.

20. **Prove:** If $\beta \circ \alpha$ is onto and β is one-to-one, then α is onto.

16

Isomorphisms

We first encountered groups that are *essentially the same* in Example 9.7, when we compared the tables of $\mathbb{Z}_3 = \{0, 1, 2\}$ under addition and $\{\varepsilon, r_{120}, r_{240}\}$ under the operation $\circ$ in D_3. Later in Section 13.3, we concluded that the additive group $\mathbb{Z}_{12}$ and the multiplicative group U_{13} are *essentially the same*, this time by forming the correspondence between $k \in \mathbb{Z}_{12}$ and $2^k \in U_{13}$. In this chapter, we will formalize this notion of sameness between groups by introducing *isomorphisms*.

The following quote by mathematician Henri Poincaré captures the spirit of this chapter:

> *Mathematicians do not deal in objects, but in relations among objects; they are free to replace some objects by others so long as the relations remain unchanged. Content to them is irrelevant: they are interested in form only.*

16.1 Groups $\mathbb{Z}_{12}$ and $\langle g \rangle$: Elements match up

Let's begin by revisiting an earlier example. Recall that given a group element g, the cyclic subgroup $\langle g \rangle$ is defined by $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$, i.e., the set of *all* integer powers of g.

Example 13.16 Let g be an element of a multiplicative group with $\text{ord}(g) = 12$. By Theorem 13.17, the distinct elements of $\langle g \rangle$ are given by $\langle g \rangle = \{\varepsilon, g^1, g^2, g^3, \ldots, g^{11}\}$, where $\varepsilon = g^0$. We thus have the correspondence $k \leftrightarrow g^k$ between $\mathbb{Z}_{12}$ and $\langle g \rangle$. Moreover, $\langle g \rangle$ behaves like $\mathbb{Z}_{12}$; e.g.,

$$g^9 \cdot g^7 = g^{9+7} = g^{16} = g^{12+4} = g^{12} \cdot g^4 = \varepsilon \cdot g^4 = g^4,$$

so that $g^9 \cdot g^7 = g^4$, which is just like $9 + 7 = 4$ in $\mathbb{Z}_{12}$.

Thus, we've said that the groups $\mathbb{Z}_{12}$ and $\langle g \rangle$ are *essentially the same*. To make this notion of sameness more precise, consider the function $\theta : \mathbb{Z}_{12} \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}_{12}$. For instance, $\theta(7) = g^7$.

Remark. The function θ associates to each $k \in \mathbb{Z}_{12}$ the element $g^k \in \langle g \rangle$. This is no different from the correspondence $k \leftrightarrow g^k$. But using the language of a *function*

here provides a powerful tool that will allow us to derive and prove various features of groups that are essentially the same.

We will prove that θ is one-to-one and onto. To show that θ is one-to-one, we begin by assuming that $\theta(a) = \theta(b)$, where $a, b \in \mathbb{Z}_{12}$. Then we show that $a = b$.

Theorem 16.1. *Let g be an element of a multiplicative group with $\text{ord}(g) = 12$. Consider the function $\theta : \mathbb{Z}_{12} \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}_{12}$. Then θ is one-to-one.*

PROOF. Assume $\theta(a) = \theta(b)$, where $a, b \in \mathbb{Z}_{12}$. Then $g^a = g^b$, so that Theorem 13.12 implies that $12 \mid (a - b)$. Thus $a = b$ in $\mathbb{Z}_{12}$, from which we conclude that θ is one-to-one. ■

To show that θ is onto, we let $y \in \langle g \rangle$. Then we must find $k \in \mathbb{Z}_{12}$ such that $\theta(k) = y$. Note that $y \in \langle g \rangle$ implies that y is an integer power of g. For example, suppose $y = g^{1,001}$. Since $1{,}001 = 12 \cdot 83 + 5$, we have

$$g^{1,001} = g^{12 \cdot 83 + 5} = (g^{12})^{83} \cdot g^5 = \varepsilon^{83} \cdot g^5 = g^5,$$

so that $g^{1,001} = g^5$. Hence with $k = 5 \in \mathbb{Z}_{12}$, we have $\theta(k) = \theta(5) = g^5 = g^{1,001} = y$. Here is a generalization, whose proof is left for you as an exercise at the end of the chapter.

Theorem 16.2. *Let g be an element of a multiplicative group with $\text{ord}(g) = 12$. Consider the function $\theta : \mathbb{Z}_{12} \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}_{12}$. Then θ is onto.*

Therefore, θ is both one-to-one and onto. Such a function is called a *bijection*. It means that each element of the domain $\mathbb{Z}_{12}$ corresponds with exactly one element of the codomain $\langle g \rangle$; and conversely, each element of the codomain corresponds with exactly one element of the domain. This implies that $\langle g \rangle$ has 12 distinct elements just like $\mathbb{Z}_{12}$, so that the elements in $\mathbb{Z}_{12}$ and $\langle g \rangle$ "match up" as follows:

$$\mathbb{Z}_{12} = \{0,\ 1,\ 2,\ 3,\ \ldots,\ 10,\ 11\},$$
$$\langle g \rangle = \{g^0, g^1, g^2, g^3, \ldots, g^{10}, g^{11}\} \quad (\text{where } g^0 = \varepsilon).$$

For instance, $\theta(7) = g^7$ so that $7 \in \mathbb{Z}_{12}$ and $g^7 \in \langle g \rangle$ match up as elements of these groups. Of course, we had already shown this in Example 13.16. But the bijection θ does more than just match up the elements of $\mathbb{Z}_{12}$ and $\langle g \rangle$, as we'll see in the next section.

16.2 Groups $\mathbb{Z}_{12}$ and $\langle g \rangle$: Operations match up

As in Section 16.1, let g be a group element with $\text{ord}(g) = 12$. Consider again the function $\theta : \mathbb{Z}_{12} \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}_{12}$. We saw that θ is a bijection; i.e., it's one-to-one and onto. Therefore, θ allows the elements in $\mathbb{Z}_{12}$ and $\langle g \rangle$ to "match up" with each other.

We also recall from Example 13.16 that the *operations* of $\mathbb{Z}_{12}$ and $\langle g \rangle$ match up as well. For instance, since $12 = 0$ in $\mathbb{Z}_{12}$ and $g^{12} = \varepsilon$ in $\langle g \rangle$, we have $9 + 7 = 4$ in $\mathbb{Z}_{12}$, which is just like $g^9 \cdot g^7 = g^{9+7} = g^4$ in $\langle g \rangle$. Addition in $\mathbb{Z}_{12}$ *feels like* multiplication in $\langle g \rangle$, and the function θ can more precisely capture this intuition.

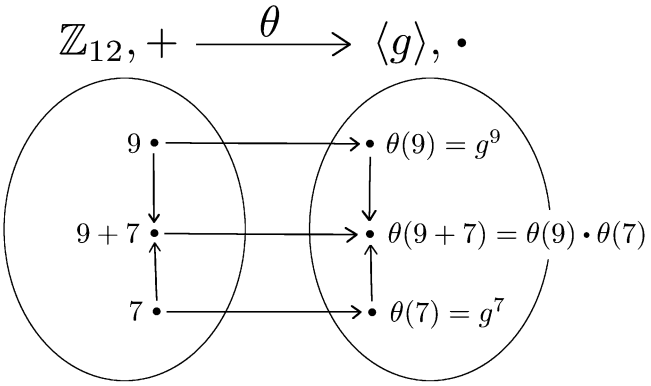
Using the law of exponents (i.e., $g^{a+b} = g^a \cdot g^b$), we have

$$\theta(9+7) = g^{9+7} = g^9 \cdot g^7 = \theta(9) \cdot \theta(7),$$

so that $\theta(9+7) = \theta(9) \cdot \theta(7)$.

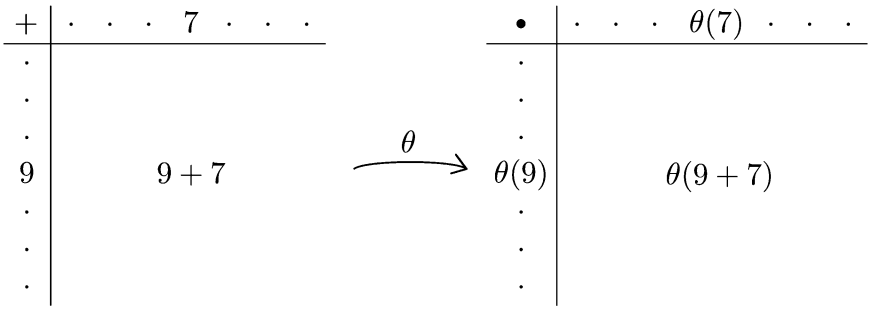
Let's dig deeper into the equation $\theta(9+7) = \theta(9) \cdot \theta(7)$. As shown in the diagram below:

- $\theta(9+7)$ means first add 9 and 7 in $\mathbb{Z}_{12}$ and then apply θ to the sum.
- $\theta(9) \cdot \theta(7)$ means first apply θ to each of 9 and 7 and then multiply them in $\langle g \rangle$.



Thus, $\theta(9+7) = \theta(9) \cdot \theta(7)$ may be interpreted as follows: It doesn't matter whether we first add in $\mathbb{Z}_{12}$ and then apply θ, or we first apply θ to each and then multiply in $\langle g \rangle$. The equation more precisely captures the notion that addition in $\mathbb{Z}_{12}$ *feels like* multiplication in $\langle g \rangle$.

Here is another way to interpret $\theta(9+7) = \theta(9) \cdot \theta(7)$. Imagine applying θ to every part of the addition table for $\mathbb{Z}_{12}$. Thus, 9, 7, and $9+7$ are mapped to $\theta(9)$, $\theta(7)$, and $\theta(9+7)$, respectively. We want the result to be the multiplication table for $\langle g \rangle$. For that to occur, we must have $\theta(9+7) = \theta(9) \cdot \theta(7)$.



As their elements and operations "match up," we say that the groups $\mathbb{Z}_{12}$ and $\langle g \rangle$ (where $\text{ord}(g) = 12$) are *isomorphic* and write $\mathbb{Z}_{12} \cong \langle g \rangle$. It's a more precise way of saying that the two groups are *essentially the same*. Below is a general definition.

Definition 16.3 (Group isomorphism)**.** Let G and H be groups with operations $*_G$ and $*_H$. A function $\theta : G \to H$ is an *isomorphism* if:

- θ is a bijection (i.e., one-to-one and onto) and
- θ is *operation preserving*; i.e., $\theta(a *_G b) = \theta(a) *_H \theta(b)$ for all $a, b \in G$.

We say that G is *isomorphic* to H and write $G \cong H$.

Remark. $G \cong H$ means they're essentially the same group.

We've seen that $\mathbb{Z}_{12} \cong \langle g \rangle$ where $\text{ord}(g) = 12$. Of course, there is nothing special about 12 here. Thus, we have the following theorem.

Theorem 16.4. *Let g be an element of a multiplicative group with $\text{ord}(g) = n$. Consider the function $\theta : \mathbb{Z}_n \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}_n$. Then, θ is an isomorphism so that $\mathbb{Z}_n$ is isomorphic to $\langle g \rangle$.*

Example 16.5. Recall that $\mathbb{R}$ is the additive group of all real numbers. Define $\mathbb{R}^{>0} = \{r \in \mathbb{R} \mid r > 0\}$, i.e., the set of all *positive* real numbers. In an exercise, you will show that $\mathbb{R}^{>0}$ is a group under multiplication. Define a function $\alpha : \mathbb{R} \to \mathbb{R}^{>0}$ where $\alpha(x) = 3^x$ for all $x \in \mathbb{R}$. We verify that α is an isomorphism:

- α is one-to-one: Assume $\alpha(a) = \alpha(b)$ where $a, b \in \mathbb{R}$. Then $3^a = 3^b$. By taking the log base 3 of both sides of $3^a = 3^b$, we obtain $\log_3 3^a = \log_3 3^b$, which simplifies to $a = b$.
- α is onto: Assume $y \in \mathbb{R}^{>0}$. Then let $x = \log_3 y \in \mathbb{R}$ so that $\alpha(x) = 3^x = 3^{\log_3 y} = y$.
- α is operation preserving: For $a, b \in \mathbb{R}$, we have $\alpha(a + b) = 3^{a+b} = 3^a \cdot 3^b = \alpha(a) \cdot \alpha(b)$.

Proof know-how. In Example 16.5 above, we rely on the "working backwards" technique yet again. To show that α is onto, we had to find $x \in \mathbb{R}$ such that $\alpha(x) = y$, or equivalently, $3^x = y$. Working backwards from this goal, we solved $3^x = y$ for x by taking the log base 3 of both sides. Thus, we found $x = \log_3 y$. As before, this process of solving for x is scratch work and does *not* belong in the proof. Instead, the focus of the proof (that α is onto) is showing that $\alpha(x) = y$ for $x = \log_3 y$.

Example 16.6 (Non-example)**.** Consider the additive group $\mathbb{R}$ and define the function $f : \mathbb{R} \to \mathbb{R}$, where $f(x) = x^3$ for all $x \in \mathbb{R}$. Then f is a bijection, i.e., it's one-to-one and onto, but not operation preserving. (You'll show these in an exercise at the end of the chapter.) Thus, f is *not* an isomorphism.

If group G is isomorphic to group H, then they are essentially the same. Thus, if one of them is commutative, then so is the other. Here is a theorem which captures that idea. Its converse is also true, but we'll leave that for you to prove as an exercise at the end of the chapter.

Theorem 16.7. *Let G and H be groups with operations $*_G$ and $*_H$. Suppose $\theta : G \to H$ is an isomorphism. If G is commutative, then H is commutative.*

PROOF. Assume G is commutative. To show that H is commutative, suppose $h_1, h_2 \in H$. We must show that $h_1 *_H h_2 = h_2 *_H h_1$. Since θ is onto, there exist $g_1, g_2 \in G$, such that $h_1 = \theta(g_1)$ and $h_2 = \theta(g_2)$. Therefore,

$$\begin{aligned} h_1 *_H h_2 &= \theta(g_1) *_H \theta(g_2) \\ &= \theta(g_1 *_G g_2) && \text{(since } \theta \text{ is operation preserving)} \\ &= \theta(g_2 *_G g_1) && \text{(since } G \text{ is commutative)} \\ &= \theta(g_2) *_H \theta(g_1) && \text{(since } \theta \text{ is operation preserving)} \\ &= h_2 *_H h_1. \end{aligned}$$

Thus $h_1 *_H h_2 = h_2 *_H h_1$ as desired. ■

Proof know-how. In Theorem 16.7 above, we were given that θ is an isomorphism. But to say that θ is an isomorphism means (1) θ is one-to-one, (2) θ is onto, and (3) θ is operation preserving. In a proof, you should specifically refer to one of these three properties, e.g., "Since θ is onto," rather than saying "Since θ is an isomorphism." (In fact, we did not use the fact that θ is one-to-one in the above proof.) This adds to the clarity of your argument.

16.3 Elements with infinite order revisited

Now, let g be a group element with infinite order. As we saw in Section 12.5, here are a couple of examples of such an element:

- $1 \in \mathbb{Z}$ where $\mathbb{Z}$ is a group under addition.
- $3 \in \mathbb{R}^*$ where $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$ is a group under multiplication.

We dig deeper into the second example. Consider the following cyclic subgroup of $\mathbb{R}^*$:

$$\begin{aligned} \langle 3 \rangle &= \{3^k \mid k \in \mathbb{Z}\} \\ &= \{\ldots, 3^{-4}, 3^{-3}, 3^{-2}, 3^{-1}, 3^0, 3^1, 3^2, 3^3, 3^4, \ldots\}. \end{aligned}$$

Then, define a function $\theta : \mathbb{Z} \to \langle 3 \rangle$ where $\theta(k) = 3^k$ for all $k \in \mathbb{Z}$. We first show that θ is one-to-one. Suppose $\theta(a) = \theta(b)$ where $a, b \in \mathbb{Z}$. Then $3^a = 3^b$. Since $\text{ord}(3)$ is infinite, Theorem 12.29 implies that $a = b$, showing that θ is one-to-one. To prove that θ is onto, let $y \in \langle 3 \rangle$ so that $y = 3^k$ for some integer k. Then $\theta(k) = 3^k = y$, and hence θ is onto.

To show that θ is operation preserving, let $a, b \in \mathbb{Z}$. Then

$$\theta(a + b) = 3^{a+b} = 3^a \cdot 3^b = \theta(a) \cdot \theta(b)$$

as desired. We showed that θ is a bijection and operation preserving, hence an isomorphism. In fact, there is nothing special about $3 \in \mathbb{R}^*$ here. Thus, we have the following theorem.

Theorem 16.8. *Let g be an element of a multiplicative group with infinite order. Define $\theta : \mathbb{Z} \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}$. Then θ is an isomorphism so that $\mathbb{Z}$ is isomorphic to $\langle g \rangle$.*

Remark. Theorems 16.4 and 16.8 together categorize all cyclic groups. Given a cyclic group $G = \langle g \rangle$, where g is a generator of G, we have the following:

- If $\text{ord}(g) = n$, then $\mathbb{Z}_n$ is isomorphic to G.
- If $\text{ord}(g)$ is infinite, then $\mathbb{Z}$ is isomorphic to G.

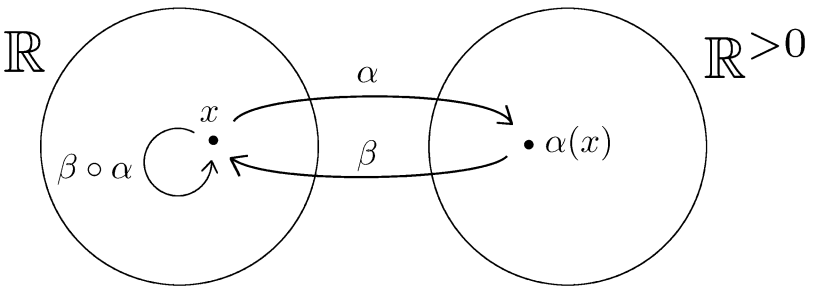
16.4 Inverse isomorphisms

Given groups G and H, we say that G is *isomorphic* to H if there exists an isomorphism $\alpha : G \to H$, i.e., a bijection that is operation preserving. This means that G and H are essentially the same group. But this relationship seems a bit one-sided, since in the isomorphism α, the domain is G and the codomain is H. If G is isomorphic to H, could we also say that H is isomorphic to G? Is there a corresponding isomorphism, say β, in which H is the domain and G is the codomain?

Before answering these questions, let's look at an example.

Example 16.9. In Example 16.5, we defined a function $\alpha : \mathbb{R} \to \mathbb{R}^{>0}$ where $\alpha(x) = 3^x$ for all $x \in \mathbb{R}$. (Note that $\mathbb{R}^{>0}$ is the set of all *positive* real numbers, which is a multiplicative group.) We also verified that α is an isomorphism. Define a function $\beta : \mathbb{R}^{>0} \to \mathbb{R}$ where $\beta(x) = \log_3 x$ for all $x \in \mathbb{R}^{>0}$. You'll show in an exercise at the end of the chapter that β is one-to-one and onto. We now show that β is operation preserving. For $a, b \in \mathbb{R}^{>0}$, we have $\beta(a \cdot b) = \log_3(a \cdot b) = \log_3 a + \log_3 b = \beta(a) + \beta(b)$. Here, $\log_3(a \cdot b) = \log_3 a + \log_3 b$ follows from one of the laws of logarithms. Therefore, β is an isomorphism.

Observe that β is not just *any* isomorphism from $\mathbb{R}^{>0}$ to $\mathbb{R}$. It corresponds to α in the following way. Consider the composite function $\beta \circ \alpha : \mathbb{R} \to \mathbb{R}$. For $x \in \mathbb{R}$, we have $(\beta \circ \alpha)(x) = \beta(\alpha(x)) = \beta(3^x) = \log_3(3^x) = x$, so that $(\beta \circ \alpha)(x) = x$. In other words, $\beta \circ \alpha$ is the identity function from the set $\mathbb{R}$ to itself. The figure below depicts this situation:



Similarly, the composite function $\alpha \circ \beta : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$ is the identity function from the set $\mathbb{R}^{>0}$ to itself. For $x \in \mathbb{R}^{>0}$, we have $(\alpha \circ \beta)(x) = \alpha(\beta(x)) = \alpha(\log_3 x) = 3^{\log_3 x} = x$, so that $(\alpha \circ \beta)(x) = x$.

Remark. In Example 16.9 above, we say that β is an *inverse isomorphism* of α and we write $\beta = \alpha^{-1}$. Conversely, α is an inverse isomorphism of β, which is denoted by $\alpha = \beta^{-1}$.

To generalize, suppose $\alpha : S \to T$ is a bijection from set S to set T. We define a function $\beta : T \to S$ as follows. Let $t \in T$. Since α is onto, there exists an element $s \in S$, such that $\alpha(s) = t$. We claim that s is the only such element, for if there exists another element $r \in S$ such that $\alpha(r) = t$, then $\alpha(r) = \alpha(s)$ as both are equal to t. And since α is one-to-one, we deduce that $r = s$. Thus for each $t \in T$, there exists a *unique* element $s \in S$ such that $\alpha(s) = t$. We can then unambiguously define $\beta(t) = s$.

Proof know-how. In defining the function β above, we wanted to show that s is the unique element in S such that $\alpha(s) = t$. We accomplished this by assuming that there are two such elements and showing that those two elements must be the same. (Compare this with the proof of Theorem 8.9, where we showed that a group has a unique identity element.)

Definition 16.10. Let $\alpha : S \to T$ be a bijection from set S to set T. Define a function $\beta : T \to S$ as follows. For each $t \in T$, let $\beta(t) = s$, where s is the unique element in S such that $\alpha(s) = t$. Then β is said to be an *inverse function* of α and is denoted $\beta = \alpha^{-1}$.

Example 16.11. Consider the function $\alpha : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$ where $\alpha(x) = \sqrt{x}$ for all $x \in \mathbb{R}^{>0}$. You'll show in an exercise that α is a bijection. To define $\beta = \alpha^{-1}$, let $t \in \mathbb{R}^{>0}$ (the codomain of α). For instance, say $t = 5$. We need to find $s \in \mathbb{R}^{>0}$ (the domain of α) such that $\alpha(s) = t$, or equivalently $\sqrt{s} = 5$. We have $s = 25$, so that $\beta(5) = 25$. As another example, let's say $t = \pi$. We must find s such that $\sqrt{s} = \pi$. Then $s = \pi^2$, so that $\beta(\pi) = \pi^2$. Indeed, we have a formula for β, namely, $\beta(t) = t^2$ for all $t \in \mathbb{R}^{>0}$.

Let's examine the composite function $\beta \circ \alpha : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$. For $x \in \mathbb{R}^{>0}$, we have

$$(\beta \circ \alpha)(x) = \beta(\alpha(x)) = \beta(\sqrt{x}) = (\sqrt{x})^2 = x,$$

so that $(\beta \circ \alpha)(x) = x$ for all $x \in \mathbb{R}^{>0}$. Similarly, $(\alpha \circ \beta)(x) = \alpha(\beta(x)) = \alpha(x^2) = \sqrt{x^2} = x$, where $\sqrt{x^2} = x$ since x is positive. Thus, both $\beta \circ \alpha$ and $\alpha \circ \beta$ are the identity function from $\mathbb{R}^{>0}$ to itself.

The function β in Definition 16.10 has the properties that we expect based on the examples that we've studied. The proof of the following theorem is left for you as an exercise.

Theorem 16.12. *Let $\alpha : S \to T$ be a bijection, and let $\beta : T \to S$ be an inverse function of α. Then:*

- *β is a bijection,*
- *$\beta \circ \alpha$ is the identity function from S to itself.*
- *$\alpha \circ \beta$ is the identity function from T to itself.*

Now, we're ready to answer the questions from the beginning of this section.

Theorem 16.13. *Let G and H be groups with operations $*_G$ and $*_H$, and suppose $\alpha : G \to H$ is an isomorphism. Let $\beta : H \to G$ be an inverse function of α. Then β is an isomorphism.*

PROOF. By Theorem 16.12, β is a bijection. Thus it suffices to show that β is operation preserving. Let $h_1, h_2 \in H$. We must show that $\beta(h_1 *_H h_2) = \beta(h_1) *_G \beta(h_2)$. Let $g_1 = \beta(h_1)$ and $g_2 = \beta(h_2)$, so that $\alpha(g_1) = h_1$ and $\alpha(g_2) = h_2$. Since α is operation preserving, $\alpha(g_1 *_G g_2) = \alpha(g_1) *_H \alpha(g_2) = h_1 *_H h_2$. Hence, α maps $g_1 *_G g_2 \in G$ to $h_1 *_H h_2 \in H$, which implies that $\beta(h_1 *_H h_2) = g_1 *_G g_2$. But $g_1 = \beta(h_1)$ and $g_2 = \beta(h_2)$, so that $\beta(h_1 *_H h_2) = \beta(h_1) *_G \beta(h_2)$, as desired. ■

Remark. In the above theorem, we call β an *inverse isomorphism* of α. Thus, we conclude that if G is isomorphic to H via an isomorphism α, then H is isomorphic to G via an inverse isomorphism β. Hence, we can simply say that G and H are isomorphic to each other.

Exercises

1. In Example 9.11, we compared the groups $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ under addition and $C = \{1, i, -1, -i\}$ under multiplication. Below are their group tables:

Table for $\mathbb{Z}_4$:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Table for C:

$*$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

   (a) Explain how these tables show that $\mathbb{Z}_4$ is isomorphic to C.
   (b) Find the order of each element of $\mathbb{Z}_4$ and each element of C. What conjectures do you have? (This exercise is referenced in Section 17.3.)

2. Prove Theorem 16.2.

3. Define $\mathbb{R}^{>0} = \{r \in \mathbb{R} \mid r > 0\}$, i.e., the set of all *positive* real numbers. Explain why $\mathbb{R}^{>0}$ is a group under multiplication. (See Example 16.5.)

4. Consider the additive group $\mathbb{R}$ and define the function $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^3$ for all $x \in \mathbb{R}$.
   (a) Show that f is a bijection; i.e., it's one-to-one and onto.
   (b) Explain why f is *not* operation preserving.

5. Consider the additive group $\mathbb{Z}$ and define the function $f : \mathbb{Z} \to \mathbb{Z}$ where $f(n) = n + 1$ for all $n \in \mathbb{Z}$.
   (a) **Prove:** f is a bijection.
   (b) Explain why f is *not* operation preserving.

6. Consider the function $\alpha : U_{35} \to U_{35}$ where $\alpha(x) = 3x$ for all $x \in U_{35}$. In Chapter 15, we showed that α is a bijection. Explain why α is *not* operation preserving.

7. Consider the function $\beta : \mathbb{R}^{>0} \to \mathbb{R}$ where $\beta(x) = \log_3 x$ for all $x \in \mathbb{R}^{>0}$. Show that β is a bijection; i.e., it's one-to-one and onto. (See Example 16.9.)

8. Consider the function $\alpha : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$ where $\alpha(x) = \sqrt{x}$ for all $x \in \mathbb{R}^{>0}$.
   (a) Show that α is a bijection. (See Example 16.11.)
   (b) Show that α is operation preserving.

   Thus, we may conclude that α is an isomorphism.

9. Consider the additive groups $\mathbb{Z}$ and $5\mathbb{Z}$. Prove that they are isomorphic by taking these steps:
   - Define a function $\theta : \mathbb{Z} \to 5\mathbb{Z}$ (or $\theta : 5\mathbb{Z} \to \mathbb{Z}$).
   - Show that θ is a bijection.
   - Show that θ is operation preserving.

10. Let $H = \left\{ \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \mid b \in \mathbb{Z}_m \right\}$. In Chapter 11, Exercise #11, we showed that H is a subgroup of $G(\mathbb{Z}_m)$. Prove that $\mathbb{Z}_m$ and H are isomorphic by taking these steps:
    - Define a function $\theta : \mathbb{Z}_m \to H$ (or $\theta : H \to \mathbb{Z}_m$).
    - Show that θ is a bijection.
    - Show that θ is operation preserving.

11. Fix a matrix $\gamma \in G(\mathbb{Z}_{10})$. Define a function $\theta : G(\mathbb{Z}_{10}) \to G(\mathbb{Z}_{10})$ such that $\theta(\alpha) = \gamma \cdot \alpha \cdot \gamma^{-1}$ for all $\alpha \in G(\mathbb{Z}_{10})$. Prove that θ is an isomorphism. (See Example 12.20.)

12. Let G be a group and fix an element $g \in G$. Define a function $\theta : G \to G$ such that $\theta(a) = gag^{-1}$ for all $a \in G$. Prove that θ is an isomorphism.

    Note: θ is often called the *conjugation function*. (See Definition 12.21.)

13. Consider the additive group $\mathbb{R}$ and define the function $\theta : \mathbb{R} \to \mathbb{R}$ where $\theta(x) = 3x$ for all $x \in \mathbb{R}$.
    (a) Show that θ is an isomorphism.
    (b) Change the domain and codomain of θ from $\mathbb{R}$ to $\mathbb{Z}$. Is θ still an isomorphism? Why or why not?

14. Let G be a group and consider the function $\theta : G \to G$ where $\theta(g) = g^{-1}$ for all $g \in G$. We've shown that θ is a bijection. (See Chapter 15, Exercise #10.) Now, prove that θ is operation preserving if and only if G is commutative.

15. (a) Explain why U_{10} is *not* isomorphic to U_7.
    (b) Explain why U_{10} is *not* isomorphic to U_8.

16. Each pair of groups has the same number of elements. Determine whether or not they're isomorphic.
    (a) $\mathbb{Z}_2 \times \mathbb{Z}_3$ and $\mathbb{Z}_6$.
    (b) $\mathbb{Z}_2 \times \mathbb{Z}_4$ and $\mathbb{Z}_8$.
    (c) $\mathbb{Z}_4 \times \mathbb{Z}_6$ and $\mathbb{Z}_3 \times \mathbb{Z}_8$.

17. Explain why the additive group $\mathbb{R}$ is *not* isomorphic to the multiplicative group $\mathbb{R}^*$.

    Hint: If $\theta : \mathbb{R} \to \mathbb{R}^*$ were an isomorphism, then there exists $r \in \mathbb{R}$ such that $\theta(r) = -1$. Explain why that would lead to a contradiction.

18. Let g be a group element with $\text{ord}(g) = 18$.

    (a) Find the order of each element of $\mathbb{Z}_{18}$.

    (b) Find the order of each element of $\langle g \rangle$.

    (c) What conjectures do you have?

19. Consider the function $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ where $\theta(n) = 4n$ for all $n \in \mathbb{Z}_{12}$.

    (a) Verify that θ is operation preserving. (Note that the operation of $\mathbb{Z}_{12}$ is addition.)

    (b) Is θ an isomorphism? Why or why not?

20. Fix $k \in \mathbb{Z}_{12}$ and consider the function $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ where $\theta(n) = k \cdot n$ for all $n \in \mathbb{Z}_{12}$.

    (a) Suppose $k = 5$. Verify that θ is an isomorphism.

    (b) For which values of k is θ an isomorphism?

    (c) What conjectures do you have?

21. Fix $k \in \mathbb{Z}_m$ and consider the function $\theta : \mathbb{Z}_m \to \mathbb{Z}_m$ where $\theta(n) = k \cdot n$ for all $n \in \mathbb{Z}_m$.

    (a) **Prove:** θ is an isomorphism if and only if $\gcd(k, m) = 1$.

    (b) **Prove:** *Every* isomorphism $\theta : \mathbb{Z}_m \to \mathbb{Z}_m$ has the formula $\theta(n) = k \cdot n$ with $\gcd(k, m) = 1$.

22. Consider the function $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ where $\theta(n) = 5n$ for all $n \in \mathbb{Z}_{12}$. We verified in Exercise #20 that θ is an isomorphism.

    (a) Find the order of $1 \in \mathbb{Z}_{12}$ (the domain of θ) and the order of $\theta(1) \in \mathbb{Z}_{12}$ (the codomain of θ).

    (b) Find the order of $2 \in \mathbb{Z}_{12}$ (the domain of θ) and the order of $\theta(2) \in \mathbb{Z}_{12}$ (the codomain of θ).

    (c) Find the order of $4 \in \mathbb{Z}_{12}$ (the domain of θ) and the order of $\theta(4) \in \mathbb{Z}_{12}$ (the codomain of θ).

    (d) What conjectures do you have?

    (This exercise is referenced in Section 17.3.)

23. Consider the function $\theta : U_{13} \to U_{13}$ where $\theta(n) = n^5$ for all $n \in U_{13}$.

    (a) Elizabeth says, "This is the U_{13} version of the function in Exercise #22. So it *should* be an isomorphism." What might she mean?

    (b) Verify that θ is indeed an isomorphism.

    Hint: To show that θ is a bijection, try computing the values $\theta(1)$, $\theta(2)$, $\theta(3)$, …, $\theta(12)$.

    (This exercise is referenced in Example 17.14.)

24. Consider the function $\theta : U_{13} \to U_{13}$ where $\theta(n) = n^5$ for all $n \in U_{13}$. We verified in Exercise #23 that θ is an isomorphism.

    (a) Find the order of $2 \in U_{13}$ (the domain of θ) and the order of $\theta(2) \in U_{13}$ (the codomain of θ).

    (b) Find the order of $4 \in U_{13}$ (the domain of θ) and the order of $\theta(4) \in U_{13}$ (the codomain of θ).

    (c) Find the order of $3 \in U_{13}$ (the domain of θ) and the order of $\theta(3) \in U_{13}$ (the codomain of θ).

    (d) What conjectures do you have?

    (This exercise is referenced in Section 17.3.)

25. **Prove:** Let G and H be groups with operations $*_G$ and $*_H$. Let $\theta : G \to H$ be an isomorphism. If H is commutative, then G is commutative.

    Note: You may *not* use the inverse isomorphism θ^{-1} in your proof.

26. Prove Theorem 16.12.

17

Homomorphisms, Part I

The next two chapters focus on *homomorphisms*. Given groups G and H, a homomorphism $\theta : G \to H$ is a function that is operation preserving, but not necessarily a bijection. Thus, an isomorphism is a special type of a homomorphism, just as a square is a special type of a rectangle. Since θ need not be a bijection, we cannot say that the elements in G and H "match up." Nonetheless, θ still preserves many essential group properties including the identity element, inverses, and the order of an element (sort of). For example, θ maps the identity element of G to the identity element of H, as we'll prove in this chapter.

Homomorphisms will also play a key role in our study of quotient groups in the next unit. In Chapter 18, we'll get a sneak preview of that role as we examine how a homomorphism $\theta : G \to H$ partitions the domain G into equal-sized subsets.

17.1 Group homomorphism

Let g be a group element with $\text{ord}(g) = 12$. We've seen how the groups $\mathbb{Z}_{12}$ and $\langle g \rangle$ are essentially the same. In particular, their *operations* match up. For instance, we have $9 + 7 = 4$ in $\mathbb{Z}_{12}$, which is just like $g^9 \cdot g^7 = g^{9+7} = g^4$ in $\langle g \rangle$. Hence, addition in $\mathbb{Z}_{12}$ *feels like* multiplication in $\langle g \rangle$.

In Chapter 16, as a way of articulating and understanding this observation more precisely, we studied the function $\theta : \mathbb{Z}_{12} \to \langle g \rangle$ where $\theta(k) = g^k$ for all $k \in \mathbb{Z}_{12}$. We found that θ is *operation preserving*; i.e.,

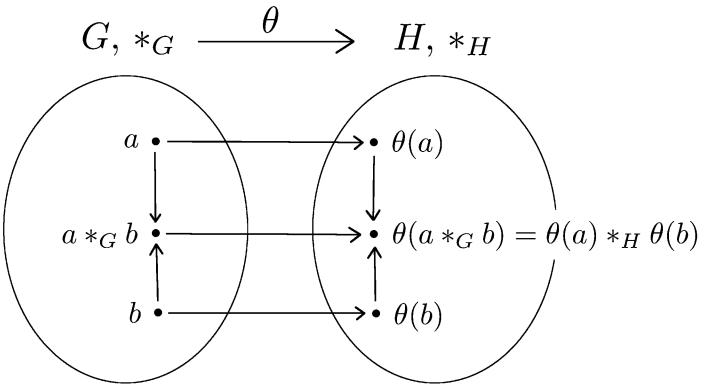
$$\theta(a + b) = g^{a+b} = g^a \cdot g^b = \theta(a) \cdot \theta(b)$$

for all $a, b \in \mathbb{Z}_{12}$. The function θ is an example of a *homomorphism*, which is defined below.

Definition 17.1 (Group homomorphism)**.** Let G and H be groups with operations $*_G$ and $*_H$. A function $\theta : G \to H$ is a *homomorphism* if it is operation preserving; i.e., $\theta(a *_G b) = \theta(a) *_H \theta(b)$ for all $a, b \in G$.

We elaborate on the equation $\theta(a *_G b) = \theta(a) *_H \theta(b)$. As shown in the diagram below:

- $\theta(a *_G b)$ means first multiply a and b in G and then apply θ to the product.
- $\theta(a) *_H \theta(b)$ means first apply θ to each of a and b and then multiply them in H.



Thus, $\theta(a*_G b) = \theta(a)*_H \theta(b)$ may be interpreted as follows: It doesn't matter whether we first multiply in G and then apply θ, or we first apply θ to each and then multiply in H. The equation more precisely captures the notion that the operation of G *feels like* the operation of H.

Remark. An isomorphism is a special type of a homomorphism that is also a bijection. In particular, all properties of homomorphisms that we will prove apply to isomorphisms as well.

As the example below shows, a homomorphism need not be an isomorphism.

Example 17.2. Consider $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. For instance, we have

$$\varphi(26 + 17) = \varphi(43) = 43 \pmod 5 = 3 \pmod 5.$$

Furthermore, $\varphi(26)+\varphi(17) = 26 \pmod 5 + 17 \pmod 5 = 1 \pmod 5 + 2 \pmod 5 = 3 \pmod 5$. Therefore, $\varphi(26+17) = \varphi(26)+\varphi(17)$. In an exercise at the end of the chapter, you'll explain why $\varphi(a + b) = \varphi(a) + \varphi(b)$ for all $a, b \in \mathbb{Z}$. In other words, it doesn't matter whether we first add in $\mathbb{Z}$ and then reduce mod 5, or first reduce each mod 5 and then add in $\mathbb{Z}_5$. Thus, φ is a homomorphism.

However, φ is *not* an isomorphism. An isomorphism is a bijection, i.e., one-to-one and onto. Thus, the domain and codomain of a bijection must have the same number of elements. This is *not* the case in φ, where the domain $\mathbb{Z}$ has infinitely many elements and the codomain $\mathbb{Z}_5$ has five elements.

Example 17.3. Consider the function $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$. For instance, we have $\gamma(7+10) = \gamma(5) = 6 \cdot 5 = 30 = 12$. Note here that the sum $7+10$ is computed in $\mathbb{Z}_{12}$ and the reduction $30 = 12$ is done in $\mathbb{Z}_{18}$. Moreover, $\gamma(7) + \gamma(10) = 6 \cdot 7 + 6 \cdot 10 = 42 + 60 = 102 = 12$, where the reduction $102 = 12$ is done in $\mathbb{Z}_{18}$. Thus, $\gamma(7 + 10) = \gamma(7) + \gamma(10)$, as both sides are equal to 12 (in $\mathbb{Z}_{18}$).

For $a, b \in \mathbb{Z}_{12}$, we have $\gamma(a+b) = 6(a+b) = 6a+6b = \gamma(a)+\gamma(b)$, so that $\gamma(a+b) = \gamma(a)+\gamma(b)$. The key step $6(a+b) = 6a+6b$ is due to the distributive law. Thus, γ is a homomorphism.

Example 17.4. Consider the function $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. For instance, we have

$$\lambda(5\cdot 2) = \lambda(10) = 10^3 = 1{,}000 = 12 \text{ and } \lambda(5)\cdot\lambda(2) = 5^3\cdot 2^3 = 125\cdot 8 = 1{,}000 = 12.$$

Thus, $\lambda(5\cdot 2) = \lambda(5)\cdot\lambda(2)$, as both sides are equal to 12.

For $a, b \in U_{13}$, we have $\lambda(a\cdot b) = (a\cdot b)^3 = a^3\cdot b^3 = \lambda(a)\cdot\lambda(b)$, so that $\lambda(a\cdot b) = \lambda(a)\cdot\lambda(b)$. The key step $(a\cdot b)^3 = a^3\cdot b^3$ is due to an exponent law, which holds because multiplication in U_{13} is commutative. Therefore, λ is a homomorphism.

Example 17.5. Consider the function $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det\alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. Recall that matrices in $G(\mathbb{Z}_{10})$ have multiplicative inverses, and thus Theorem 7.21 ensures that $\det\alpha \in U_{10}$. For a concrete example, let $\alpha, \beta \in G(\mathbb{Z}_{10})$ where $\alpha = \left[\begin{smallmatrix}1&5\\2&7\end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix}5&2\\3&1\end{smallmatrix}\right]$. Then $\delta(\alpha\cdot\beta) = \delta\left(\left[\begin{smallmatrix}0&7\\1&1\end{smallmatrix}\right]\right) = \det\left[\begin{smallmatrix}0&7\\1&1\end{smallmatrix}\right] = 3$, and $\delta(\alpha)\cdot\delta(\beta) = \det\left[\begin{smallmatrix}1&5\\2&7\end{smallmatrix}\right]\cdot\det\left[\begin{smallmatrix}5&2\\3&1\end{smallmatrix}\right] = 7\cdot 9 = 3$. Thus, $\delta(\alpha\cdot\beta) = \delta(\alpha)\cdot\delta(\beta)$, as both sides are equal to 3.

For matrices $\alpha, \beta \in G(\mathbb{Z}_{10})$, we have $\delta(\alpha\cdot\beta) = \det(\alpha\cdot\beta) = \det\alpha\cdot\det\beta = \delta(\alpha)\cdot\delta(\beta)$, so that $\delta(\alpha\cdot\beta) = \delta(\alpha)\cdot\delta(\beta)$. The key property is $\det(\alpha\cdot\beta) = \det\alpha\cdot\det\beta$; i.e., the determinant of the matrix product is equal to the product of the determinants. (See Theorem 7.24.) Thus, δ is a homomorphism.

These above examples illustrate how homomorphisms provide a *unifying language* to describe familiar algebraic properties. The opening example of this chapter shows that an exponent law $g^{a+b} = g^a\cdot g^b$ can be written as $\theta(a+b) = \theta(a)\cdot\theta(b)$. Similarly, Examples 17.2 through 17.5 show the following:

- Reduction law $a+b \pmod 5 = a \pmod 5 + b \pmod 5$ can be written as $\varphi(a+b) = \varphi(a)+\varphi(b)$.
- Distributive law $6(a+b) = 6a+6b$ can be written as $\gamma(a+b) = \gamma(a)+\gamma(b)$.
- Exponent law $(ab)^3 = a^3b^3$ can be written as $\lambda(a\cdot b) = \lambda(a)\cdot\lambda(b)$.
- Determinant property $\det(\alpha\beta) = \det\alpha\cdot\det\beta$ can be written as $\delta(\alpha\cdot\beta) = \delta(\alpha)\cdot\delta(\beta)$.

Example 17.6. Let G and H be groups. Consider the function $\theta : G \to H$ where $\theta(g) = \varepsilon_H$ for all $g \in G$. (Here, ε_H is the identity element of H.) For $a, b \in G$, we have $\theta(a\cdot b) = \varepsilon_H$, since $a\cdot b \in G$ by closure and θ maps every element of G to ε_H. Moreover, $\theta(a)\cdot\theta(b) = \varepsilon_H\cdot\varepsilon_H = \varepsilon_H$, so that $\theta(a\cdot b) = \theta(a)\cdot\theta(b)$, as both sides are equal to ε_H. Hence, θ is a homomorphism, and it's typically called the *trivial homomorphism*.

Example 17.7 (Non-example)**.** Consider the additive group $\mathbb{R}$ and define the function $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^3$ for all $x \in \mathbb{R}$. We have $f(2+4) = f(6) = 6^3 = 216$ and $f(2)+f(4) = 2^3+4^3 = 8+64 = 72$, so that $f(2+4) \neq f(2)+f(4)$. This counterexample suffices to show that f is *not* a homomorphism.

17.2 Properties of homomorphisms

In this section, we will derive and prove various properties of homomorphisms. Recall that since an isomorphism is a special type of a homomorphism, all properties in this section also apply to isomorphisms.

We will continue to work with these homomorphisms from Section 17.1:

- $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$.
- $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.
- $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$.
- $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$.

Example 17.8. We have $\varphi(0) = 0 \pmod 5$, so that the homomorphism φ maps the identity element of $\mathbb{Z}$ to the identity element of $\mathbb{Z}_5$. (Note that both $\mathbb{Z}$ and $\mathbb{Z}_5$ are *additive* groups.) Similarly, we have the following:

- $\gamma(0) = 6 \cdot 0 = 0$, where 0 is the additive identity in both $\mathbb{Z}_{12}$ and $\mathbb{Z}_{18}$.
- $\lambda(1) = 1^3 = 1$, where 1 is the multiplicative identity of U_{13}.
- $\delta(\varepsilon) = 1$ where $\varepsilon = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is the identity matrix in $G(\mathbb{Z}_{10})$ and 1 is the multiplicative identity of U_{10}.

The calculations in Example 17.8 suggest the following theorem. Note that the sentence "Let $\theta : G \to H$ be a group homomorphism" means that G and H are groups and θ is a homomorphism from G to H.

Theorem 17.9. *Let $\theta : G \to H$ be a group homomorphism. Then θ maps the identity element of G to the identity element of H; i.e., $\theta(\varepsilon_G) = \varepsilon_H$.*

PROOF. In G, we have $\varepsilon_G \cdot \varepsilon_G = \varepsilon_G$. Applying θ to both sides, we obtain $\theta(\varepsilon_G \cdot \varepsilon_G) = \theta(\varepsilon_G)$. Since θ is operation preserving, we have $\theta(\varepsilon_G \cdot \varepsilon_G) = \theta(\varepsilon_G) \cdot \theta(\varepsilon_G)$, so that $\theta(\varepsilon_G) \cdot \theta(\varepsilon_G) = \theta(\varepsilon_G)$.

In H, we have $\theta(\varepsilon_G) \cdot \varepsilon_H = \theta(\varepsilon_G)$. Thus, $\theta(\varepsilon_G) \cdot \theta(\varepsilon_G) = \theta(\varepsilon_G) \cdot \varepsilon_H$, since both are equal to $\theta(\varepsilon_G)$. Then, left cancellation in H yields $\theta(\varepsilon_G) = \varepsilon_H$ as desired. ■

Proof know-how. Key to this proof was to write $\varepsilon_G \cdot \varepsilon_G = \varepsilon_G$ in G and $\theta(\varepsilon_G) \cdot \varepsilon_H = \theta(\varepsilon_G)$ in H. These applications of the "inserting the identity" technique eventually allowed us to use left cancellation. (Compare this with the proof of Theorem 9.6.)

Example 17.10. Consider the homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. Observe that 23 and -23 are additive inverses in the domain $\mathbb{Z}$, while $\varphi(23) = 3 \pmod 5$ and $\varphi(-23) = 2 \pmod 5$ are additive inverses in the codomain $\mathbb{Z}_5$. Thus, $\varphi(-23)$ is the additive inverse of $\varphi(23)$ in $\mathbb{Z}_5$, which is written symbolically as $\varphi(-23) = -\varphi(23)$.

Example 17.11. Consider the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. Then $\alpha = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 8 & 3 \\ 5 & 4 \end{smallmatrix}\right]$ are multiplicative inverses in the domain $G(\mathbb{Z}_{10})$. (You should verify that $\alpha \cdot \beta = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $\beta \cdot \alpha = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.) We have

$\delta(\alpha) = 3$ and $\delta(\beta) = 7$, which are multiplicative inverses in the codomain U_{10}, since $3 \cdot 7 = 1$ modulo 10. By writing $\alpha^{-1} = \beta$, we see that $\delta(\alpha^{-1})$ is the multiplicative inverse of $\delta(\alpha)$, which is written symbolically as $\delta(\alpha^{-1}) = \delta(\alpha)^{-1}$.

In the exercises at the end of the chapter, you will work with similar examples involving the homomorphisms γ and λ. For now, here is the generalization based on Examples 17.10 and 17.11.

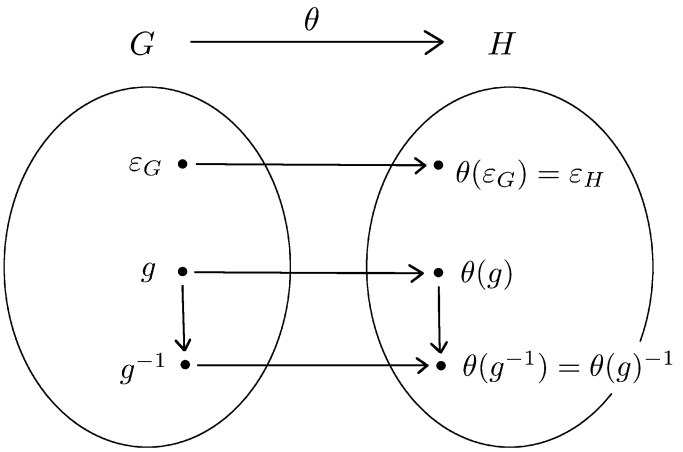
Theorem 17.12. *Let $\theta : G \to H$ be a group homomorphism. Then $\theta(g^{-1}) = \theta(g)^{-1}$ for all $g \in G$.*

PROOF. Let $g \in G$. Then $g \cdot g^{-1} = \varepsilon_G$ in G. Applying θ to both sides, we get $\theta(g \cdot g^{-1}) = \theta(\varepsilon_G)$ and thus $\theta(g) \cdot \theta(g^{-1}) = \varepsilon_H$. We can similarly show that $\theta(g^{-1}) \cdot \theta(g) = \varepsilon_H$. Therefore, $\theta(g^{-1})$ is the multiplicative inverse of $\theta(g)$ in H. In other words, $\theta(g^{-1}) = \theta(g)^{-1}$ as desired. ■

Proof know-how. In the above proof, the goal was to show that $\theta(g^{-1}) = \theta(g)^{-1}$, which translates to "$\theta(g^{-1})$ is the multiplicative inverse of $\theta(g)$." To show this, we multiplied $\theta(g)$ by $\theta(g^{-1})$ and vice versa and verified that the products equal the identity element ε_H. This approach is similar to the one employed in the proof of the socks-shoes property, i.e., Theorem 8.11.

Remark. As usual, the above theorems and proofs assume that G and H are multiplicative groups. If one (or both) is an additive group, then the theorem (and its proof) would have to be adjusted accordingly. For instance, if G is multiplicative and H is additive, Theorem 17.12 becomes $\theta(g^{-1}) = -\theta(g)$ for all $g \in G$.

The diagram below summarizes Theorems 17.9 and 17.12. Note that $\theta(g^{-1}) = \theta(g)^{-1}$ means it doesn't matter whether we first invert in G and then apply θ, or first apply θ and then invert in H.



Once again, let $\theta : G \to H$ be a group homomorphism. Theorem 17.9 says $\theta(\varepsilon_G) = \varepsilon_H$. For any $g \in G$, we have $g^0 = \varepsilon_G$ by the definition of g^0 in G. Moreover, $\theta(g)$ is some element of H and we have $\theta(g)^0 = \varepsilon_H$ by the definition of taking the 0^{th} power of an element in H. Thus, $\theta(\varepsilon_G) = \varepsilon_H$ can be rewritten as $\theta(g^0) = \theta(g)^0$, and this is true for all $g \in G$.

Theorem 17.12 says $\theta(g^{-1}) = \theta(g)^{-1}$ for all $g \in G$. Generalizing from these two examples, we will now show that $\theta(g^k) = \theta(g)^k$ for all integer exponents k (both positive and negative). For instance, suppose $k = 3$. Since θ is operation preserving, we have $\theta(g^3) = \theta(g \cdot g \cdot g) = \theta(g) \cdot \theta(g) \cdot \theta(g) = \theta(g)^3$ for all $g \in G$. Likewise, if k is any positive integer, then

$$\theta(g^k) = \theta(\underbrace{g \cdot g \cdot g \cdot \cdots \cdot g}_{k \text{ terms}}) = \underbrace{\theta(g) \cdot \theta(g) \cdot \theta(g) \cdot \cdots \cdot \theta(g)}_{k \text{ terms}} = \theta(g)^k.$$

Let's consider a negative exponent. Suppose $k = -3$. Using the interpretation $g^{-3} = (g^{-1})^3$, we have $\theta(g^{-3}) = \theta((g^{-1})^3) = \theta(g^{-1} \cdot g^{-1} \cdot g^{-1})$. Moreover,

$$\theta(g^{-1} \cdot g^{-1} \cdot g^{-1}) = \theta(g^{-1}) \cdot \theta(g^{-1}) \cdot \theta(g^{-1}),$$

since θ is operation preserving. Using $\theta(g^{-1}) = \theta(g)^{-1}$, we obtain

$$\theta(g^{-1}) \cdot \theta(g^{-1}) \cdot \theta(g^{-1}) = \theta(g)^{-1} \cdot \theta(g)^{-1} \cdot \theta(g)^{-1}.$$

Finally, $\theta(g)^{-1} \cdot \theta(g)^{-1} \cdot \theta(g)^{-1} = (\theta(g)^{-1})^3 = \theta(g)^{-3}$, where we used the interpretation $h^{-3} = (h^{-1})^3$ for any element $h \in H$, with $h = \theta(g)$. Putting these altogether gives $\theta(g^{-3}) = \theta(g)^{-3}$ as desired.

In an exercise at the end of the chapter, you'll generalize from the $k = -3$ case to show that the relationship holds for all negative exponents k. Thus, we have the following result.

Theorem 17.13. *Let $\theta : G \to H$ be a group homomorphism. Then $\theta(g^k) = \theta(g)^k$ for all $g \in G$ and $k \in \mathbb{Z}$.*

17.3 Order of an element

Suppose $\theta : G \to H$ is a group isomorphism. In Chapter 16, Exercises #1(b), #22, and #24, we compared the order of an element $g \in G$ and the order of the corresponding element $\theta(g) \in H$. Here is an example.

Example 17.14. Consider the function $\theta : U_{13} \to U_{13}$ where $\theta(n) = n^5$ for all $n \in U_{13}$. In Chapter 16, Exercise #23, we showed that θ is an isomorphism. Let $g = 4 \in U_{13}$ (the domain of θ). We have

$$4^1 = 4, \;\; 4^2 = 3, \;\; 4^3 = 12, \;\; 4^4 = 9, \;\; 4^5 = 10, \;\; 4^6 = 1$$

so that $\text{ord}(g) = 6$. Moreover, $\theta(g) = 4^5 = 1{,}024 = 10$, which is an element of the codomain U_{13}. To find its order, we compute as follows:

$$10^1 = 10, \;\; 10^2 = 9, \;\; 10^3 = 12, \;\; 10^4 = 3, \;\; 10^5 = 4, \;\; 10^6 = 1$$

which implies $\text{ord}(\theta(g)) = 6$.

The proof of the following theorem is left to you as an exercise. This result shouldn't be too surprising, since $G \cong H$ (i.e., G is isomorphic to H) means that these two groups are essentially the same. Thus, the corresponding elements $g \in G$ and $\theta(g) \in H$ should have the same order.

Theorem 17.15. *Let $\theta : G \to H$ be a group isomorphism. Then* $\text{ord}(\theta(g)) = \text{ord}(g)$ *for all $g \in G$.*

What if we have a homomorphism, but not necessarily an isomorphism? Then $\text{ord}(\theta(g)) = \text{ord}(g)$ no longer holds, as we'll see in the next example. But in what way is the order of an element still "preserved"?

Example 17.16. Recall the function $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$. In Example 17.3, we verified that γ is a homomorphism. We have the following data:

- $\text{ord}(10) = 6$ in $\mathbb{Z}_{12}$ and $\text{ord}(\gamma(10)) = \text{ord}(6) = 3$ in $\mathbb{Z}_{18}$.
- $\text{ord}(7) = 12$ in $\mathbb{Z}_{12}$ and $\text{ord}(\gamma(7)) = \text{ord}(6) = 3$ in $\mathbb{Z}_{18}$.
- $\text{ord}(8) = 3$ in $\mathbb{Z}_{12}$ and $\text{ord}(\gamma(8)) = \text{ord}(12) = 3$ in $\mathbb{Z}_{18}$.
- $\text{ord}(6) = 2$ in $\mathbb{Z}_{12}$ and $\text{ord}(\gamma(6)) = \text{ord}(0) = 1$ in $\mathbb{Z}_{18}$.

In all these cases, we see that $\text{ord}(\gamma(g))$ is a divisor of $\text{ord}(g)$.

Theorem 17.17. *Let $\theta : G \to H$ be a group homomorphism. Then* $\text{ord}(\theta(g)) \mid \text{ord}(g)$ *for all $g \in G$.*

PROOF. Let $g \in G$. Suppose $m = \text{ord}(g)$ so that $g^m = \varepsilon_G$. Then, $\theta(g)^m = \theta(g^m) = \theta(\varepsilon_G) = \varepsilon_H$, so that $\theta(g)^m = \varepsilon_H$. Thus by Theorem 12.18, the order of $\theta(g)$ is a divisor of m; i.e., $\text{ord}(\theta(g)) \mid \text{ord}(g)$. ■

Proof know-how. In the above proof, we let $m = \text{ord}(g)$ so that $g^m = \varepsilon_G$. But the key to the proof is to apply the exponent m to the *other element*, namely $\theta(g)$, and show that $\theta(g)^m = \varepsilon_H$. This allows us to conclude that $\text{ord}(\theta(g))$ is a divisor of m. (Compare this with the proof of Theorem 12.23.)

Example 17.18. Suppose $\theta : U_{16} \to \mathbb{Z}_5$ is a homomorphism. Here's a fact about the orders of elements in each group, whose verification is left up to you:

- For $U_{16} = \{1, 3, 5, 7, 9, 11, 13, 15\}$, we have $\text{ord}(g) = 1$, 2, or 4 for all $g \in U_{16}$.
- For $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$, we have $\text{ord}(h) = 1$ or 5 for all $h \in \mathbb{Z}_5$.

Let $g \in U_{16}$ and $h = \theta(g) \in \mathbb{Z}_5$. Theorem 17.17 implies that $\text{ord}(h)$ is a divisor of $\text{ord}(g)$. Then $\text{ord}(h) \neq 5$, since 5 is not a divisor of 1, 2, or 4. Hence, $\text{ord}(h) = 1$ so that $h = \varepsilon_H$. Thus, $\theta(g) = \varepsilon_H$ for all $g \in U_{16}$. We conclude that the only homomorphism from U_{16} to $\mathbb{Z}_5$ is the trivial homomorphism.

Exercises

1. Consider the function $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$.
   (a) Explain why $\varphi(a + b) = \varphi(a) + \varphi(b)$ for all $a, b \in \mathbb{Z}$. (See Example 17.2.)
   (b) While you're at it, explain why $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ for all $a, b \in \mathbb{Z}$.
   Note: This will be useful when we study *ring homomorphisms* later in the textbook.

2. In Examples 17.3, 17.4, and 17.5, we saw that the functions below are homomorphisms. Explain why each is *not* a bijection, hence *not* an isomorphism.
   (a) $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.
   (b) $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$.
   (c) $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$.

3. Consider the function $f : D_4 \to \mathbb{R}^*$ where
$$f(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is a rotation,} \\ -1 & \text{if } \sigma \text{ is a reflection.} \end{cases}$$
   Recall that D_4 is the group of symmetries of a square with four rotations and four reflections and that $\mathbb{R}^* = \{a \in \mathbb{R} \mid a \text{ has a multiplicative inverse}\}$ is a group under multiplication. (Group table for D_4 is in Appendix B.)
   (a) Compute $f(r_{180} \circ r_{270})$ and $f(r_{180}) \cdot f(r_{270})$ and verify that they are equal.
   (b) Compute $f(r_{90} \circ v)$ and $f(r_{90}) \cdot f(v)$ and verify that they are equal.
   (c) Compute $f(d \circ h)$ and $f(d) \cdot f(h)$ and verify that they are equal.
   (d) Explain why $f(\sigma \circ \tau) = f(\sigma) \cdot f(\tau)$ for all $\sigma, \tau \in D_4$. (Thus, f is a homomorphism.)

   (This exercise is referenced in Chapter 18, Exercise #7 and Example 24.18.)

4. Consider $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. Let $\alpha = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$.
   (a) Compute α^{-2} in two ways: via the interpretations $\alpha^{-2} = (\alpha^{-1})^2$ and $\alpha^{-2} = (\alpha^2)^{-1}$.
   (b) Use the result in part (a) to compute $\delta(\alpha^{-2})$.
   (c) Compute $\delta(\alpha)$ and use that to compute $\delta(\alpha)^{-2}$.
   (d) Compare $\delta(\alpha^{-2})$ with $\delta(\alpha)^{-2}$. Is the outcome surprising?

5. Consider again the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. For $\beta = \left[\begin{smallmatrix} 4 & 7 \\ 5 & 1 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$, it turns out that $\operatorname{ord}(\beta) = 6$ in $G(\mathbb{Z}_{10})$. What could $\operatorname{ord}(\delta(\beta))$ be? Verify by actually computing $\operatorname{ord}(\delta(\beta))$ in U_{10}.

6. Consider the function $f : \mathbb{Z}_{15} \to \mathbb{Z}_{20}$ where $f(a) = 4a$ for all $a \in \mathbb{Z}_{15}$. Show that f is a homomorphism. (This exercise is referenced in Chapter 18, Exercise #8.)

7. Consider again the function $f : \mathbb{Z}_{15} \to \mathbb{Z}_{20}$ where $f(a) = 4a$ for all $a \in \mathbb{Z}_{15}$. We showed in Exercise #6 that f is a homomorphism. For order computations below, note that $\mathbb{Z}_{15}$ and $\mathbb{Z}_{20}$ are *additive* groups.
   (a) Find $\operatorname{ord}(7)$ in $\mathbb{Z}_{15}$ and $\operatorname{ord}(f(7))$ in $\mathbb{Z}_{20}$.
   (b) Find $\operatorname{ord}(6)$ in $\mathbb{Z}_{15}$ and $\operatorname{ord}(f(6))$ in $\mathbb{Z}_{20}$.
   (c) Find $\operatorname{ord}(10)$ in $\mathbb{Z}_{15}$ and $\operatorname{ord}(f(10))$ in $\mathbb{Z}_{20}$.
   (d) Are the above results surprising? Why or why not?

8. Consider the function $\theta : \mathbb{R}^* \to \mathbb{R}^*$ where $\theta(x) = |x|$ for all $x \in \mathbb{R}^*$. Here, $|x|$ denotes the absolute value of x. Show that θ is a homomorphism. (This exercise is referenced in Chapter 18, Exercise #9.)

9. Consider the homomorphism $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.

   (a) How are 2 and 10 related in the domain $\mathbb{Z}_{12}$?

   (b) Compute $\gamma(2)$ and $\gamma(10)$. How are they related in the codomain $\mathbb{Z}_{18}$?

   (c) Repeat parts (a) and (b), starting with the pair of elements 4 and 8 in $\mathbb{Z}_{12}$.

   (d) Are the above results surprising? Why or why not?

10. Consider the homomorphism $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$.

    (a) How are 2 and 7 related in the domain U_{13}?

    (b) Compute $\lambda(2)$ and $\lambda(7)$. How are they related in the codomain U_{13}?

    (c) Repeat parts (a) and (b), starting with the pair of elements 4 and 10 in the domain U_{13}.

    (d) Are the above results surprising? Why or why not?

11. Determine all homomorphisms from U_{20} to $\mathbb{Z}_7$.

12. Rewrite the proof of Theorem 17.12, assuming that the operations of G and H are multiplication and addition, respectively. In other words, prove that $\theta(g^{-1}) = -\theta(g)$ for all $g \in G$.

13. Complete the proof of Theorem 17.13 by proving the case $k < 0$.

14. Rewrite the proof of Theorem 17.13, assuming that the operations of G and H are both addition. In other words, prove that $\theta(k \cdot g) = k \cdot \theta(g)$ for all $g \in G$ and $k \in \mathbb{Z}$.

15. Prove Theorem 17.15.

    Hint: Use the "swap the exponents" technique employed in the proof of Theorem 12.23.

    Note: When referring to an identity element, you must write ε_G or ε_H, rather than just ε.

16. Let $\theta : G \to H$ be a group homomorphism that is *onto*. Consider an element $h \in H$ with $\text{ord}(h) = n$. Prove that there exists an element in G with order n.

17. Let G and H be groups and consider the function $\theta : G \to H$. Moreover, suppose G is commutative. Theorem 16.7 says that if θ is an isomorphism, then H must also be commutative.

    (a) What if θ is a homomorphism, but not necessarily an isomorphism? Must H be commutative then? If so, prove it. If not, provide a counterexample.

    (b) Suppose θ is a homomorphism that's one-to-one. Must H be commutative then? If so, prove it. If not, provide a counterexample.

    (c) Suppose θ is a homomorphism that's onto. Must H be commutative then? If so, prove it. If not, provide a counterexample.

18. Let G be a group and consider the function $\theta : G \to G$ where $\theta(g) = g$ for all $g \in G$.

    Note: This function is often called the *identity function*. Do you see why?

    (a) Prove that θ is a homomorphism.

    (b) Is θ an isomorphism? If so, prove it. If not, explain why not.

19. **Prove:** Let $\alpha : G \to H$ and $\beta : H \to K$ be group homomorphisms. Then $\beta \circ \alpha : G \to K$ is also a homomorphism. (This exercise is referenced in Chapter 18, Exercise #16.)

20. **Prove:** Let $\theta : G \to H$ be a group isomorphism. If G is cyclic with a generator g, then H is cyclic with a generator $\theta(g)$.

    Note: In other words, show that if $G = \langle g \rangle$, then $H = \langle \theta(g) \rangle$.

21. Let H and K be subgroups of a commutative group G. Define $HK = \{hk \mid h \in H, k \in K\}$, which is a subgroup of G. (See Chapter 11, Exercise #24.) Consider the function $\theta : H \times K \to HK$ where $\theta((h, k)) = hk$ for all $(h, k) \in H \times K$. Prove each of these statements:

    (a) θ is an onto homomorphism.

    (b) If $H \cap K = \{\varepsilon\}$, then θ is one-to-one (hence an isomorphism).

22. **Prove:** Let $\alpha : G \to H$ and $\beta : G \to H$ be group homomorphisms. Define $K = \{g \in G \mid \alpha(g) = \beta(g)\}$. Then K is a subgroup of G.

    Note: When referring to an identity element, you must write ε_G or ε_H, rather than just ε.

18

Homomorphisms, Part II

Suppose we want to divide the set $U_{13} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ into 4 equal-sized subsets, with 3 elements each. How would you do it? Here is one possibility:

$$\{1, 2, 3\}, \{4, 5, 6\}, \{7, 8, 9\}, \{10, 11, 12\}.$$

There are many ways to do this, but you might be surprised to learn that there is a *correct* answer, at least according to group theory. (And it's *not* the one just given!)

The above task motivates the focus of this chapter, as we continue our exploration into homomorphisms. In particular, we will encounter the *kernel* and *image*, two important subgroups that are associated to each homomorphism. Suppose $\theta : G \to H$ is a group homomorphism. Then the kernel contains the elements of the domain that are mapped to the identity of the codomain, i.e., elements $g \in G$ such that $\theta(g) = \varepsilon_H$. The image contains the elements of the codomain that are "hit" by the function θ, i.e., elements $h \in H$ such that $h = \theta(g)$ for some $g \in G$. Both the kernel and image will play a key role in the *First Isomorphism Theorem* in Chapter 25, which culminates our study of group theory.

18.1 Kernel of a homomorphism

Example 18.1. Consider the homomorphism $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. Here are the values of $\lambda(a)$ for each input a in the domain U_{13}:

$\lambda(1) = 1,$	$\lambda(4) = 12,$	$\lambda(7) = 5,$	$\lambda(10) = 12,$
$\lambda(2) = 8,$	$\lambda(5) = 8,$	$\lambda(8) = 5,$	$\lambda(11) = 5,$
$\lambda(3) = 1,$	$\lambda(6) = 8,$	$\lambda(9) = 1,$	$\lambda(12) = 12.$

Define the set $K = \{a \in U_{13} \mid \lambda(a) = 1\}$. Thus, K is a subset of the *domain* U_{13} consisting of those elements that map to the identity element 1 in the *codomain* U_{13}. We have $\lambda(1) = \lambda(3) = \lambda(9) = 1$, so that $K = \{1, 3, 9\}$. As shown in the table below,

set K is closed, $1 \in K$, and every element of K has an inverse in K (i.e., 3 and 9 are an inverse pair, and 1 is a self-inverse). Thus, K is a subgroup of U_{13}.

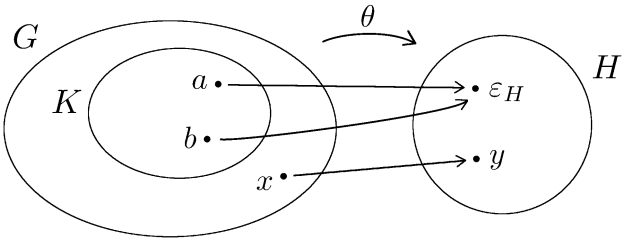
$\cdot$	1	3	9
1	1	3	9
3	3	9	1
9	9	1	3

Example 18.2. Consider the homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. Define $K = \{a \in \mathbb{Z} \mid \varphi(a) = 0\}$. Thus, K is a subset of the *domain* $\mathbb{Z}$ consisting of those elements that map to the identity element 0 in the *codomain* $\mathbb{Z}_5$. Then $K = 5\mathbb{Z}$, the set of multiples of 5, which is a subgroup of $\mathbb{Z}$.

Example 18.3. Consider the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. Define $K = \{\alpha \in G(\mathbb{Z}_{10}) \mid \delta(\alpha) = 1\}$. Thus, K is a subset of the *domain* $G(\mathbb{Z}_{10})$ consisting of those elements that map to the identity element 1 in the *codomain* U_{10}. In other words, the set K contains matrices in $G(\mathbb{Z}_{10})$ whose determinant is 1. We saw in Section 10.3 that $K = S(\mathbb{Z}_{10})$, which is a subgroup of $G(\mathbb{Z}_{10})$.

Example 18.4 (Non-example)**.** Consider again the homomorphism $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. This time, define the set $L = \{a \in U_{13} \mid \lambda(a) = 8\}$. Thus, L is a subset of the *domain* U_{13} containing elements that map to the element 8 in the *codomain* U_{13}. We have $\lambda(2) = \lambda(5) = \lambda(6) = 8$, so that $L = \{2, 5, 6\}$. Unlike K, the set L is *not* a subgroup of U_{13}. (We'll leave it up to you to explain why.)

In each example above, set K is a subset of the domain consisting of those elements that map to the identity element in the codomain. In the figure below, we have $a, b \in K$, because $\theta(a) = \varepsilon_H$ and $\theta(b) = \varepsilon_H$. However, $x \notin K$, because $\theta(x) \neq \varepsilon_H$. This subset K is called the *kernel* of the homomorphism.



Definition 18.5 (Kernel of a homomorphism)**.** Let $\theta : G \to H$ be a group homomorphism. Define

$$K = \{a \in G \mid \theta(a) = \varepsilon_H\},$$

where ε_H is the identity element of H. Then K is called the *kernel* of θ and is denoted $\ker \theta$.

Theorem 18.6 (Kernel is a subgroup)**.** *Let $\theta : G \to H$ be a group homomorphism. Then $K = \ker \theta$ is a subgroup of the domain G.*

PROOF. Let $a, b \in K$ so that $\theta(a) = \varepsilon_H$ and $\theta(b) = \varepsilon_H$. Since θ is operation preserving, we have $\theta(ab) = \theta(a)\theta(b) = \varepsilon_H\varepsilon_H = \varepsilon_H$, so that $\theta(ab) = \varepsilon_H$. Thus, $ab \in K$ and we

conclude that K is closed. We have $\theta(\varepsilon_G) = \varepsilon_H$, and thus $\varepsilon_G \in K$. Finally, $\theta(a^{-1}) = \theta(a)^{-1} = {\varepsilon_H}^{-1} = \varepsilon_H$, so that $\theta(a^{-1}) = \varepsilon_H$. This implies $a^{-1} \in K$. Hence, K is a subgroup of G. ∎

Proof know-how. In the above proof, we used the definition of the kernel in two subtly different ways:

- Assuming that $a \in K$ allowed us to conclude that $\theta(a) = \varepsilon_H$ and to use that fact in the proof. In essence, we're applying the following implication: If $g \in K$, then $\theta(g) = \varepsilon_H$.

- Showing that $\theta(ab) = \varepsilon_H$ allowed us to conclude that $ab \in K$. Here, we're using the converse of the above implication; namely: If $\theta(g) = \varepsilon_H$, then $g \in K$.

The *size* of the kernel provides information about the behavior of the homomorphism. For instance, suppose $\theta : G \to H$ is a group homomorphism with $\ker\theta = G$; i.e., the kernel is all of the domain G. Then *every* element in G maps to the identity ε_H. Thus, θ is the trivial homomorphism. (See Example 17.6.)

Observe that $\ker\theta = G$ is the largest possible subgroup of G. On the other extreme is $\ker\theta = \{\varepsilon_G\}$, which is the smallest possible subgroup of G. What can we say about θ in this case? The following theorem provides the answer.

Theorem 18.7. *Let $\theta : G \to H$ be a group homomorphism. Then $\ker\theta = \{\varepsilon_G\}$ if and only if θ is one-to-one.*

PROOF. We must prove two implications:

- If $\ker\theta = \{\varepsilon_G\}$, then θ is one-to-one.

- If θ is one-to-one, then $\ker\theta = \{\varepsilon_G\}$.

We will prove the second implication. The proof of the first implication is left for you as an exercise.

Assume θ is one-to-one. To show that $\ker\theta = \{\varepsilon_G\}$, we must show that ε_G is the *only* element in $\ker\theta$. First, we have $\theta(\varepsilon_G) = \varepsilon_H$, so that $\varepsilon_G \in \ker\theta$. Next, consider $g \in \ker\theta$. Then $\theta(g) = \varepsilon_H$. Thus, $\theta(g) = \theta(\varepsilon_G)$, as both are equal to ε_H. But since θ is one-to-one, we obtain $g = \varepsilon_G$. Hence, ε_G is the only element in $\ker\theta$, which implies that $\ker\theta = \{\varepsilon_G\}$. ∎

Proof know-how. In the above proof, we showed that $\ker\theta$ is equal to a single-element set, namely $\{\varepsilon_G\}$. More generally, here's how to show $S = \{x\}$, i.e., that a set S contains just one particular element x.

- Verify that $x \in S$, so that the element x is actually in the set S. This shows that $\{x\} \subseteq S$.

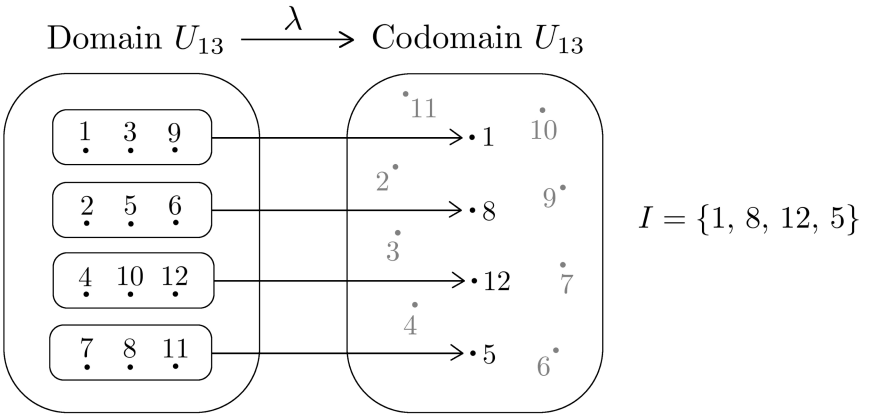
- Show that x is the *only* element in S. This can be done by assuming that there is a second element $y \in S$ and showing that $y = x$. This shows $S \subseteq \{x\}$. (Compare this with the proof of Theorem 8.9.)

18.2 Image of a homomorphism

Example 18.8. Consider again the homomorphism $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. Now define the set $I = \{\lambda(a) \mid a \in U_{13}\}$. Thus, I is a subset of the *codomain* U_{13} consisting of all of the actual outputs of the function. In looking at the calculations shown in Example 18.1, we have the following:

- $\lambda(1) = \lambda(3) = \lambda(9) = 1$.
- $\lambda(2) = \lambda(5) = \lambda(6) = 8$.
- $\lambda(4) = \lambda(10) = \lambda(12) = 12$.
- $\lambda(7) = \lambda(8) = \lambda(11) = 5$.

Therefore, $I = \{1, 8, 12, 5\}$. The figure below depicts the set I.



As shown in the table below, set I is closed, $1 \in I$, and every element of I has an inverse in I (i.e., 5 and 8 are an inverse pair, and 1 and 12 are self-inverses). Thus, I is a subgroup of U_{13}.

$\cdot$	1	8	12	5
1	1	8	12	5
8	8	12	5	1
12	12	5	1	8
5	5	1	8	12

Example 18.9. Consider again the homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. Define the set $I = \{\varphi(a) \mid a \in \mathbb{Z}\}$. Thus, I is a subset of the *codomain* $\mathbb{Z}_5$ consisting of all of the actual outputs of the function. In fact, every element of $\mathbb{Z}_5$ is "hit" by the function φ, since $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(2) = 2$, $\varphi(3) = 3$, and $\varphi(4) = 4$. Therefore, $I = \mathbb{Z}_5$. Also, note that $I = \mathbb{Z}_5$ is a subgroup of $\mathbb{Z}_5$ itself.

The set I in the above examples is called the *image* of the homomorphism.

Definition 18.10 (Image of a homomorphism)**.** Let $\theta : G \to H$ be a group homomorphism. Define

$$I = \{\theta(a) \mid a \in G\}.$$

Then I is called the *image* of θ and is denoted $\operatorname{im}\theta$.

Remark. Let $\theta : G \to H$ be a group homomorphism. As we saw in Example 18.9, θ is onto precisely when the image $I = \operatorname{im}\theta$ is equal to the codomain H.

You will prove the following theorem in an exercise at the end of the chapter.

Theorem 18.11 (Image is a subgroup)**.** *Let $\theta : G \to H$ be a group homomorphism. Then $I = \operatorname{im}\theta$ is a subgroup of the codomain H.*

18.3 Partitioning the domain

Let's return to the question posed at the beginning of this chapter.

> Suppose we want to divide the set $U_{13} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ into 4 equal-sized subsets, with 3 elements each. How would you do it?

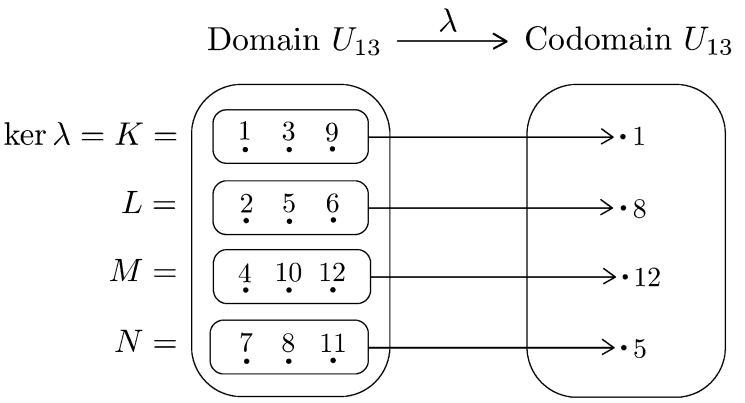
Answering this question is the goal of this section.

Example 18.12. Consider yet again the homomorphism $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. We will use λ to divide the domain U_{13} into 4 equal-sized subsets. From Example 18.8, the image of λ is given by $\operatorname{im}\lambda = \{\lambda(a) \mid a \in U_{13}\} = \{1, 8, 12, 5\}$.

For each element of $\operatorname{im}\lambda$, we define a corresponding subset of the domain U_{13}:

$$\begin{aligned}
\ker\lambda = K &= \{a \in U_{13} \mid \lambda(a) = \ 1\} = \{1, \ 3, \ 9\},\\
L &= \{a \in U_{13} \mid \lambda(a) = \ 8\} = \{2, \ 5, \ 6\},\\
M &= \{a \in U_{13} \mid \lambda(a) = 12\} = \{4, 10, 12\},\\
N &= \{a \in U_{13} \mid \lambda(a) = \ 5\} = \{7, \ 8, 11\}.
\end{aligned}$$

For instance, the subset L contains elements of the domain U_{13} that map to 8 in the codomain U_{13}. We have $\lambda(2) = \lambda(5) = \lambda(6) = 8$, so that $L = \{2, 5, 6\}$. Notice that the domain U_{13} has been divided into 4 equal-sized subsets, namely K, L, M, and N. The diagram below illustrates this scenario.



In Example 18.12 above, the homomorphism λ *partitions* the domain into equal-sized subsets. Here, "partition" means that the subsets are disjoint (i.e., they don't overlap) and that the subsets together cover the entire domain. This alone is not a big deal, since any function (whether or not it's a homomorphism) does the same. What *is* a big deal is that these subsets are *equal-sized*. In the next unit, we will see the role that a homomorphism plays in ensuring that these subsets are equal-sized.

Example 18.13 (Non-example)**.** Consider the function $f : \mathbb{Z}_8 \to \mathbb{Z}_8$ where $f(a) = a^2$ for all $a \in \mathbb{Z}_8$. We have $f(1+3) = f(4) = 4^2 = 0$ and $f(1)+f(3) = 1^2+3^2 = 1+9 = 2$, so that $f(1+3) \neq f(1)+f(3)$. Thus, f is *not* a homomorphism. Moreover, we have the following:

- $f(0) = f(4) = 0$.
- $f(1) = f(3) = f(5) = f(7) = 1$.
- $f(2) = f(6) = 4$.

Hence, the function f divides the domain $\mathbb{Z}_8$ into 3 subsets:

$$\{a \in \mathbb{Z}_8 \mid f(a) = 0\} = \{0, 4\},$$
$$\{a \in \mathbb{Z}_8 \mid f(a) = 1\} = \{1, 3, 5, 7\},$$
$$\{a \in \mathbb{Z}_8 \mid f(a) = 4\} = \{2, 6\}.$$

These subsets do form a partition of the domain $\mathbb{Z}_8$, since they do not overlap and together cover all of $\mathbb{Z}_8$. However, they are *not* equal-sized.

18.4 Finding homomorphisms

How many different functions are there from $\mathbb{Z}_{12}$ to itself? Consider a function $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$. Then the output $\theta(0)$ could be any element in the codomain $\mathbb{Z}_{12}$, so there are 12 possible values for $\theta(0)$. Likewise, there are 12 possible values for $\theta(1)$, for $\theta(2)$, for $\theta(3)$, …, and for $\theta(11)$. Altogether, there are 12^{12} (which is almost 9 trillion!) different functions from $\mathbb{Z}_{12}$ to itself.

What if we require that θ be a homomorphism? First, we must have $\theta(0) = 0$, since a homomorphism maps the identity element of the domain to the identity element of the codomain (Theorem 17.9). As we'll see in the next example, there are far fewer than 9 trillion homomorphisms from $\mathbb{Z}_{12}$ to itself.

Example 18.14. Let $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ be a homomorphism. Furthermore, suppose $\theta(1) = 4$. Since θ is operation preserving, we have the following:

- $\theta(2) = \theta(1+1) = \theta(1)+\theta(1) = 4+4 = 8$.
- $\theta(3) = \theta(1+1+1) = \theta(1)+\theta(1)+\theta(1) = 4+4+4 = 0$.
- $\theta(4) = \theta(1+1+1+1) = \theta(1)+\theta(1)+\theta(1)+\theta(1) = 4+4+4+4 = 4$.

$\vdots$

In general, for any $n \in \mathbb{Z}_{12}$, we have

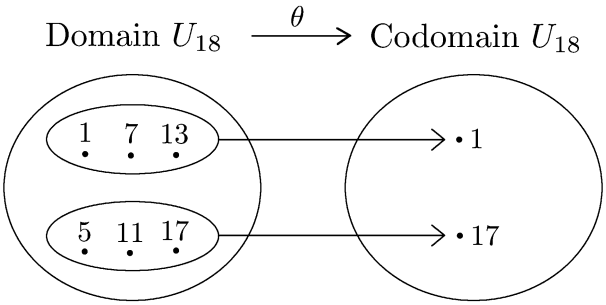
$$\begin{aligned}\theta(n) &= \theta(\underbrace{1+1+1+\cdots+1}_{n \text{ terms}}) \\ &= \underbrace{\theta(1)+\theta(1)+\theta(1)+\cdots+\theta(1)}_{n \text{ terms}} \\ &= \underbrace{4+4+4+\cdots+4}_{n \text{ terms}} \\ &= 4 \cdot n,\end{aligned}$$

so that $\theta(n) = 4n$ for all $n \in \mathbb{Z}_{12}$.

But there is nothing special about the multiplier 4 in the rule $\theta(n) = 4n$. We could have replaced 4 with any fixed element of $\mathbb{Z}_{12}$, say $\theta(n) = 7n$ for all $n \in \mathbb{Z}_{12}$. Thus, a homomorphism $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ must have the rule $\theta(n) = k \cdot n$ for all $n \in \mathbb{Z}_{12}$, where k is a fixed element of $\mathbb{Z}_{12}$. Since there are 12 choices for the multiplier k, we conclude that there are 12 different homomorphisms from $\mathbb{Z}_{12}$ to itself.

Example 18.15. Now let $\theta : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ be an *isomorphism*. Since every isomorphism is a homomorphism, we know from Example 18.14 that θ must have the rule $\theta(n) = k \cdot n$ for all $n \in \mathbb{Z}_{12}$, where k is a fixed element of $\mathbb{Z}_{12}$. Thus, we have $\theta(1) = k \cdot 1 = k$. By Theorem 17.15, we also know that $\operatorname{ord}(\theta(1)) = \operatorname{ord}(1)$, or equivalently, $\operatorname{ord}(k) = \operatorname{ord}(1)$. Because $\operatorname{ord}(1) = 12$, we find that $\operatorname{ord}(k) = 12$ as well. Thus, the possible values of k are $k = 1$, 5, 7, and 11 (i.e., the elements of $\mathbb{Z}_{12}$ that have order 12). Hence, there are only 4 different isomorphisms from $\mathbb{Z}_{12}$ to itself.

Example 18.16. Consider the multiplicative group $U_{18} = \{1, 5, 7, 11, 13, 17\}$. Let $\theta : U_{18} \to U_{18}$ be a homomorphism with $\ker\theta = \{1, 7, 13\}$ and $\theta(5) = 17$. Since the elements in the kernel map to the identity element of the codomain, we have $\theta(1) = \theta(7) = \theta(13) = 1$. According to our conjecture from Section 18.3, θ partitions the domain U_{18} into equal-sized subsets. We know 3 elements of the domain, namely 1, 7, 13, map to 1 in the codomain. Thus, since $\theta(5) = 17$, we conjecture that 3 elements of the domain map to 17 in the codomain; i.e., $\theta(5) = \theta(11) = \theta(17) = 17$. This analysis is depicted in the figure below.



Exercises

1. Consider the homomorphism $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.
   (a) Let $K = \{a \in \mathbb{Z}_{12} \mid \gamma(a) = 0\}$. Find the elements of K.
   (b) Is K a subset of the domain or the codomain?
   (c) Create an addition table for K and verify that it's a subgroup of $\mathbb{Z}_{12}$.

2. Consider the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det\alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$.
   (a) Let $I = \{\delta(\alpha) \mid \alpha \in G(\mathbb{Z}_{10})\}$. Find the elements of I.
   (b) Is I a subset of the domain or the codomain?
   (c) Anita says, "The set I is all of the codomain U_{10}, because the function δ is onto." What might she mean?

3. Consider the homomorphism $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.
   (a) Let $I = \{\gamma(a) \mid a \in \mathbb{Z}_{12}\}$. Find the elements of I.
   (b) Is I a subset of the domain or the codomain?
   (c) Create an addition table for I and verify that it's a subgroup of $\mathbb{Z}_{18}$.

4. Consider again the homomorphism $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$. Similar to what we did with λ in Example 18.12, use γ to partition the domain $\mathbb{Z}_{12}$ into equal-sized subsets.

5. Consider again the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$.
   (a) Proceed as in Example 18.12 and use δ to partition the domain $G(\mathbb{Z}_{10})$ into subsets.
   (b) How many subsets are there, and how do you describe the matrices in each subset?
   (c) **(Challenge)** Verify that the subsets created by δ are equal-sized.

6. Consider again the homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$.
   (a) Proceed as in Example 18.12 and use φ to partition the domain $\mathbb{Z}$ into subsets.
   (b) How many subsets are there, and how do you describe the integers in each subset?
   (c) The subsets created by φ have infinitely many elements. But in what sense are they "equal-sized"?

7. Consider the function $f : D_4 \to \mathbb{R}^*$ where
$$f(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is a rotation,} \\ -1 & \text{if } \sigma \text{ is a reflection.} \end{cases}$$
   We showed in Chapter 17, Exercise #3 that f is a homomorphism.
   (a) Find $K = \ker f$ and verify that it's a subgroup of the domain D_4.
   (b) Find $I = \operatorname{im} f$ and verify that it's a subgroup of the codomain $\mathbb{R}^*$.
   (c) Proceed as in Example 18.12 and use f to partition the domain D_4 into subsets.
   (d) Verify that the subsets created by f are equal-sized.

8. Consider the function $f : \mathbb{Z}_{15} \to \mathbb{Z}_{20}$ where $f(a) = 4a$ for all $a \in \mathbb{Z}_{15}$. We showed in Chapter 17, Exercise #6 that f is a homomorphism. Repeat Exercise #7 above using this homomorphism.

9. Consider the function $\theta : \mathbb{R}^* \to \mathbb{R}^*$ where $\theta(x) = |x|$ for all $x \in \mathbb{R}^*$. Here, $|x|$ denotes the absolute value of x. We showed in Chapter 17, Exercise #8 that θ is a homomorphism. Repeat Exercise #7 above using this homomorphism.

10. Let $f : S \to T$ be a function from set S to set T. (**Note:** f need *not* be a group homomorphism.) Supposing that we proceed as in Example 18.12, explain why the function f partitions the domain S. In other words, explain why the subsets created by f are disjoint (i.e., they don't overlap) and that the subsets together cover the entire domain S. (This exercise is referenced in Section 25.1.)

11. Let $\theta : U_{16} \to U_{16}$ be a homomorphism with $\ker\theta = \{1, 7, 9, 15\}$ and $\theta(11) = 9$. Find the output values $\theta(a)$ for all $a \in U_{16}$.

12. Let $\theta : U_{13} \to U_{13}$ be a homomorphism with $\ker\theta = \{1, 5, 8, 12\}$ and $\theta(4) = \theta(6) = \theta(7) = \theta(9) = 3$. Find the ouput values $\theta(a)$ for all $a \in U_{13}$.

13. Complete the proof of Theorem 18.7 by proving its first implication.

14. Prove Theorem 18.11.

    Note: When referring to an identity element, you must write ε_G or ε_H, rather than just ε.

15. Let $\theta : G \to H$ be a group homomorphism. By Theorem 18.11, the set $I = \text{im}\,\theta$ is a subgroup of the codomain H. Below, you'll show that the image I retains some of the properties of the domain G.

    (a) **Prove:** If G is commutative, then I is commutative.

    (b) **Prove:** If G is cyclic, then I is cyclic.

16. Let $\alpha : G \to H$ and $\beta : H \to K$ be group homomorphisms. In Chapter 17, Exercise #19, you showed that $\beta \circ \alpha : G \to K$ is also a homomorphism.

    (a) $\ker\alpha$ is a subset (in fact, a subgroup) of G, H, or K?

    (b) $\ker\beta \circ \alpha$ is a subset of G, H, or K?

    (c) How are $\ker\alpha$ and $\ker\beta \circ \alpha$ related? Explain.

17. Let $\theta : G \to H$ be a group homomorphism, and let Q be a subgroup of H. Define the set

$$P = \{a \in G \mid \theta(a) \in Q\}.$$

    (a) Show that P is a subgroup of G.

    (b) Anita says, "P is a more general version of the kernel of θ." What might she mean?

    (This exercise is referenced in Section 24.4.)

18. (a) Find all homomorphisms from $\mathbb{Z}_{18}$ to itself.

    (b) Find all isomorphisms from $\mathbb{Z}_{18}$ to itself.

19. (a) Find all homomorphisms from U_{19} to itself.

    (b) Find all isomorphisms from U_{19} to itself.

    Hint: It turns out that $U_{19} = \langle 2 \rangle$. (See Chapter 13, Exercise #7.) How does that help?

20. (a) Find all homomorphisms from $\mathbb{Z}_m$ to itself.

    (b) Find all isomorphisms from $\mathbb{Z}_m$ to itself.

21. (a) Find all homomorphisms from $\mathbb{Z}$ to itself.

    (b) Find all isomorphisms from $\mathbb{Z}$ to itself.

22. Let $\theta : \mathbb{Z} \to \mathbb{Q}$ be a homomorphism, where $\mathbb{Q}$ is the additive group of rational numbers.
    (a) Suppose $\theta(1) = \frac{1}{7}$. Find the image $\operatorname{im}\theta = \{\theta(a) \mid a \in \mathbb{Z}\}$.
    (b) Suppose $\theta(1) = \frac{3}{8}$. Find $\operatorname{im}\theta$.
    (c) Suppose $\theta(1) = -\frac{6}{19}$. Find $\operatorname{im}\theta$.
    (d) Explain why θ cannot be an isomorphism.

Unit V: Quotient Groups

Chapter 19 introduces *cosets*, which are subsets of a group with powerful consequences. One such result is *Lagrange's theorem*, arguably the most important theorem about groups, which states that (spoiler alert!) if H is a subgroup of a finite group G, then the size of H is a divisor of the size of G.

I often tell students, "If proving Lagrange's theorem is all that cosets are good for, then cosets would still hold a special place in group theory." But there's so much more to cosets! Chapter 21 describes how we can turn a set of cosets into its own group called the *quotient group*. After an extensive investigation into quotient groups, we end our study of group theory in Chapter 25, where we resolve our conjecture (from Chapter 18) about how a homomorphism partitions the domain into equal-sized subsets.

Here is a taste of what you'll be able to accomplish in this unit:

- Prove that distinct cosets of a subgroup H form a *partition* of group G. This becomes a key ingredient to the proof of Lagrange's theorem.
- Discover a nifty shortcut for multiplying a pair of cosets. Then learn about *normal* subgroups by analyzing when this coset multiplication shortcut fails and when it works.
- Write lots of proofs involving quotient groups, carefully navigating between group G (whose elements have the form $a \in G$) and quotient group G/H (whose elements are *cosets* of the form $aH \in G/H$).

19

Introduction to Cosets

This chapter introduces a new object called a *coset*, which is obtained by taking a subgroup of a group and multiplying each element of the subgroup by a fixed element. In fact, cosets made a surreptitious appearance in Chapter 18, although they were disguised enough that you likely did not notice them. Despite their relatively simply construction, cosets play a powerful role in group theory. We will use them to prove *Lagrange's theorem* in the next chapter. Cosets will also be used to create a new type of a group called a *quotient group*, which will be the primary focus of the rest of this unit.

19.1 Multiplicative group example

Example 19.1. Consider the multiplicative group $U_{13} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ and its subgroup $H = \{1, 3, 9\}$. Choose an element $6 \in U_{13}$. Then the *coset* $6H$ is obtained by multiplying each element of H by 6; i.e., $6H = \{6\cdot 1, 6\cdot 3, 6\cdot 9\} = \{6, 5, 2\}$.

Just as we found $6H$, let's compute the coset aH for each $a \in U_{13}$. Since there are 12 elements in U_{13}, we would expect to obtain 12 cosets:

$$1H = \{1\cdot 1, 1\cdot 3, 1\cdot 9\} = \{1, 3, 9\}, \qquad 7H = \{7\cdot 1, 7\cdot 3, 7\cdot 9\} = \{7, 8, 11\},$$
$$2H = \{2\cdot 1, 2\cdot 3, 2\cdot 9\} = \{2, 6, 5\}, \qquad 8H = \{8\cdot 1, 8\cdot 3, 8\cdot 9\} = \{8, 11, 7\},$$
$$3H = \{3\cdot 1, 3\cdot 3, 3\cdot 9\} = \{3, 9, 1\}, \qquad 9H = \{9\cdot 1, 9\cdot 3, 9\cdot 9\} = \{9, 1, 3\},$$
$$4H = \{4\cdot 1, 4\cdot 3, 4\cdot 9\} = \{4, 12, 10\}, \qquad 10H = \{10\cdot 1, 10\cdot 3, 10\cdot 9\} = \{10, 4, 12\},$$
$$5H = \{5\cdot 1, 5\cdot 3, 5\cdot 9\} = \{5, 2, 6\}, \qquad 11H = \{11\cdot 1, 11\cdot 3, 11\cdot 9\} = \{11, 7, 8\},$$
$$6H = \{6\cdot 1, 6\cdot 3, 6\cdot 9\} = \{6, 5, 2\}, \qquad 12H = \{12\cdot 1, 12\cdot 3, 12\cdot 9\} = \{12, 10, 4\}.$$

There are several duplicates in this list of cosets. Recall that in a set, the order in which the elements are listed does *not* matter. For instance, $4H = \{4, 12, 10\}$, $10H = \{10, 4, 12\}$, and $12H = \{12, 10, 4\}$ are the same set, because they all contain the same elements. More generally, these duplicates occur because the cosets aH and bH can be equal (i.e., they contain the same elements) even when $a \neq b$. For instance, we have $4H = 10H = 12H$, even though 4, 10, and 12 are distinct elements in U_{13}.

Consolidating the duplicates, there are only four *distinct* cosets within the above list:

- $1H = \ 3H = \ 9H = \{1, \ 3, \ 9\}$ (original subgroup).
- $2H = \ 5H = \ 6H = \{2, \ 5, \ 6\}$.
- $4H = 10H = 12H = \{4, 10, 12\}$.
- $7H = \ 8H = 11H = \{7, \ 8, 11\}$.

Aside from the original subgroup (i.e., $1H = 3H = 9H$), none of the other cosets are subgroups of U_{13}. In an exercise, you'll describe which properties of a subgroup are violated by, say, $4H = 10H = 12H$.

Below are observations about these cosets, which will be generalized in Section 19.4.

(1) The coset aH contains the element a. For example, $10H = \{4, 10, 12\}$ contains the element 10.

(2) We have $1H = 3H = 9H = H$, the original subgroup; and 1, 3, 9 are precisely the elements of H.

(3) All cosets have the same size, namely the size of H.

(4) The distinct cosets form a *partition* of U_{13}. Recall from Section 18.3 that "partition" means that the distinct cosets do not overlap and together cover all of U_{13}.

Example 19.2. Consider the group $U_{20} = \{1, 3, 7, 9, 11, 13, 17, 19\}$ and its subgroup $H = \{1, 9\}$. How many *distinct* cosets do we expect to find? Let's compute the coset aH for each $a \in U_{20}$:

$$\begin{aligned} 1H &= \{1 \cdot 1, 1 \cdot 9\} = \{1, 9\}, & 11H &= \{11 \cdot 1, 11 \cdot 9\} = \{11, 19\}, \\ 3H &= \{3 \cdot 1, 3 \cdot 9\} = \{3, 7\}, & 13H &= \{13 \cdot 1, 13 \cdot 9\} = \{13, 17\}, \\ 7H &= \{7 \cdot 1, 7 \cdot 9\} = \{7, 3\}, & 17H &= \{17 \cdot 1, 17 \cdot 9\} = \{17, 13\}, \\ 9H &= \{9 \cdot 1, 9 \cdot 9\} = \{9, 1\}, & 19H &= \{19 \cdot 1, 19 \cdot 9\} = \{19, 11\}. \end{aligned}$$

Again, we see duplicates in this list of cosets. For instance, we have $3H = 7H$, even though 3 and 7 are distinct elements in U_{20}. Consolidating the duplicates, we obtain four distinct cosets:

- $1H = \ 9H = \{1, 9\}$ (original subgroup).
- $3H = \ 7H = \{3, 7\}$.
- $11H = 19H = \{11, 19\}$.
- $13H = 17H = \{13, 17\}$.

The observations that we made about the cosets in Example 19.1 apply here as well. After computing $3H = \{3, 7\}$, it's reasonable to suspect that $7H$ must be the same coset, because $7H$ should contain the element 7. We also have $1H = 9H = \{1, 9\}$, the original subgroup; and 1, 9 are precisely the elements of H. All the cosets have the same size, namely 2 elements each. And these four cosets do form a partition of U_{20}.

Example 19.3. Consider the group $D_4 = \{\varepsilon, r_{90}, r_{180}, r_{270}, h, v, d, d'\}$ and its subgroup $H = \{\varepsilon, d\}$. Since D_4 and H contain 8 and 2 elements, respectively, we would expect four distinct cosets. (Do you see why?) Referring to Appendix B for the group table of D_4, we compute the coset aH for each $a \in D_4$:

$$\begin{aligned}
\varepsilon H &= \{\varepsilon \cdot \varepsilon, \varepsilon \cdot d\} = \{\varepsilon, d\}, & hH &= \{h \cdot \varepsilon, h \cdot d\} = \{h, r_{90}\},\\
r_{90}H &= \{r_{90} \cdot \varepsilon, r_{90} \cdot d\} = \{r_{90}, h\}, & vH &= \{v \cdot \varepsilon, v \cdot d\} = \{v, r_{270}\},\\
r_{180}H &= \{r_{180} \cdot \varepsilon, r_{180} \cdot d\} = \{r_{180}, d'\}, & dH &= \{d \cdot \varepsilon, d \cdot d\} = \{d, \varepsilon\},\\
r_{270}H &= \{r_{270} \cdot \varepsilon, r_{270} \cdot d\} = \{r_{270}, v\}, & d'H &= \{d' \cdot \varepsilon, d' \cdot d\} = \{d', r_{180}\}.
\end{aligned}$$

Consolidating the duplicates, we obtain four distinct cosets, as we had expected:

- $\varepsilon H = dH = \{\varepsilon, d\}$ (original subgroup).
- $r_{90}H = hH = \{r_{90}, h\}$.
- $r_{180}H = d'H = \{r_{180}, d'\}$.
- $r_{270}H = vH = \{r_{270}, v\}$.

We will leave it up to you to verify that the observations that we made about the cosets in Example 19.1 are satisfied in this example as well.

Example 19.4. Consider the matrix group $G(\mathbb{Z}_{10}) = \{\alpha \in M(\mathbb{Z}_{10}) \mid \alpha \text{ has a multiplicative inverse}\}$. Let $H = S(\mathbb{Z}_{10}) = \{\alpha \in G(\mathbb{Z}_{10}) \mid \det \alpha = 1\}$ and recall from Section 10.3 that $S(\mathbb{Z}_{10})$ is a subgroup of $G(\mathbb{Z}_{10})$.

Fix an element $\mu = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$ with $\det \mu = 3$. To obtain the coset μH, we multiply each element of H by μ; i.e., $\mu H = \{\mu \cdot h \mid h \in H\}$. For instance, let $h = \left[\begin{smallmatrix} 7 & 2 \\ 5 & 3 \end{smallmatrix}\right]$ with $\det h = 7 \cdot 3 - 2 \cdot 5 = 1$, so that $h \in H$. Thus the following matrix is in the coset μH: $\mu \cdot h = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} 7 & 2 \\ 5 & 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 9 & 7 \\ 5 & 2 \end{smallmatrix}\right]$. We note that $\det(\mu \cdot h) = \det \left[\begin{smallmatrix} 9 & 7 \\ 5 & 2 \end{smallmatrix}\right] = 9 \cdot 2 - 7 \cdot 5 = 3$, and so $\mu \cdot h$ has determinant 3, just like μ. In fact, you'll show in an exercise that every matrix in the coset μH has determinant 3.

Conversely, you'll also show that every matrix with determinant 3 is in the coset μH. For instance, let $\beta = \left[\begin{smallmatrix} 5 & 7 \\ 1 & 2 \end{smallmatrix}\right]$ and note that $\det \beta = 5 \cdot 2 - 7 \cdot 1 = 3$. Let $h = \left[\begin{smallmatrix} 3 & 2 \\ 9 & 3 \end{smallmatrix}\right]$ with $\det h = 3 \cdot 3 - 2 \cdot 9 = 1$, so that $h \in H$. We have $\mu \cdot h = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} 3 & 2 \\ 9 & 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 5 & 7 \\ 1 & 2 \end{smallmatrix}\right]$, so that $\mu \cdot h = \beta$. Therefore, $\beta \in \mu H$ as desired. (How did we come up with the matrix h here? That's for you to explore in the exercises!)

19.2 Additive group example

As a default, we assume that an operation of a group is multiplication. But cosets of additive groups will be particularly important when we study *rings* later in the book. So we will take an in-depth look at the additive case in the following examples.

Example 19.5. Consider the additive group $\mathbb{Z}_{12} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ and its subgroup $H = \{0, 4, 8\}$. Choose $6 \in \mathbb{Z}_{12}$. Then the *coset* $6 + H$ is obtained by adding 6 to each element of H; i.e., $6 + H = \{6 + 0, 6 + 4, 6 + 8\} = \{6, 10, 2\}$. In an exercise, you'll compute the coset $a + H$ for each $a \in \mathbb{Z}_{12}$. You should find that there

are only four *distinct* cosets:

- $0 + H = 4 + H = 8 + H = \{0, 4, 8\}$ (original subgroup).
- $1 + H = 5 + H = 9 + H = \{1, 5, 9\}$.
- $2 + H = 6 + H = 10 + H = \{2, 6, 10\}$.
- $3 + H = 7 + H = 11 + H = \{3, 7, 11\}$.

Aside from the original subgroup $0 + H = 4 + H = 8 + H$, none of the other cosets are subgroups of $\mathbb{Z}_{12}$. Notice again that the duplicates occur, because the cosets $a + H$ and $b + H$ can be equal (i.e., they contain the same elements) even when $a \neq b$. For instance, $5 + H = 9 + H$, even though $5 \neq 9$ in $\mathbb{Z}_{12}$.

Here are some observations about these additive cosets. Note how they're the same as the observations about the cosets in Example 19.1, but written in the language of addition.

(1) The coset $a + H$ contains the element a. For example, $10 + H = \{2, 6, 10\}$ contains the element 10.

(2) We have $0 + H = 4 + H = 8 + H = H$, the original subgroup; and $H = \{0, 4, 8\}$.

(3) All cosets have the same size, namely the size of H.

(4) The distinct cosets form a partition of $\mathbb{Z}_{12}$.

Here is an example of cosets where the group and subgroup have infinitely many elements.

Example 19.6. Consider the additive group $\mathbb{Z}$ and its subgroup

$$H = 5\mathbb{Z} = \{\ldots, -20, -15, -10, -5, 0, 5, 10, 15, 20, \ldots\}.$$

As an example, here's the coset $7 + H$, which is obtained by adding 7 to each element of H:

$$\begin{aligned} 7 + H &= \{7 + h \mid h \in H\} \\ &= \{\ldots, 7 + (-20), 7 + (-15), 7 + (-10), 7 + (-5), \\ &\qquad\qquad 7 + 0, 7 + 5, 7 + 10, 7 + 15, 7 + 20, \ldots\} \\ &= \{\ldots, -13, -8, -3, 2, 7, 12, 17, 22, 27, \ldots\}. \end{aligned}$$

There are five distinct cosets of H in $\mathbb{Z}$, as shown:

- $\cdots = -5 + H = 0 + H = 5 + H = 10 + H = 15 + H = \cdots$ (original subgroup).
- $\cdots = -4 + H = 1 + H = 6 + H = 11 + H = 16 + H = \cdots$.
- $\cdots = -3 + H = 2 + H = 7 + H = 12 + H = 17 + H = \cdots$.
- $\cdots = -2 + H = 3 + H = 8 + H = 13 + H = 18 + H = \cdots$.
- $\cdots = -1 + H = 4 + H = 9 + H = 14 + H = 19 + H = \cdots$.

The observations made in earlier examples apply here as well. For instance, we have $7+H = 12+H$, even though 7 and 12 are distinct elements in $\mathbb{Z}$. After computing

$$7 + H = \{\ldots, -13, -8, -3, 2, 7, \mathbf{12}, 17, 22, 27, \ldots\},$$

it's reasonable to suspect that $12 + H$ must be the same coset, because $12 + H$ should contain the element 12. We also have $\cdots = -10 + H = -5 + H = 0 + H = 5 + H = 10 + H = 15 + H = \cdots = H$, the original subgroup; and $0, \pm5, \pm10, \pm15, \ldots$ are the elements of H. And these five distinct cosets do form a partition of $\mathbb{Z}$.

19.3 Right cosets

In Section 19.1, we considered *left* cosets of the form aH, where we multiplied each element of H *on the left* by a. We can also consider *right* cosets Ha, as shown in the example below.

Example 19.7. Consider again the group U_{13} and its subgroup $H = \{1, 3, 9\}$. The left coset $6H$ is given by $6H = \{6 \cdot 1, 6 \cdot 3, 6 \cdot 9\} = \{6, 5, 2\}$, and we have the right coset $H6 = \{1 \cdot 6, 3 \cdot 6, 9 \cdot 6\} = \{6, 5, 2\}$. Observe that $6H = H6$; i.e., the left and right cosets are equal, because U_{13} is commutative.

As seen in Example 19.7, the distinction between left and right cosets is irrelevant in a commutative group. In particular, additive groups are always commutative, so there is no distinction between the left coset $a + H$ and the right coset $H + a$. Thus, we will only consider left cosets $a + H$ with additive groups.

Here is a non-commutative example, where things get a bit more interesting.

Example 19.8. Let $H = \{\varepsilon, d\}$ be a subgroup of D_4. Let's compute and compare the left coset $r_{90}H$ and the right coset Hr_{90}:

- $r_{90}H = \{r_{90} \cdot \varepsilon, r_{90} \cdot d\} = \{r_{90}, h\}$.
- $Hr_{90} = \{\varepsilon \cdot r_{90}, d \cdot r_{90}\} = \{r_{90}, v\}$.

Therefore, the left and right cosets are not the same; i.e., $r_{90}H \neq Hr_{90}$.

Example 19.9. Let $K = C(h) = \{\varepsilon, r_{180}, h, v\}$ be a subgroup of D_4. (It's the *centralizer* of h in D_4. See Section 5.3.) Let's compute and compare the left coset dK and the right coset Kd:

- $dK = \{d \cdot \varepsilon, d \cdot r_{180}, d \cdot h, d \cdot v\} = \{d, d', r_{270}, r_{90}\}$.
- $Kd = \{\varepsilon \cdot d, r_{180} \cdot d, h \cdot d, v \cdot d\} = \{d, d', r_{90}, r_{270}\}$.

Thus we have a *coset equality* $dK = Kd$, because these sets contain the same four elements. But this does not imply that we have an *element-by-element equality*; i.e., $dk = kd$ for all $k \in K$. Indeed, we have $dh \neq hd$ and $dv \neq vd$, where $h, v \in K$.

You'll show in an exercise at the end of the chapter that $\varepsilon K = K\varepsilon$, $r_{90}K = Kr_{90}$, $r_{180}K = Kr_{180}$, and so on. In fact, it turns out that $aK = Ka$ for all $a \in D_4$, so that left and right cosets are always equal in this example. But be careful: Coset equality does *not* imply element-by-element equality.

Remark. In the next section, we will prove various properties of cosets. The proofs will be written using left cosets. But for each property about left cosets, an analogous theorem holds true for right cosets.

19.4 Properties of cosets

We've seen plenty of examples of cosets thus far, and now we're ready for a general definition.

Definition 19.10 (Coset)**.** Let G be a group, H a subgroup of G, and $a \in G$. Then:

- The set $aH = \{ah \mid h \in H\}$ is the *left coset* of H generated by a.
- The set $Ha = \{ha \mid h \in H\}$ is the *right coset* of H generated by a.

The element a is called the *coset representative* of aH and Ha.

Remark. If G is an *additive* group, then the left and right cosets are $a + H = \{a + h \mid h \in H\}$ and $H + a = \{h + a \mid h \in H\}$, respectively. Recall that we always have $a + H = H + a$, since additive groups are commutative. Given this lack of distinction between left and right cosets, we will only consider left cosets $a + H$ with additive groups.

Below are the first three properties of cosets observed in Examples 19.1 and 19.5. (The fourth property about how the distinct cosets partition the group will be addressed in the next chapter.) While they are stated in the context of left cosets, as will be typical of coset theorems, analogous statements are true for right cosets. Each proof is written for a multiplicative group, and the proofs for an additive group are left for you as an exercise.

The following example motivates the proof of the first theorem.

Example 19.11. Consider again the group U_{13} and its subgroup $H = \{\mathbf{1}, 3, 9\}$. Since H is a subgroup, it must contain the identity element $\mathbf{1}$. Thus the coset $6H$ must contain the element $6 \cdot \mathbf{1}$ or 6; i.e.,

$$6H = \{6 \cdot \mathbf{1}, 6 \cdot 3, 6 \cdot 9\} = \{\mathbf{6}, 5, 2\}.$$

Theorem 19.12. *A coset representative is contained in the coset that it generates. Specifically, let G be a group, H a subgroup of G, and $a \in G$. Then:*

- (*Multiplicative*) *The coset aH contains the element a; i.e., $a \in aH$.*
- (*Additive*) *The coset $a + H$ contains the element a; i.e., $a \in a + H$.*

PROOF. Since H is a subgroup, it contains the identity element ε. Then $a = a\varepsilon \in aH$. ∎

The next theorem says that the elements of G whose cosets are the same as the original subgroup are precisely those elements that are in H. Here are some examples we've seen that illustrate the theorem.

Example 19.13.

- For the group U_{13} and subgroup $H = \{1, 3, 9\}$, we have $1H = 3H = 9H = H$, the original subgroup; and 1, 3, 9 are precisely the elements of H.
- For the group $\mathbb{Z}_{12}$ and subgroup $H = \{0, 4, 8\}$, we have $0+H = 4+H = 8+H = H$, the original subgroup; and 0, 4, 8 are precisely the elements of H.

- For the group D_4 and subgroup $H = \{\varepsilon, d\}$, we have $\varepsilon H = dH = H$, the original subgroup; and ε, d are precisely the elements of H.

Theorem 19.14. *Let G be a group, H a subgroup of G, and $a \in G$. Then:*

- (*Multiplicative*) *$aH = H$ if and only if $a \in H$.*
- (*Additive*) *$a + H = H$ if and only if $a \in H$.*

PROOF. We must prove two implications:

- If $aH = H$, then $a \in H$.
- If $a \in H$, then $aH = H$.

We will prove the second implication. The proof of the first implication is left for you as an exercise.

Assume $a \in H$. To prove $aH = H$, we must show that $aH \subseteq H$ and $H \subseteq aH$. We begin with $aH \subseteq H$. Let $g \in aH$ so that $g = ah$ for some $h \in H$. Since a and h are both in H, we know $ah \in H$ by the closure of H. Then $g \in H$ and thus $aH \subseteq H$. Next, we will show $H \subseteq aH$. Let $g \in H$. To show that $g \in aH$, we must show that $g = ah$ for some $h \in H$. Let $h = a^{-1}g$, which is in H, because a and g are in H. And we have $ah = a(a^{-1}g) = (aa^{-1})g = g$, so that $g = ah \in aH$. This shows that $H \subseteq aH$, so that $aH = H$. ■

Proof know-how. Proofs about cosets often involve a group element contained in a coset. Note how the remarks below are similar to those given after the proof of Theorem 18.6.

- In the first part of the proof, *assuming* that $g \in aH$ allowed us to conclude that $g = ah$ for some $h \in H$. In essence, we're applying the following implication: If $g \in aH$, then $g = ah$ for some $h \in H$.
- Later in the proof, *showing* that $g = ah$ for some $h \in H$ allowed us to conclude that $g \in aH$. Here, we're using the converse of the above implication; namely: If $g = ah$ for some $h \in H$, then $g \in aH$.

Coming up with the element $h = a^{-1}g$ employed the familiar "working backwards" technique. Our goal was to show that $g = ah$ for some $h \in H$, so we solved this equation for h by left-multiplying each side by a^{-1}, which yielded $h = a^{-1}g$. As before, this process of solving for h is scratch work and does *not* belong in the proof. Instead, the focus of the argument is showing that $g = ah$ for $h = a^{-1}g$.

Theorem 19.15. *Let H be a subgroup of a group G. Then all the left cosets of H have the same size, namely the size of H.*

PROOF. Let $a \in G$. We will define a bijection from H to the coset aH. This will show that all cosets of H have the same size as H. Consider the function $f : H \to aH$ where $f(h) = ah$ for all $h \in H$. To show that f is one-to-one, suppose $f(h) = f(k)$ for some $h, k \in H$. Then $ah = ak$ and left cancellation would imply $h = k$. To show that f is onto, let $ah \in aH$ where $h \in H$. Then $f(h) = ah$. ■

Remark. Even if H were infinite, the above proof is valid. It would show that all cosets of H are infinite, each having a bijection from H.

19.5 When are cosets equal?

We have seen that cosets can be equal even when their coset representatives are different. For instance, let's revisit Example 19.1 with the group U_{13} and subgroup $H = \{1, 3, 9\}$. We found $2H = \{2, 6, 5\}$ and $6H = \{6, 5, 2\}$, so that $2H = 6H$, even though $2 \neq 6$ in U_{13}. But could we have determined that $2H = 6H$ *without* computing these cosets? More generally, is there a relationship between the coset representatives a and b that ensures that the cosets aH and bH are equal?

To answer these questions, we study an example whose operation is addition, since additive relationships tend to be easier to detect than their multiplicative counterparts. We revisit Example 19.5 with group $\mathbb{Z}_{12}$ and subgroup $H = \{0, 4, 8\}$. Here is what we found:

- $0 + H = 4 + H = 8 + H = \{0, 4, 8\}$.
- $1 + H = 5 + H = 9 + H = \{1, 5, 9\}$.
- $2 + H = 6 + H = 10 + H = \{2, 6, 10\}$.
- $3 + H = 7 + H = 11 + H = \{3, 7, 11\}$.

For instance, we have $2 + H = 6 + H$, and we seek an additive relationship between the coset representatives 2 and 6. We do have $2 + 6 = 8$, which is contained in the subgroup H. Perhaps the rule is: $a + H = b + H$ if and only if $a + b \in H$. But we also have $3 + H = 11 + H$ where $3 + 11 = 2$ (in $\mathbb{Z}_{12}$), which is not in H. Thus, our conjectured rule does not work in all cases.

Alternatively, we might try *subtracting* the coset representatives. For $2+H = 6+H$, we have $2-6 = 4$ (and $6-2 = 4$), which is in H. For $3+H = 11+H$, we have $3-11 = 4$ and $11 - 3 = 8$, and both differences are in H. Thus, we conjecture the following:

$$a + H = b + H \text{ if and only if } a - b \in H \text{ and } b - a \in H.$$

This rule even works with $5 + H = 5 + H$, since $5 - 5 = 0$ is in H. We also have $3 + H \neq 10 + H$, and $3 - 10 = 5$ and $10 - 3 = 7$, neither of which is in H. Thus, our conjecture seems promising.

Now let's translate this conjecture into the language of multiplicative groups. The expressions $a - b$ and $b - a$ could translate to $a \cdot b^{-1}$ and $b \cdot a^{-1}$. Thus, a conjecture for multiplicative groups may be

$$aH = bH \text{ if and only if } a \cdot b^{-1} \in H \text{ and } b \cdot a^{-1} \in H.$$

Let's verify this with the group U_{13} and subgroup $H = \{1, 3, 9\}$. For $2H = 6H$ (i.e., $a = 2$ and $b = 6$), we note that $2^{-1} = 7$ as $2 \cdot 7 = 1$ modulo 13 and $6^{-1} = 11$ as $6 \cdot 11 = 1$ modulo 13. Thus, $a \cdot b^{-1} = 2 \cdot 6^{-1} = 2 \cdot 11 = 9 \in H$ and $b \cdot a^{-1} = 6 \cdot 2^{-1} = 6 \cdot 7 = 3 \in H$. We also have $2H \neq 4H$, and $2 \cdot 4^{-1} = 2 \cdot 10 = 7$ and $4 \cdot 2^{-1} = 4 \cdot 7 = 2$, neither of which is in H. (Here, $4^{-1} = 10$, because $4 \cdot 10 = 1$ modulo 13.) Thus, the conjecture seems to work, both in concluding that $aH = bH$ and that $aH \neq bH$.

With the multiplicative case, it's instructive to test the conjecture with a non-commutative example. Let's use the group D_4 and subgroup $H = \{\varepsilon, d\}$ from Example 19.3. For $r_{90}H = hH$ (i.e., $a = r_{90}$ and $b = h$), we have $a \cdot b^{-1} = r_{90} \cdot h^{-1} = r_{90} \cdot h = d'$ and $b \cdot a^{-1} = h \cdot r_{90}^{-1} = h \cdot r_{270} = d'$, and yet $d' \notin H$. The conjecture fails, but how can we fix it? In a non-commutative group, the product $a \cdot b^{-1}$ does not necessarily equal $b^{-1} \cdot a$. Likewise, the products $b \cdot a^{-1}$ and $a^{-1} \cdot b$ need not be equal.

As a salvage, therefore, we rewrite $a \cdot b^{-1}$ and $b \cdot a^{-1}$ as $b^{-1} \cdot a$ and $a^{-1} \cdot b$, respectively. This change won't affect commutative groups such as U_{13}. Hence, our revised conjecture is

$$aH = bH \text{ if and only if } b^{-1} \cdot a \in H \text{ and } a^{-1} \cdot b \in H.$$

For $r_{90}H = hH$ (i.e., $a = r_{90}$ and $b = h$), we have $b^{-1} \cdot a = h^{-1} \cdot r_{90} = h \cdot r_{90} = d$ and $a^{-1} \cdot b = r_{90}^{-1} \cdot h = r_{270} \cdot h = d$, and $d \in H$. We also have $r_{90}H \neq vH$, and $v^{-1} \cdot r_{90} = v \cdot r_{90} = d'$ and $r_{90}^{-1} \cdot v = r_{270} \cdot v = d'$, which is not in H. The revised conjecture seems to correctly conclude that $aH = bH$ and that $aH \neq bH$.

We now state the revised conjecture as a theorem. The proof is written for a multiplicative group, and the proof for an additive group is left for you as an exercise.

Theorem 19.16. *Let G be a group, H a subgroup of G, and $a, b \in G$. Then:*

- *(Multiplicative) $aH = bH$ if and only if $b^{-1} \cdot a \in H$ and $a^{-1} \cdot b \in H$.*
- *(Additive) $a + H = b + H$ if and only if $a - b \in H$ and $b - a \in H$.*

Remark. The elements $b^{-1} \cdot a$ and $a^{-1} \cdot b$ are multiplicative inverses of each other. (Think socks-shoes.) Thus, they're both in H or neither is in H. Therefore, Theorem 19.16 could be stated as follows: $aH = bH$ *if and only if* $b^{-1} \cdot a \in H$. But, as we saw above with the group D_4 and subgroup $H = \{\varepsilon, d\}$, we cannot use the conditions $a \cdot b^{-1} \in H$ and $b \cdot a^{-1} \in H$.

PROOF. We must prove two implications:

- If $aH = bH$, then $b^{-1} \cdot a \in H$ and $a^{-1} \cdot b \in H$.
- If $b^{-1} \cdot a \in H$ and $a^{-1} \cdot b \in H$, the $aH = bH$.

We will prove the second implication. The proof of the first implication is left for you as an exercise.

Assume $b^{-1}a \in H$ and $a^{-1}b \in H$. To prove $aH = bH$, we will show that $aH \subseteq bH$ and $bH \subseteq aH$.

We start with $aH \subseteq bH$. Let $g \in aH$ so that $g = ah$ for some $h \in H$. (We must show that $g \in bH$). Moreover, since $b^{-1}a \in H$, we have $b^{-1}a = j$ for some $j \in H$. Left-multiplying both sides of $b^{-1}a = j$ by b, we obtain $a = bj$. Combining $g = ah$ and $a = bj$, we find $g = ah = (bj)h = b(jh) \in bH$, where $jh \in H$. Therefore, $g \in bH$ so that $aH \subseteq bH$.

By symmetry, we can deduce that $bH \subseteq aH$. Hence, we conclude that $aH = bH$, as desired. ∎

Proof know-how. The above proof contains the sentence, "By symmetry, we can deduce that $bH \subseteq aH$." This means that the argument for $bH \subseteq aH$ is *identical* to that for $aH \subseteq bH$, with the roles of a and b swapped. For instance, the first step in the argument would be "Let $g \in bH$ so that $g = bh$ for some $h \in H$." Rather than repeating what is essentially the same argument, we invoked the phrase "By symmetry." This proof-writing technique is often called *proof by symmetry*.

In Theorem 19.16, it's easy to get confused between the conditions $b^{-1} \cdot a,\ a^{-1} \cdot b \in H$ (which is correct) and $a \cdot b^{-1},\ b \cdot a^{-1} \in H$ (which is incorrect). Here's a mnemonic

device that can help. Starting with $aH = bH$, left-multiply both sides by b^{-1} to obtain $b^{-1}aH = b^{-1}bH$, which simplifies to $(b^{-1}a)H = H$. Then Theorem 19.14 implies that $b^{-1}a \in H$. Likewise, starting with $aH = bH$ and left-multiplying both sides by a^{-1} results in $a^{-1}b \in H$. We caution that this is merely a mnemonic for remembering the correct condition, and it does *not* constitute a proof of Theorem 19.16. In particular, the coset equality $b^{-1}aH = b^{-1}bH$ would require a more rigorous justification in an actual proof.

Example 19.17. Consider again the additive group $\mathbb{Z}$ and it subgroup $H = 5\mathbb{Z}$. Theorem 19.16 implies the following:

$$a + H = b + H \iff a - b \in H \iff 5 \mid (a - b) \iff a = b \text{ in } \mathbb{Z}_5.$$

Here, the symbol $\iff$ is a shorthand for "if and only if." Therefore, the cosets $a + H$ and $b + H$ are related in a way that resembles how a and b are related in $\mathbb{Z}_5$. We'll dig much more into this soon!

Exercises

When working with the group D_4, refer to Appendix B for its group table.

1. Consider the group U_{13} and its subgroup $H = \{1, 3, 9\}$. Then the coset $4H = \{4, 12, 10\}$ is *not* a subgroup of U_{13}. Describe all the group properties that are violated by $4H$.

2. A group G has 100 elements and its subgroup H has 5 elements. Determine the number of distinct cosets of H. Explain your reasoning.

3. Can a group G with 100 elements have a subgroup H with 12 elements? Why or why not? (This exercise is referenced in Chapter 20.)

4. Let $\alpha \in S_5$ be defined by

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

   Let $H = \langle \alpha \rangle$, i.e., the cyclic subgroup generated by α. Find the number of distinct cosets of H.

5. Let $H = \{0, 4, 8\}$ be a subgroup of $\mathbb{Z}_{12}$. For each $a \in \mathbb{Z}_{12}$, compute the coset $a + H$. (See Example 19.5.)

6. Let $H = \{0, 5, 10\}$ be a subgroup of $\mathbb{Z}_{15}$.

   (a) How many distinct cosets of H do you expect? Explain your reasoning.

   (b) For each $a \in \mathbb{Z}_{15}$, compute the coset $a + H$.

   (c) Verify Theorems 19.12, 19.14, 19.15, and 19.16 using the cosets you computed in part (b).

   (d) Verify that the distinct cosets of H form a partition of $\mathbb{Z}_{15}$.

7. Consider the multiplicative group U_{28} and its subset $H = \{1, 9, 25\}$.
   (a) Verify that H is indeed a subgroup of U_{28}.
   (b) How many distinct cosets of H do you expect? Explain your reasoning.
   (c) For each $a \in U_{28}$, compute the coset aH.
   (d) Verify Theorems 19.12, 19.14, 19.15, and 19.16 using the cosets you computed in part (c).
   (e) Verify that the distinct cosets of H form a partition of U_{28}.

8. Consider the additive group $\mathbb{Z}$ and its subgroup $H = 5\mathbb{Z}$. (See Example 19.6.)
   (a) Compute the cosets $12 + H$, $-1 + H$, $203 + H$, $-25 + H$, and $101 + H$.
   (b) Find all distinct cosets of H.
   (c) Verify that the distinct cosets of H form a partition of $\mathbb{Z}$.

9. Consider the additive group $\mathbb{Z}$ and its subgroup $H = 5\mathbb{Z}$. Determine whether or not the following cosets of H are equal.
   (a) $436 + H$ and $721 + H$.
   (b) $-43 + H$ and $111 + H$.
   (c) $317 + H$ and $532 + H$.

10. Consider the multiplicative group U_{35} and its subset $H = \{1, 8, 22, 29\}$.
    (a) Verify that H is indeed a subgroup of U_{35}.
    (b) How many distinct cosets of H do you expect? Explain your reasoning.
    (c) Without computing these cosets, determine if $11H = 18H$. (**Hint:** $18 \cdot 2 = 1$ modulo 35.)
    (d) Without computing these cosets, determine if $9H = 13H$. (**Hint:** $9 \cdot 4 = 1$ modulo 35.)
    (e) Without computing these cosets, determine if $3H = 24H$. (**Hint:** $3 \cdot 12 = 1$ modulo 35.)

11. Let $H = \{\varepsilon, \upsilon\}$ be a subgroup of D_4.
    (a) How many distinct cosets of H do you expect? Explain your reasoning.
    (b) For each $a \in D_4$, compute the left coset aH.
    (c) Find a pair of distinct elements $a, b \in D_4$ for which $aH = bH$. Verify that $b^{-1} \cdot a,\ a^{-1} \cdot b \in H$.
    (d) Find $a, b \in D_4$ for which $aH = bH$, but $b \cdot a^{-1}, a \cdot b^{-1} \notin H$.
    Note: So, you should be careful when using Theorem 19.16.

12. Write the proofs of Theorems 19.12, 19.14, 19.15, and 19.16 when the group operations is *addition*.

13. (a) For Theorem 19.16, give an analogous statement for right cosets.
    (b) Using an example, verify that the statement in part (a) correctly concludes that $Ha = Hb$ and that $Ha \neq Hb$.
    Note: It's recommended that you work on Exercise #14 in conjunction with this one.

    (This exercise is referenced in Chapter 20, Exercise #2.)

14. Again, let $H = \{\varepsilon, v\}$ be a subgroup of D_4.

    (a) For each $a \in D_4$, compute the right coset Ha.

    (b) Verify Theorems 19.12, 19.14, 19.15, and 19.16 (that are appropriately restated for right cosets) using the cosets you computed in part (a).
    Note: It's recommended that you work on Exercise #13 in conjunction with this one.

    (c) **True or False:** $aH = Ha$ for all $a \in D_4$.

    (This exercise is referenced in Example 24.14.)

15. Let $K = C(h) = \{\varepsilon, r_{180}, h, v\}$ be a subgroup of D_4. (See Example 19.9.)

    (a) For each $a \in D_4$, compute the left and right cosets aK and Ka.

    (b) Verify that $aK = Ka$ for all $a \in D_4$.

    (c) **True or False:** $aK = Ka$ means $ak = ka$ for each $k \in K$ (i.e., element-by-element equality).

    (This exercise is referenced in Example 24.5.)

16. Repeat Exercise #15 with $K = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$. (This exercise is referenced in Example 24.9 and Chapter 24, Exercise #5.)

17. Consider the group U_{13} and its subgroup $H = \{1\}$.

    (a) How many distinct cosets of H do you expect?

    (b) For each $a \in U_{13}$, compute the coset aH.

    (c) Using your result in part (b), complete this statement: $aH = bH$ if and only if _____.

18. Consider the group U_{13} and its subgroup $H = U_{13}$.

    (a) How many distinct cosets of H do you expect?

    (b) For each $a \in U_{13}$, compute the coset aH.

    (c) Using your result in part (b), complete this statement: $aH = bH$ if and only if _____.

19. Let G be a group and H its subgroup. Restate Theorem 19.16 for the cases $H = \{\varepsilon\}$ and $H = G$. How do your restatements compare with your answers in Exercises #17 and #18?

20. Let's generalize our work from Example 19.4. Consider the matrix group $G(\mathbb{Z}_{10})$ and its subgroup $H = S(\mathbb{Z}_{10}) = \{\alpha \in G(\mathbb{Z}_{10}) \mid \det \alpha = 1\}$. Define the following two sets:

    - Coset μH where μ is a fixed element $\mu = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$ with $\det \mu = 3$.
    - Set $T = \{\beta \in G(\mathbb{Z}_{10}) \mid \det \beta = 3\}$, i.e., the set of all matrices in $G(\mathbb{Z}_{10})$ with determinant 3.

    (a) Choose an element of the coset μH, i.e., a product $\mu \cdot h$ where $h \in H$. Show that this product is in set T by showing that it has determinant 3.

(b) Choose an element of set T, i.e., a matrix β with determinant 3. Show that β is in the coset μH by finding $h \in H$ such that $\beta = \mu \cdot h$.

Note: You must create examples that are different from the ones we used in Example 19.4.

21. Consider the matrix group $G(\mathbb{Z}_{10})$ and its subgroup $H = S(\mathbb{Z}_{10})$. Let $\mu \in G(\mathbb{Z}_{10})$ be a fixed element with $\det \mu = 3$, and define $T = \{\beta \in G(\mathbb{Z}_{10}) \mid \det \beta = 3\}$. Prove that $\mu H = T$. (This exercise is referenced in Section 25.2.)

Note: You must let μ be a general element of $G(\mathbb{Z}_{10})$ with $\det \mu = 3$, not a specific one like $\mu = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right]$.

22. **Prove:** Let H and K be subgroups of a group G. If $aH \subseteq bK$ for some $a, b \in G$, then $H \subseteq K$.

23. Complete the proof of Theorem 19.14 by proving its first implication.

24. Complete the proof of Theorem 19.16 by proving its first implication.

20

Lagrange's Theorem

Chapter 19, Exercise #3 asked the following:

> Can a group G with 100 elements have a subgroup H with 12 elements? Why or why not?

Here is a related question:

> Suppose there are lots of eggs and lots of dozen egg cartons, like the one shown in the figure below. If all the eggs are in cartons and all the cartons are full, can there be 100 eggs? Why or why not?

The answer to both questions is "No," and understanding the reason behind it will be the focus of this chapter. In particular, we will prove Lagrange's theorem, which states that if H is a subgroup of G with $\#H$ and $\#G$ elements, respectively, then $\#H$ is a divisor of $\#G$. Cosets will play a prominent role in proving this theorem, one of the most important results about finite groups.

20.1 Motivating Lagrange's theorem

Let G be a *finite* group, and let H be a subgroup of G. Denote the number of elements of G and H by $\#G$ and $\#H$, respectively. Then *Lagrange's theorem*, named after Joseph-Louis Lagrange, states that $\#H$ is a divisor of $\#G$. In this section, we will review several examples that motivate this theorem, highlighting the aspects of those examples that will play a role in its proof.

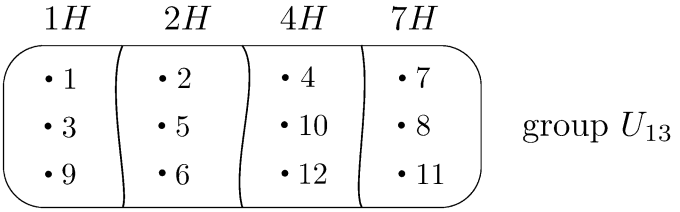
Example 20.1. Consider the group U_{13} and its subgroup $H = \{1, 3, 9\}$. In Example 19.1, we computed the cosets aH for each $a \in U_{13}$, i.e., $1H$, $2H$, $3H$, …, $12H$. We found several duplicates, and after consolidating those duplicates, we found four *distinct* cosets:

- $1H = \{1, 3, 9\}$.
- $2H = \{2, 5, 6\}$.
- $4H = \{4, 10, 12\}$.
- $7H = \{7, 8, 11\}$.

To say that these are the distinct cosets of H means that *any* coset of H must be equal to one of these. For instance, the coset $8H$ (which isn't on the above list) is equal to $7H$. More generally, a coset aH, where $a \in U_{13}$, must be equal to one of $1H$, $2H$, $4H$, or $7H$.

We recall two observations about these distinct cosets that will help in proving Lagrange's theorem. First, all cosets of H have the same size; namely $\#H = 3$. In fact, this was proved in Theorem 19.15.

Second, these distinct cosets form a *partition* of U_{13}, which means that the distinct cosets do not overlap and together cover all of U_{13}. This second observation is depicted in the figure below. Note how each element of U_{13} is contained in *exactly one* of the distinct cosets.



Example 20.2. Consider the group D_4 and its subgroup $H = \{\varepsilon, d\}$. In Example 19.3, we found the distinct left cosets of H, which are listed below along with the distinct right cosets of H.

Distinct left cosets of H:

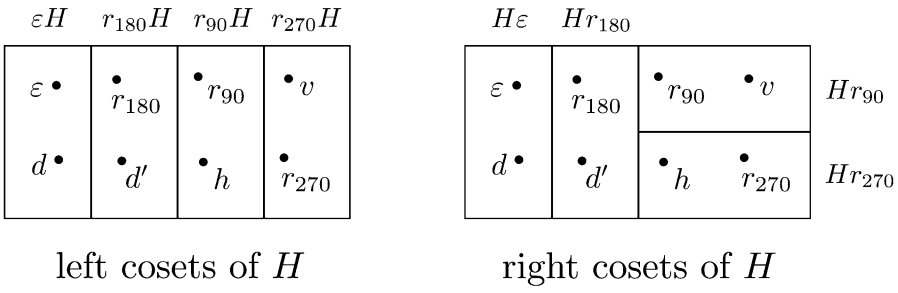
- $\varepsilon H = \{\varepsilon, d\}$.
- $r_{90}H = \{r_{90}, h\}$.
- $r_{180}H = \{r_{180}, d'\}$.
- $r_{270}H = \{r_{270}, v\}$.

Distinct right cosets of H:

- $H\varepsilon = \{\varepsilon, d\}$.
- $Hr_{90} = \{r_{90}, v\}$.
- $Hr_{180} = \{r_{180}, d'\}$.
- $Hr_{270} = \{r_{270}, h\}$.

As we observed in Example 19.8, the same coset representative can generate different left and right cosets, such as $r_{90}H \neq Hr_{90}$. This isn't too surprising, since D_4 is a *non-commutative* group. Nonetheless, all left and right cosets of H have the same size, namely $\#H = 2$, and the number of distinct left cosets equals the number of distinct right cosets (i.e., there are four of each type). The distinct left cosets of H form a partition of D_4, and the distinct right cosets of H do so as well. However, the manner in

which D_4 is partitioned differs between the left and right cosets, as shown in the figure below:



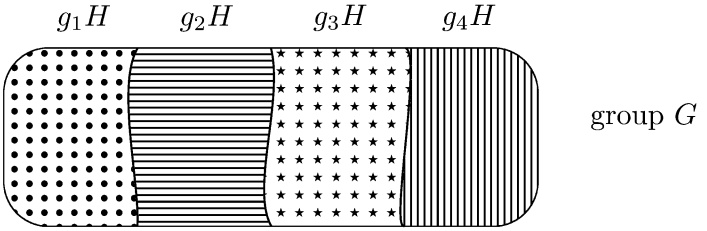
20.2 Proving Lagrange's theorem

We are now ready to prove Lagrange's theorem, which says: If H is a subgroup of a finite group G, then $\#H$ is a divisor of $\#G$. (Recall that $\#H$ and $\#G$ denote the number of elements of H and G, respectively.) Here are the key ingredients needed for its proof:

(1) All the cosets of H have the same size, namely $\#H$. (Proved in Theorem 19.15.)

(2) The distinct cosets of H form a *partition* of G; i.e., they cover all of G without overlapping.

Let's see why these two ingredients suffice to prove Lagrange's theorem. Suppose g_1H, g_2H, g_3H, …, g_nH are the distinct left cosets of H. The case of $n = 4$, i.e., four distinct cosets, is shown in the figure below:



Since these cosets form a partition of G, the number of elements in G equals the sum of the number of elements in each coset; i.e., $\#G = \#(g_1H) + \#(g_2H) + \#(g_3H) + \cdots + \#(g_nH)$. But all the cosets of H have the same size, namely $\#H$, and thus we obtain

$$\#G = \underbrace{\#H + \#H + \#H + \cdots + \#H}_{n \text{ terms}} = n \cdot \#H,$$

where n is the number of left cosets of H. Hence $\#G = n \cdot \#H$, so that $\#H$ is a divisor of $\#G$.

Remark. A helpful analogy is to think of the cosets of H as "tiling" the group G. And if the cosets of H were to tile G without any overlap, then it must be the case that $\#H$ is a divisor of $\#G$.

It remains to show that the distinct cosets of H form a *partition* of G. Thus, we will prove the following:

- Every element of G is contained in one of these cosets.
- Distinct cosets of H do not overlap with each other.

Remark. Like the theorems in Chapter 19, Theorems 20.3 and 20.4 below are stated in terms of left cosets. Analogous statements are true for right cosets (which you'll prove in an exercise at the end of the chapter), as well as for groups whose operation is addition.

Theorem 20.3. *Let H be a subgroup of a finite group G, and suppose g_1H, g_2H, g_3H, …, g_nH are the distinct left cosets of H. Then every element of G is contained in one of these cosets.*

PROOF. Consider an element $a \in G$. Then $a \in aH$ by Theorem 19.12. Moreover, aH must be equal to one of $g_1H, g_2H, g_3H, \ldots, g_nH$, say $aH = g_iH$. Since $a \in aH$, we have $a \in g_iH$ as desired. ∎

Next, we will show that distinct cosets of H do not overlap with each other. In other words, given cosets aH and bH, we must prove the following: If $aH \neq bH$, then aH and bH do not share any common element. Instead, we will prove the contrapositive; namely: If aH and bH do share a common element, then $aH = bH$.

Proof know-how. Why prove the contrapositive? Inequalities such as $aH \neq bH$ can be difficult to work with, since they indicate a *lack* of something. Instead, the hypothesis of the contrapositive, "aH and bH do share a common element," gives us something concrete to use, namely an element common to aH and bH.

Theorem 20.4. *Let H be a subgroup of a finite group G. If cosets aH and bH share a common element, then $aH = bH$.*

Remark. The last step of the proof below uses Theorem 19.16; i.e., $aH = bH$ if and only if $b^{-1}a \in H$.

PROOF. Assume aH and bH share a common element. Let g be an element contained in aH and bH. Thus, $g = ah$ and $g = bk$ for some $h, k \in H$, so that $ah = bk$. Take the equation $ah = bk$ and left-multiply by b^{-1} and right-multiply by h^{-1} to obtain $b^{-1}a = kh^{-1}$. Since $h, k \in H$ and H is a subgroup, we have $kh^{-1} \in H$. Hence $b^{-1}a \in H$, from which we conclude $aH = bH$. ∎

Proof know-how. In the above proof, one might wonder how we knew to take the equation $ah = bk$ and left-multiply by b^{-1} and right-multiply by h^{-1}. Our goal was to show $aH = bH$. Due to Theorem 19.16, that meant showing $b^{-1}a \in H$. Once we determined our goal (i.e., show that $b^{-1}a \in H$), the proof boiled down to manipulating the equation $ah = bk$ to solve for $b^{-1}a$ (and show that it's in H).

With its proof complete, we now state Lagrange's theorem.

Theorem 20.5 (Lagrange's theorem)**.** *Let H be a subgroup of a finite group G. Then $\#H$ is a divisor of $\#G$.*

Example 20.6. At the beginning of this chapter, we asked, "Can a group G with 100 elements have a subgroup H with 12 elements?" The answer is "No," because 12 is not a divisor of 100, and thus such a group and subgroup would violate Lagrange's theorem.

Consider a finite group G and its subgroup H. In the proof of Lagrange's theorem, we found that $\#G = n \cdot \#H$ where n is the number of distinct left cosets of H. Therefore, we concluded that $\#H$ is a divisor of $\#G$. From $\#G = n \cdot \#H$, we obtain the formula $n = \frac{\#G}{\#H}$. But we could have proved Lagrange's theorem using *right* cosets instead and derived the same conclusions. Thus, the same formula $n = \frac{\#G}{\#H}$ also applies to the number of distinct right cosets of H.

The above discussion prompts the following definition and theorem.

Definition 20.7 (Index of a subgroup)**.** Let G be a group, and let H be a subgroup of G. Then the *index* of H in G, denoted $[G : H]$, is the number of distinct left (or right) cosets of H in G.

Theorem 20.8. *Let H be a subgroup of a finite group G. Then $[G : H] = \frac{\#G}{\#H}$.*

Remark. Using the tiling analogy again, the index tells us how many cosets of H are needed to tile G.

We end the section with some examples of index calculations.

Example 20.9. In Example 19.1, we considered the group U_{13} and its subgroup $H = \{1, 3, 9\}$. There are four distinct left cosets of H, and so $[U_{13} : H] = 4$. Note that $\frac{\#U_{13}}{\#H} = \frac{12}{3} = 4$, which confirms Theorem 20.8.

Example 20.10. In Example 19.6, we computed the distinct cosets of $5\mathbb{Z}$ in $\mathbb{Z}$. Since $\mathbb{Z}$ is an infinite group, Theorem 20.8 does not apply. However, we found five distinct cosets, so that $[\mathbb{Z} : 5\mathbb{Z}] = 5$.

Example 20.11. In Example 20.2, we considered the group D_4 and its subgroup $H = \{\varepsilon, d\}$. We found four distinct left cosets and four distinct right cosets, and thus $[D_4 : H] = 4$. Note that $\frac{\#D_4}{\#H} = \frac{8}{2} = 4$, which confirms Theorem 20.8.

20.3 Applications of Lagrange's theorem

Lagrange's theorem is a powerful result that can provide many insights into the structure of finite groups. For instance, here is a conjecture that we had made in earlier chapters, which can now be proved using Lagrange's theorem. (See Chapter 4, Exercise #11; Chapter 5, Exercise #10; Chapter 6, Exercise #7.)

Theorem 20.12. *Let G be a finite group, and let $g \in G$. Then* $\operatorname{ord}(g)$ *is a divisor of $\#G$.*

PROOF. Let $n = \operatorname{ord}(g)$. By Theorem 13.17, the cyclic subgroup $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$ contains n distinct elements; namely $\langle g \rangle = \{\varepsilon, g^1, g^2, g^3, \ldots, g^{n-1}\}$. Since $\langle g \rangle$ is a subgroup of G, Lagrange's theorem implies that $\#\langle g \rangle$ is a divisor of $\#G$. Thus, $n = \operatorname{ord}(g)$ is a divisor of $\#G$, as desired. ∎

Example 20.13. Suppose a group G contains p elements, where p is prime. Let $g \in G$ be a non-identity element; i.e., $g \neq \varepsilon$. By Theorem 20.12, $\text{ord}(g)$ is a divisor of $\#G = p$. Since p is prime, its only positive divisors are 1 and p. And since $g \neq \varepsilon$, we know that $\text{ord}(g) \neq 1$. Thus we must have $\text{ord}(g) = p$, which implies that the cyclic subgroup $\langle g \rangle$ contains p elements; namely $\langle g \rangle = \{\varepsilon, g^1, g^2, g^3, \ldots, g^{p-1}\}$. Since G also has p elements, we have $G = \langle g \rangle$, so that G is cyclic with generator g.

The result of Example 20.13 is summarized in the following theorem.

Theorem 20.14. *Let G be a group with p elements, where p is prime. Then G is cyclic with $G = \langle g \rangle$, where g is any non-identity element of G.*

The next theorem is a direct consequence of Theorem 20.14. Moreover, its proof introduces a new proof-writing technique, which is described in the Proof know-how below. (**Note:** Recall that $H \cap K$ is the *intersection* of H and K, i.e., the set of elements that are contained in both subgroups.)

Theorem 20.15. *Let G be a group with subgroups H and K. Suppose $\#H = \#K = p$, where p is prime. Then $H = K$ or $H \cap K = \{\varepsilon\}$.*

Proof know-how. In this theorem, we must prove an "or" statement; i.e., we must prove that (1) $H = K$ or (2) $H \cap K = \{\varepsilon\}$. Here is a possible approach. We know that conclusion (2) is either true or false. If it's true, then we're done with the proof. Thus, we will *assume* that (2) is false and *prove* that (1) is true. In other words, we will prove the following implication: If (2) is false, then (1) is true. In the actual proof, we typically leave out the rationale behind this approach, i.e., about how (2) is either true or false and we're done with the proof if (2) is true. Instead, we start right away with the assumption that (2) is false.

Now, we could also assume that (1) is false and prove that (2) is true, which happens to be the contrapositive of "if (2) is false, then (1) is true." In this case, it turns out that "if (2) is false, then (1) is true" is easier to prove. Like many aspects of proof writing, a feel for choosing which implication to prove comes with lots of experience. If you're not sure which one to prove, try both and see what happens!

PROOF. Assume that $H \cap K \neq \{\varepsilon\}$. Hence, there is a non-identity element g that is contained in both H and K. We will show that $H = K$. Since H has p elements, Theorem 20.14 implies that it is cyclic with $H = \langle g \rangle$. Likewise, we have $K = \langle g \rangle$. Therefore $H = K$, as both are equal to $\langle g \rangle$. ■

Example 20.16. Let G be a group with 35 elements. We'll prove that G contains an element of order 5. Let $g \in G$ be a non-identity element. By Theorem 20.12, $\text{ord}(g)$ is a divisor of $\#G = 35$. The positive divisors of 35 are 1, 5, 7, and 35. Since $g \neq \varepsilon$, we have $\text{ord}(g) = 5$, 7, or 35. If $\text{ord}(g) = 5$, then we're done with our proof. If $\text{ord}(g) = 35$, then $\text{ord}(g^7) = 5$ by Theorem 12.7 and we're also done.

But what if all 34 non-identity elements of G have order 7? We'll show this isn't possible, and thus G must have an element of order 5. Suppose $g_1 \in G$ has order 7. Then the cyclic subgroup $\langle g_1 \rangle$ contains 6 non-identity elements, each with order 7. Next, let $g_2 \in G$ have order 7, with $g_2 \notin \langle g_1 \rangle$. Then $\langle g_2 \rangle$ also contains 6 non-identity

elements, each with order 7. Moreover, Theorem 20.15 implies that $\langle g_1 \rangle \cap \langle g_2 \rangle = \{\varepsilon\}$. Therefore, elements of order 7 come in disjoint "clumps" of 6 elements. Since 6 is not a divisor of 34, there cannot be 34 elements of order 7 in G.

Exercises

1. Consider the group D_4 and its subgroup $H = \{\varepsilon, r_{180}, d, d'\}$.

   Note: We have $H = C(d)$, i.e., the *centralizer* of d in D_4. Thus, H is indeed a subgroup.

   (a) Find $[D_4 : H]$.

   (b) Let $a \in H$. *Without* using the group table for D_4, compute the left coset aH and right coset Ha.

   (c) Repeat part (b) for $a \notin H$. Explain your reasoning.

   (d) Explain why $aH = Ha$ for all $a \in D_4$.

2. For Theorems 20.3 and 20.4, write analogous statements for right cosets and prove them.

   Note: For Theorem 20.4, you should first complete Chapter 19, Exercise #13.

3. Let G be a finite group, H a subgroup of G, and K a subgroup of H (thus $K \subseteq H \subseteq G$). Prove that $[G : K] = [G : H] \cdot [H : K]$.

4. Let G be a group, and let H and K be its subgroups. Define $M = H \cap K = \{g \in G \mid g \in H \text{ and } g \in K\}$; i.e., M is the *intersection* of H and K. If $\#H = 15$ and $\#K = 28$, find $\#M$. Explain your reasoning.

   Hint: See Chapter 11, Exercise #13.

5. **Prove:** Suppose G is a group and H and K are subgroups of G containing m and n elements, respectively. If $\gcd(m, n) = 1$, then $H \cap K = \{\varepsilon\}$.

   Note: Compare with Chapter 14, Exercise #24.

6. Prove each of the following statements.

   (a) If the group U_m contains k elements, then $a^k = 1$ for all $a \in U_m$. (See Chapter 4, Exercise #10.)

   (b) If a group G contains k elements, then $g^k = \varepsilon$ for all $g \in G$.

7. **Prove:** Let G be a group with k elements. Suppose $\gcd(k, n) = 1$. If $g \in G$ and $g^n = \varepsilon$, then $g = \varepsilon$.

   Hint: Use Theorem 3.9, i.e., the GCD theorem.

8. Consider the prime number $p = 3$.

   (a) Choose an integer a, compute $a^p - a$, and verify that p is a divisor of $a^p - a$.

   (b) Repeat part (a) with another integer a of your choice.

   (c) Repeat part (a) again, this time with a negative integer a.

9. (a) Repeat Exercise #8 with prime $p = 5$; with prime $p = 7$; with prime $p = 11$.
   (b) Repeat Exercise #8 with one more prime number of your choice.
   (c) What conjecture do you have?

10. **(Fermat's little theorem)** Let p be a prime number. Prove that p is a divisor of $a^p - a$ for all $a \in \mathbb{Z}$.

11. Let G be a group with 15 elements, and let H be a *proper* subgroup of G. Explain why H is cyclic.

    Note: A *proper* subgroup of G is a subgroup that is not G itself.

12. Repeat Exercise #11 for a group G with 21 elements; with 33 elements; with 91 elements.

13. **Prove:** Suppose G is a group with pq elements, where p and q are distinct prime numbers. If H is a proper subgroup of G, then H is cyclic.

14. Prove each of the following statements.
    (a) If G is a group with 27 elements, then there exists an element $g \in G$ with $\text{ord}(g) = 3$.
    (b) Suppose G is a group with p^n elements, where p is a prime number and n is a positive integer. Then there exists an element $g \in G$ with $\text{ord}(g) = p$.

15. Let G be a group with 49 elements. If G is *not* cyclic, what can you say about the order of each element in G?

16. (a) Repeat Exercise #15 for a group with 25 elements; with 169 elements; with 289 elements.
    (b) Write a statement that generalizes part (a). Then prove your statement.

17. Let G be a group with 55 elements.
    (a) For $g \in G$, find the possible values of $\text{ord}(g)$.
    (b) **Prove:** There exists an element of G with order 11.

18. Let G be a group with 40 elements.
    (a) For $g \in G$, find the possible values of $\text{ord}(g)$.
    (b) **Prove:** There exists an element of G with order 2.

19. **Prove:** Let G be a group with $p \cdot 2^n$ elements, where p is an odd prime and n is a positive integer. Then there exists an element $g \in G$ with $\text{ord}(g) = 2$.

20. **Prove:** Let G be a group with n elements. If n is odd, then G has no element of order 2.

21. **Prove:** Let G be a commutative group with n elements. Let a be the product of all elements of G. If n is odd, then $a = \varepsilon$.

22. Let G be a commutative group with 21 elements. Consider the function $\theta : G \to G$ where $\theta(g) = g^{16}$ for all $g \in G$. Show that θ is an isomorphism.

23. Let G be a commutative group with n elements. Consider the function $\theta : G \to G$ where $\theta(g) = g^m$ for all $g \in G$. If $\gcd(m, n) = 1$, then show that θ is an isomorphism.

21

Multiplying/Adding Cosets

In Chapter 20, we saw the instrumental role that cosets play in the proof of Lagrange's theorem. As noted in the introduction to this unit, I often tell students, "If proving Lagrange's theorem is all that cosets are good for, then cosets would still hold a special place in group theory." However, cosets can do so much more, and we'll learn about their further exploits in the next few chapters.

Given a group G and its subgroup H, the distinct cosets of H form a group called a *quotient group*, provided that H meets a certain condition (TBA). In this chapter, we'll learn the group operation for the quotient group, i.e., how to multiply (or add) cosets.

21.1 Turning a set of cosets into a group

We begin by revisiting Example 19.1. Let $H = \{1, 3, 9\}$ be a subgroup of U_{13}. The distinct cosets of H are the following:

- $1H = \{1, \ 3, \ 9\} = \ 3H = \ 9H$.
- $2H = \{2, \ 6, \ 5\} = \ 6H = \ 5H$.
- $4H = \{4, 12, 10\} = 12H = 10H$.
- $7H = \{7, \ 8, 11\} = \ 8H = 11H$.

We can use the property $a \in aH$ (Theorem 19.12) and that distinct cosets do not overlap (Theorem 20.4) when finding these coset representatives. For instance, once we compute $2H = \{2 \cdot 1, 2 \cdot 3, 2 \cdot 9\} = \{2, 6, 5\}$, then we know that this coset also equals $6H$ and $5H$, because 6 and 5 are contained in $2H$.

Notation. We define U_{13}/H (read "U_{13} mod H") to be the set of distinct cosets of H. Thus,

$$U_{13}/H = \{1H, \ 2H, \ 4H, \ 7H\}.$$

Since different coset representatives can generate the same coset (e.g., $2H = 6H$), we could have written U_{13}/H slightly differently, perhaps like this:

$$U_{13}/H = \{1H, \ 6H, \ 10H, \ 11H\}.$$

It does make sense to use $1H$ instead of $3H$ or $9H$, given that $1 \in U_{13}$ is a special element, namely the multiplicative identity. (We'll soon see the special role that $1H$ plays in U_{13}/H.) However, using $2H$ instead of $6H$ is simply a matter of choice.

Here's a crazy idea: We wish to turn the set U_{13}/H into a *group*. That means we need an *operation*, i.e., a way to "multiply" a pair of cosets. For instance, what would $2H \cdot 4H$ equal? By closure, it would have to equal $1H$, $2H$, $4H$, or $7H$. But which one? To answer this question, we begin by defining what it means to multiply two subsets of a group. (**Recall:** This was first defined in Chapter 8, Exercise #9.)

Definition 21.1. Let S and T be subsets of a group G. Then the *set product* of S and T is the set

$$S \cdot T = \{s \cdot t \mid s \in S,\ t \in T\},$$

where the multiplication $s \cdot t$ is done in G.

Example 21.2. Let $H = \{1, 3, 9\}$ be a subgroup of U_{13}. To compute the set product $2H \cdot 4H$, we multiply every element of $2H = \{2, 6, 5\}$ by every element of $4H = \{4, 12, 10\}$, as shown below:

$$\begin{aligned} 2H \cdot 4H &= \{2, 6, 5\} \cdot \{4, 12, 10\} \\ &= \{2 \cdot 4,\ 2 \cdot 12,\ 2 \cdot 10,\ 6 \cdot 4,\ 6 \cdot 12,\ 6 \cdot 10,\ 5 \cdot 4,\ 5 \cdot 12,\ 5 \cdot 10\} \\ &= \{8, 11, 7, 11, 7, 8, 7, 8, 11\}. \end{aligned}$$

Since $2H$ and $4H$ contain 3 elements each, it may seem that $2H \cdot 4H$ would contain 9 elements of U_{13}. But we see several duplicates in the set product, and duplicate elements are not counted in a set. After consolidating the duplicates, we obtain $2H \cdot 4H = \{7, 8, 11\}$. Therefore, $2H \cdot 4H = 7H$.

Example 21.3. Consider again the subgroup $H = \{1, 3, 9\}$ of U_{13}. Let's compute the set product $4H \cdot 2H$ and compare the result with $2H \cdot 4H$, which we found in Example 21.2 above:

$$\begin{aligned} 4H \cdot 2H &= \{4, 12, 10\} \cdot \{2, 6, 5\} \\ &= \{4 \cdot 2,\ 4 \cdot 6,\ 4 \cdot 5,\ 12 \cdot 2,\ 12 \cdot 6,\ 12 \cdot 5,\ 10 \cdot 2,\ 10 \cdot 6,\ 10 \cdot 5\} \\ &= \{8, 11, 7, 11, 7, 8, 7, 8, 11\} \\ &= \{7, 8, 11\} \\ &= 7H. \end{aligned}$$

We obtain $4H \cdot 2H = 7H$, so that $2H \cdot 4H = 4H \cdot 2H$. Perhaps this was expected, since multiplication of cosets in this example and in Example 21.2 is based on multiplication in U_{13}, which is commutative. In an exercise at the end of the chapter, you'll prove that if a group G is commutative, then the corresponding coset multiplication is also commutative; i.e., $aH \cdot bH = bH \cdot aH$.

Example 21.4. For the subgroup $H = \{1, 3, 9\}$ of U_{13}, we have the set of distinct cosets

$$U_{13}/H = \{1H, 2H, 4H, 7H\}.$$

We'll now see the special role that $1H$ plays in coset multiplication. Consider the set product:

$$\begin{aligned} 1H \cdot 4H &= \{1, 3, 9\} \cdot \{4, 12, 10\} \\ &= \{1 \cdot 4, 1 \cdot 12, 1 \cdot 10, 3 \cdot 4, 3 \cdot 12, 3 \cdot 10, 9 \cdot 4, 9 \cdot 12, 9 \cdot 10\} \\ &= \{4, 12, 10, 12, 10, 4, 10, 4, 12\} \\ &= \{4, 10, 12\} \\ &= 4H. \end{aligned}$$

Therefore, $1H \cdot 4H = 4H$. Based on Example 21.3, we also have $4H \cdot 1H = 4H$. In an exercise at the end of the chapter, you'll show that $1H \cdot aH = aH$ and $aH \cdot 1H = aH$ for all cosets aH. Hence, $1H$ is the multiplicative identity element of U_{13}/H.

Example 21.5. Let $H = \{1, 3, 9\}$ be a subgroup of U_{13} and consider the set of distinct cosets

$$U_{13}/H = \{1H, 2H, 4H, 7H\}.$$

We'll leave it up to you to verify the following set products: $1H \cdot 1H = 1H$, $2H \cdot 7H = 1H$ (and hence $7H \cdot 2H = 1H$), and $4H \cdot 4H = 1H$. Since $1H$ is the multiplicative identity of U_{13}/H (see Example 21.4), we conclude that $2H$ and $7H$ are an inverse pair and $1H$ and $4H$ are self-inverses.

After computing all the set products, we obtain the following table for $U_{13}/H = \{1H, 2H, 4H, 7H\}$ where the operation is coset multiplication:

$\cdot$	$1H$	$2H$	$4H$	$7H$
$1H$	$1H$	$2H$	$4H$	$7H$
$2H$	$2H$	$4H$	$7H$	$1H$
$4H$	$4H$	$7H$	$1H$	$2H$
$7H$	$7H$	$1H$	$2H$	$4H$

We use this table to verify the group properties for U_{13}/H.

(1) U_{13}/H is closed under coset multiplication. We can see this from the table, since every entry in the table (i.e., all possible "products") is an element of U_{13}/H.

(2) Coset multiplication is associative. See Theorem 21.6 below for a justification.

(3) U_{13}/H contains the multiplicative identity element $1H$, where $1H \cdot aH = aH$ (first row of the table) and $aH \cdot 1H = aH$ (first column of the table) for all $aH \in U_{13}/H$.

(4) Every element in U_{13}/H has a multiplicative inverse. $2H$ and $7H$ are multiplicative inverses of each other, and $1H$ and $4H$ are self-inverses.

Thus, U_{13}/H is a group under coset multiplication.

Remark. The key here is that we are treating each coset aH as an *element* of U_{13}/H.

Using the group table, we can compute the order of each $aH \in U_{13}/H$, i.e., the smallest number of times we multiply aH by itself to obtain the multiplicative identity $1H$.

- $\text{ord}(1H) = 1$, because $(1H)^1 = 1H$.
- $\text{ord}(2H) = 4$, because $(2H)^1 = 2H$, $(2H)^2 = 4H$, $(2H)^3 = 7H$, and $(2H)^4 = 1H$.
- $\text{ord}(4H) = 2$, because $(4H)^1 = 4H$ and $(4H)^2 = 1H$.
- $\text{ord}(7H) = 4$, because $(7H)^1 = 7H$, $(7H)^2 = 4H$, $(7H)^3 = 2H$, and $(7H)^4 = 1H$.

Therefore, U_{13}/H is a *cyclic* group with generators $2H$ and $7H$; i.e., $U_{13}/H = \langle 2H \rangle = \langle 7H \rangle$.

The theorem below shows that coset multiplication is associative, for any group G and its subgroup H. Note that we must show the set equality $(aH \cdot bH) \cdot cH = aH \cdot (bH \cdot cH)$, which entails showing the set inclusions $(aH \cdot bH) \cdot cH \subseteq aH \cdot (bH \cdot cH)$ and $aH \cdot (bH \cdot cH) \subseteq (aH \cdot bH) \cdot cH$.

Theorem 21.6 (Associativity of coset multiplication)**.** *Let G be a group, H a subgroup of G, and $a, b, c \in G$. Then $(aH \cdot bH) \cdot cH = aH \cdot (bH \cdot cH)$.*

PROOF. Let $(\alpha \cdot \beta) \cdot \gamma \in (aH \cdot bH) \cdot cH$ where $\alpha \in aH$, $\beta \in bH$, and $\gamma \in cH$. But α, β, and γ are elements of G. So we use the associativity of multiplication in G to get

$$(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma) \in aH \cdot (bH \cdot cH).$$

Therefore, $(aH \cdot bH) \cdot cH \subseteq aH \cdot (bH \cdot cH)$. The argument for the other set inclusion follows similarly. Thus, $(aH \cdot bH) \cdot cH = aH \cdot (bH \cdot cH)$, as desired. ■

21.2 Coset multiplication shortcut

As we saw in Examples 21.2, 21.3, and 21.4, coset multiplication can be a tedious process. Fortunately, there is a shortcut. To motivate this shortcut, let's examine the earlier examples in more depth.

Example 21.7. Consider the subgroup $H = \{1, 3, 9\}$ of U_{13}. We found the coset product $2H \cdot 4H = 7H$. But $7H$ can be written as $8H$ (i.e., they're the same coset). Thus $2H \cdot 4H = 8H$, where we observe that $2 \cdot 4 = 8$ in U_{13}. We also found that $1H \cdot 4H = 4H$, where $1 \cdot 4 = 4$ in U_{13}. For the product $4H \cdot 4H = 1H$, we observe that $1H$ can also be written as $3H$. Hence we have $4H \cdot 4H = 3H$, where $4 \cdot 4 = 3$ in U_{13}.

Example 21.7 above suggests that to find the coset product $aH \cdot bH$ in G/H, we can simply multiply their coset representatives in G; i.e., $aH \cdot bH = (ab)H$. The theorem below is for the case when G is commutative, such as $G = U_{13}$. Its proof is left for you as an exercise at the end of the chapter.

Theorem 21.8 (Coset multiplication shortcut)**.** *Let G be a* commutative *group, H a subgroup of G, and $a, b \in G$. Define the coset product by $aH \cdot bH = \{\alpha \cdot \beta \mid \alpha \in aH, \beta \in bH\}$. Then $aH \cdot bH = (ab)H$.*

Example 21.9. Consider again the cosets $2H = \{2, 6, 5\}$ and $4H = \{4, 12, 10\}$ of U_{13}. Rather than multiplying every element of $2H$ by every element of $4H$, we use the coset multiplication shortcut to compute $2H \cdot 4H$. Note that the shortcut applies in this case, because U_{13} is commutative. We have

$$2H \cdot 4H = (2 \cdot 4)H = 8H.$$

But $8H = 7H$, so that $2H \cdot 4H = 7H$ as before.

Remark. In Example 21.9 above, we have the the equality $2H \cdot 4H = (2 \cdot 4)H$. There's a subtle distinction between the multiplication symbol $\cdot$ in each side of the equation. On the left side, the expression $2H \cdot 4H$ denotes a product of cosets, i.e., multiplication in U_{13}/H. In the expression $(2 \cdot 4)H$, the product $2 \cdot 4$ depicts a product of coset representatives, which occurs in U_{13}.

Example 21.10 (Non-example)**.** Let's look at a non-commutative example. Consider the group D_4 and its subgroup $H = \{\varepsilon, \upsilon\}$. To compute the set product $r_{90}H \cdot d'H$, we multiply every element of $r_{90}H = \{r_{90}, d\}$ by every element of $d'H = \{d', r_{270}\}$, as shown below:

$$\begin{aligned} r_{90}H \cdot d'H &= \{r_{90}, d\} \cdot \{d', r_{270}\} \\ &= \{r_{90} \cdot d', r_{90} \cdot r_{270}, d \cdot d', d \cdot r_{270}\} \\ &= \{\upsilon, \varepsilon, r_{180}, h\}. \end{aligned}$$

Thus, $r_{90}H \cdot d'H = \{\upsilon, \varepsilon, r_{180}, h\}$. But $(r_{90} \cdot d')H = \upsilon H = \{\upsilon, \varepsilon\}$, so we see that $r_{90}H \cdot d'H \neq (r_{90} \cdot d')H$. The coset multiplication shortcut fails! In fact, the set product $r_{90}H \cdot d'H$ contains 4 elements, and therefore it's not even a coset of H.

Based on Example 21.10 above, it's natural to ask: Does the coset multiplication shortcut in G/H hold *only* when G is commutative? Not quite. You'll see in the exercises at the end of this chapter that the shortcut can hold in D_4/H for some subgroups H, even though D_4 is non-commutative. The precise condition for when the coset multiplication shortcut holds will be revealed in Chapter 24. Stay tuned!

21.3 Cosets of $H = 5\mathbb{Z}$ in $\mathbb{Z}$ revisited

Consider the additive group $\mathbb{Z}$ and its subgroup $H = 5\mathbb{Z}$. We saw in Example 19.6 that the distinct cosets of H in $\mathbb{Z}$ are as follows:

- $\cdots = -5 + H = \mathbf{0} + \boldsymbol{H} = 5 + H = 10 + H = 15 + H = \cdots$ (original subgroup).
- $\cdots = -4 + H = \mathbf{1} + \boldsymbol{H} = 6 + H = 11 + H = 16 + H = \cdots$.
- $\cdots = -3 + H = \mathbf{2} + \boldsymbol{H} = 7 + H = 12 + H = 17 + H = \cdots$.
- $\cdots = -2 + H = \mathbf{3} + \boldsymbol{H} = 8 + H = 13 + H = 18 + H = \cdots$.
- $\cdots = -1 + H = \mathbf{4} + \boldsymbol{H} = 9 + H = 14 + H = 19 + H = \cdots$.

We define $\mathbb{Z}/H$ (read "$\mathbb{Z}$ mod H") to be the set of distinct cosets of H. Thus,

$$\mathbb{Z}/H = \{0 + H, 1 + H, 2 + H, 3 + H, 4 + H\}.$$

As with U_{13}/H in Section 21.1, we could have written $\mathbb{Z}/H$ slightly differently, since different coset representatives can generate the same coset (e.g., $2+H = 7+H$). Here is one possibility:

$$\mathbb{Z}/H = \{0+H,\ 11+H,\ 7+H,\ -2+H,\ 1049+H\}.$$

However, $\mathbb{Z}/H = \{0+H,\ 1+H,\ 2+H,\ 3+H,\ 4+H\}$ is a natural choice, as we will see below.

We will now show that $\mathbb{Z}/H$ is an (additive) group under coset addition. To add, for example, the cosets $2+H$ and $3+H$, we add every element of $2+H$ to those of $3+H$. Thus, we have

$$\begin{aligned}(2+H)+(3+H)&\\ &= \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\} + \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\}\\ &= \{\ldots, -25, -20, -15, -10, -5, 0, 5, 10, 15, 20, 25, 30, 35, \ldots\}\\ &= 0+H.\end{aligned}$$

When computing $(2+H)+(3+H)$, we encounter duplicates, infinitely many of them, in fact. For instance, the integer 20 in the coset sum is obtained by $-13+33$, $-8+28$, $-3+23$, $2+18$, $7+13$, and so on, each of which is the sum of an element in $2+H$ and an element in $3+H$.

You might have noticed that $(2+H)+(3+H) = (2+3)+H$. This is true not only in $\mathbb{Z}/H$, but also in any G/H where G is an additive group. Below is the additive version of the coset multiplication shortcut (Theorem 21.8). Note that additive groups are always commutative, so that the shortcut should hold.

Theorem 21.11 (Coset addition shortcut)**.** *Let G be an additive group, H a subgroup of G, and $a, b \in G$. Define the coset sum by $(a+H)+(b+H) = \{\alpha+\beta \mid \alpha \in a+H,\ \beta \in b+H\}$. Then $(a+H)+(b+H) = (a+b)+H$.*

PROOF. To show this set equality, we must show that

$$(a+H)+(b+H) \subseteq (a+b)+H \quad \text{and} \quad (a+b)+H \subseteq (a+H)+(b+H).$$

First, let $\alpha+\beta \in (a+H)+(b+H)$, where $\alpha \in a+H$ and $\beta \in b+H$. Thus, $\alpha = a+h$ and $\beta = b+k$ for some $h, k \in H$. Since additive groups are commutative,

$$\alpha+\beta = (a+h)+(b+k) = (a+b)+(h+k) \in (a+b)+H.$$

Thus $\alpha+\beta \in (a+b)+H$, so that $(a+H)+(b+H) \subseteq (a+b)+H$.

Next, let $\gamma \in (a+b)+H$ so that $\gamma = (a+b)+h$ for some $h \in H$. Then,

$$\gamma = (a+b)+h = (a+0)+(b+h) \in (a+H)+(b+H).$$

Thus $\gamma \in (a+H)+(b+H)$, so that $(a+b)+H \subseteq (a+H)+(b+H)$.

Therefore, $(a+H)+(b+H) = (a+b)+H$ as desired. ■

Proof know-how. In the above proof, we showed $\gamma = (a+b)+h = (a+0)+(b+h)$ by rewriting a as $a+0$. This "inserting the (additive) identity" technique allowed us to conclude that γ is in $(a+H)+(b+H)$, since $a+0 \in a+H$ and $b+h \in b+H$. (Compare this with the proofs of Theorems 9.6 and 17.9.)

Using the shortcut, we create a group table for $\mathbb{Z}/H$:

$+$	$0+H$	$1+H$	$2+H$	$3+H$	$4+H$
$0+H$	$0+H$	$1+H$	$2+H$	$3+H$	$4+H$
$1+H$	$1+H$	$2+H$	$3+H$	$4+H$	$0+H$
$2+H$	$2+H$	$3+H$	$4+H$	$0+H$	$1+H$
$3+H$	$3+H$	$4+H$	$0+H$	$1+H$	$2+H$
$4+H$	$4+H$	$0+H$	$1+H$	$2+H$	$3+H$

Then we use the table to verify the group properties for $\mathbb{Z}/H$.

(1) $\mathbb{Z}/H$ is closed under coset addition, since every entry in the table is an element of $\mathbb{Z}/H$.

(2) Coset addition is associative. The proof looks similar to that of Theorem 21.6. You'll fill in the details in an exercise at the end of the chapter.

(3) $\mathbb{Z}/H$ contains the identity element $0+H$, where $(0+H)+(a+H)=a+H$ (first row of the table) and $(a+H)+(0+H)=a+H$ (first column of the table) for all $a+H \in \mathbb{Z}/H$.

(4) Every element in $\mathbb{Z}/H$ has an additive inverse, because each row (and column) of the table contains the additive identity element $0+H$.

Thus, $\mathbb{Z}/H$ is a group. To which familiar group is it isomorphic? (**Hint:** Ignore the "$+H$" in the table.)

Exercises

When working with the group D_4, refer to Appendix B for its group table.

1. Let $H=\{1,\ 3,\ 9\}$ be a subgroup of U_{13}. We saw in Example 21.4 that $1H \cdot 4H = 4H$.
   (a) Compute the coset product $1H \cdot 1H$ by multiplying each element of $1H$ by those of $1H$.
   (b) Repeat part (a) to compute $1H \cdot 2H$.
   (c) Repeat part (a) to compute $1H \cdot 7H$.
   (d) What conclusion can you make about the element $1H$ in U_{13}/H?

2. Let $H=\{1,\ 3,\ 9\}$ be a subgroup of U_{13}. We saw in Exercise #1 that $1H \cdot 1H = 1H$.
   (a) Compute the coset $2H \cdot 7H$ by multiplying each element of $2H$ by those of $7H$.
   (b) Repeat part (a) to compute $7H \cdot 2H$.
   (c) Repeat part (a) to compute $4H \cdot 4H$.
   (d) What conclusion can you make about the pair $2H$ and $7H$? About $4H$ itself?

3. In Chapter 8, Exercise #9, we computed set products using the following subsets of U_7:
   $$E=\{1,\ 6\},\ S=\{2,\ 5\},\ T=\{3,\ 4\}.$$
   One of these is a subgroup of U_7 and the other two are cosets of that subgroup. Find the subgroup (call it H) and write the other two sets as cosets of H.

4. Let G be a group, H a subgroup of G, and $a, b \in G$. Define the coset product by

$$aH \cdot bH = \{\alpha \cdot \beta \mid \alpha \in aH,\ \beta \in bH\}.$$

Prove: If G is commutative, then $aH \cdot bH = bH \cdot aH$. (See Example 21.3.)

5. Prove Theorem 21.6 when the group operation is addition: Let G be an *additive* group, H a subgroup of G, and $a, b, c \in G$. Then $(aH + bH) + cH = aH + (bH + cH)$.

6. Let $H = \{1, 3, 9\}$ be a subgroup of U_{13}. Shown below is the table for $U_{13}/H = \{1H,\ 2H,\ 4H,\ 7H\}$, where the operation is coset multiplication. Use this table to verify the coset multiplication shortcut for each pair $aH, bH \in U_{13}/H$. (See Example 21.9.)

$\cdot$	$1H$	$2H$	$4H$	$7H$
$1H$	$1H$	$2H$	$4H$	$7H$
$2H$	$2H$	$4H$	$7H$	$1H$
$4H$	$4H$	$7H$	$1H$	$2H$
$7H$	$7H$	$1H$	$2H$	$4H$

7. Consider the subgroup $Z = \{\varepsilon,\ r_{180}\}$ of D_4. Recall that $Z = \{z \in G \mid zg = gz \text{ for all } g \in D_4\}$ is the *center* of D_4, i.e., the set of elements of D_4 that commute with all elements of D_4.

   (a) Quick! How many distinct left (or right) cosets of Z are there? Explain how you know.

   (b) For each $a \in D_4$, compute the left and right cosets aZ and Za.

   (c) Verify that $aZ = Za$ for all $a \in D_4$.

   (d) Elizabeth says, "Since Z is the center of D_4, it's not surprising to see that the left and right cosets were the same in part (c)." What might she mean?

8. In Exercise #7, we found that the distinct left cosets of Z are $D_4/Z = \{\varepsilon Z,\ r_{90}Z,\ hZ,\ dZ\}$.

   Note: Using, for example, hZ instead of vZ (they're the same coset) is simply a matter of choice.

   (a) Compute the coset product $r_{90}Z \cdot hZ$ by multiplying each element of $r_{90}Z$ by those of hZ. You may *not* use the coset multiplication shortcut (i.e., Theorem 21.8).

   (b) Verify that the product in part (a) is indeed equal to $(r_{90} \cdot h)Z$.

   (This exercise and Exercise #9 below are referenced in Sections 22.3 and 23.1.)

9. Repeat Exercise #8, but now verify the following:

   (a) $hZ \cdot r_{90}Z = (h \cdot r_{90})Z$.

   (b) $\varepsilon Z \cdot dZ = (\varepsilon \cdot d)Z$.

   (c) $dZ \cdot hZ = (d \cdot h)Z$.

   Note: The coset multiplication shortcut does hold in D_4/Z; i.e., $aZ \cdot bZ = (ab)Z$ for all $aZ, bZ \in D_4/Z$, even though D_4 is *not* commutative (so Theorem 21.8 doesn't apply). We'll soon see why this is true.

10. Consider again the subgroup $Z = \{\varepsilon, r_{180}\}$ of D_4.

    (a) Complete the group table below:

$\cdot$	εZ	$r_{90}Z$	hZ	dZ
εZ				
$r_{90}Z$				
hZ				
dZ				

    (b) Use the table to verify that D_4/Z is a group under coset multiplication.

    (c) Is D_4/Z commutative or non-commutative?

    (d) Find the order of each $aZ \in D_4/Z$. Is the group cyclic?

    (This exercise is referenced in Example 23.2.)

11. Let $H = \{\varepsilon, h\}$ be a subgroup of D_4. (This exercise is referenced in Exercise #18 below.)

    (a) Compute the cosets $r_{90}H$ and dH.

    (b) Compute the coset product $r_{90}H \cdot dH$ by multiplying each element of $r_{90}H$ by those of dH.

    (c) Does the coset multiplication shortcut work here? What's going on?!

12. Consider the group D_4 and its subgroup $H = \{\varepsilon, r_{180}, d, d'\}$.

    Note: We have $H = C(d)$, i.e., the *centralizer* of d in D_4. Thus, H is indeed a subgroup.

    (a) Compute the coset product $r_{90}H \cdot dH$ by multiplying each element of $r_{90}H$ by those of dH.

    (b) Verify that the product in part (a) is indeed equal to $(r_{90} \cdot d)H$.

13. Repeat Exercise #12, but now verify the following:

    (a) $dH \cdot r_{90}H = (d \cdot r_{90})H$.

    (b) $\varepsilon H \cdot dH = (\varepsilon \cdot d)H$.

    (c) $hH \cdot r_{270}H = (h \cdot r_{270})H$.

    Note: Indeed, the coset multiplication shortcut holds in this setting as well. We'll see why soon.

14. Consider again the subgroup $H = \{\varepsilon, r_{180}, d, d'\}$ of D_4.

    (a) Create the group table for D_4/H and verify that it's a group under coset multiplication.

    (b) Is D_4/H commutative or non-commutative?

15. Let $H = \{0, 4, 8\}$ be a subgroup of the (additive) group $\mathbb{Z}_{12}$.

    (a) Find the distinct left cosets of H.

    (b) Compute the coset sum $(2 + H) + (3 + H)$ by adding each element of $2 + H$ to those of $3 + H$.

    (c) Verify that the sum in part (b) is indeed equal to $(2 + 3) + H$.

16. Repeat Exercise #15, but now verify the following:
    (a) $(3 + H) + (2 + H) = (3 + 2) + H$.
    (b) $(0 + H) + (3 + H) = (0 + 3) + H$.
    (c) $(2 + H) + (2 + H) = (2 + 2) + H$.

17. Consider again the subgroup $H = \{0, 4, 8\}$ of $\mathbb{Z}_{12}$.
    (a) Create the group table for $\mathbb{Z}_{12}/H$ and verify that it's a group under coset addition.
    (b) Find the order of each $a + H \in \mathbb{Z}_{12}/H$. Is the group cyclic?

18. We've seen examples where the coset multiplication shortcut fails.
    (a) With subgroup $H = \{\varepsilon, v\}$ of D_4, we saw that $r_{90}H \cdot d'H \neq (r_{90} \cdot d')H$. (See Example 21.10.)
    (b) With subgroup $H = \{\varepsilon, h\}$ of D_4, we saw that $r_{90}H \cdot dH \neq (r_{90} \cdot d)H$. (See Exercise #11.)

    In each of those cases, verify that the set inclusion $(ab)H \subseteq aH \cdot bH$ still holds.

19. Let G be a group (not necessarily commutative), H a subgroup of G, and $a, b \in G$. Define the coset product by $aH \cdot bH = \{\alpha \cdot \beta \mid \alpha \in aH,\ \beta \in bH\}$. Prove that $(ab)H \subseteq aH \cdot bH$. (This exercise is referenced in the proof of Theorem 23.5. It is also the statement of Theorem 24.2.)

20. Let G be a *commutative* group, H a subgroup of G, and $a, b \in G$. Define the coset product by $aH \cdot bH = \{\alpha \cdot \beta \mid \alpha \in aH,\ \beta \in bH\}$. Prove that $aH \cdot bH \subseteq (ab)H$. (This exercise is referenced in Example 24.3.)

    Note: Combined with the result of Exercise #19, this shows that $aH \cdot bH = (ab)H$ when G is commutative, hence completing the proof of Theorem 21.8.

21. Consider the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$, and let $K = \ker \delta$ be the kernel of δ; i.e., $K = \{\alpha \in G(\mathbb{Z}_{10}) \mid \delta(\alpha) = 1\}$.
    (a) Let $\alpha = \left[\begin{smallmatrix} 7 & 2 \\ 5 & 3 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$. Verify that $\alpha \in K$.
    (b) Let $g = \left[\begin{smallmatrix} 2 & 1 \\ 5 & 4 \end{smallmatrix}\right] \in G(\mathbb{Z}_{10})$ so that $g \cdot \alpha$ is in the left coset gK. Find $\beta \in K$ for which $g \cdot \alpha = \beta \cdot g$.
    (c) Explain why $g \cdot \alpha$ is also contained in the right coset Kg.

22. Let $\theta : G \to H$ be a group homomorphism, and let $K = \ker \theta = \{a \in G \mid \theta(a) = \varepsilon_H\}$. Show that $gK = Kg$ for all $g \in G$. (This exercise is referenced in Example 24.13.)

    Hint: Be careful! $\theta(gk) = \theta(kg)$ does *not* necessarily imply that $gk = kg$.

22

Quotient Group Examples

In Chapter 21, we learned how to multiply (or add) a pair of cosets, as well as the coset multiplication shortcut; namely $aH \cdot bH = (ab)H$. This allowed us to form a new type of group, called a *quotient group*, whose elements are cosets. In this chapter, we'll formalize the notion of the quotient group. In particular, let G be a group, H a subgroup of G, and G/H (read "G mod H") the set of distinct cosets of H. We will show that if the coset multiplication shortcut holds in G/H, then G/H satisfies the group properties and thus is a group under coset multiplication. We will begin writing some proofs about G/H, with more proofs to come in the next chapter (which is aptly named "Quotient Group Proofs").

22.1 Quotient group U_{13}/H revisited

We begin by reviewing the main example from Chapter 21. Let $H = \{1, 3, 9\}$ be a subgroup of U_{13}. The set of distinct cosets of H is $U_{13}/H = \{1H, 2H, 4H, 7H\}$. Moreover, the set U_{13}/H (read "U_{13} mod H") turned out to be a group under coset multiplication. To multiply a pair of cosets such as $2H = \{2, 5, 6\}$ and $7H = \{7, 8, 11\}$, we multiply every element of $2H$ by every element of $7H$, as shown:

$$\begin{aligned} 2H \cdot 7H &= \{2, 5, 6\} \cdot \{7, 8, 11\} \\ &= \{2 \cdot 7, 2 \cdot 8, 2 \cdot 11, 5 \cdot 7, 5 \cdot 8, 5 \cdot 11, 6 \cdot 7, 6 \cdot 8, 6 \cdot 11\} \\ &= \{1, 3, 9, 9, 1, 3, 3, 9, 1\} \\ &= \{1, 3, 9\} \\ &= 1H \end{aligned}$$

Therefore, $2H \cdot 7H = 1H$.

Rather than performing this tedious computation, we found the *coset multiplication shortcut*; namely, $aH \cdot bH = (ab)H$. In Theorem 21.8, we proved that this shortcut holds in G/H when G is commutative (such as when $G = U_{13}$). Thus, $2H \cdot 7H = (2 \cdot 7)H = 14H = 1H$ as before.

In Section 21.1, we verified that U_{13}/H satisfies the group properties under coset multiplication. In particular, $1H$ is its multiplicative identity element, where $1H \cdot aH = aH$ and $aH \cdot 1H = aH$ for all $aH \in U_{13}/H$. Thus, $2H \cdot 7H = 1H$ and $7H \cdot 2H = 1H$ imply that $2H$ and $7H$ are multiplicative inverses of each other. In symbols, we write $7H = (2H)^{-1}$; i.e., $7H$ is the multiplicative inverse of $2H$. Likewise, we write $2H = (7H)^{-1}$, which says that $2H$ is the multiplicative inverse of $7H$.

22.2 Quotient group U_{37}/H

Consider the multiplicative group $U_{37} = \{1, 2, 3, \ldots, 35, 36\}$ and its subgroup $H = \{1, 10, 26\}$. You will verify in an exercise at the end of the chapter that H is equal to $\langle 10 \rangle = \{10^k \mid k \in \mathbb{Z}\}$, i.e., the cyclic subgroup generated by 10. Thus, H is indeed a subgroup of U_{37}. Since U_{37} and H contain 36 and 3 elements, respectively, there are $\frac{36}{3} = 12$ distinct cosets of H. And as U_{37} is commutative, Theorem 21.8 ensures that the coset multiplication shortcut holds in U_{37}/H.

Note that in all of the examples in this section, the subgroup H of U_{37} refers to $H = \{1, 10, 26\}$.

Example 22.1. Let $4H, 11H \in U_{37}/H$, where $4H = \{4, 3, 30\}$ and $11H = \{11, 36, 27\}$. While the shortcut does hold in U_{37}/H, we'll find $4H \cdot 11H$ as a set product, to remind us of the underlying computation. To ease the calculation somewhat, we will write some numbers in the cosets $4H$ and $11H$ using negative values modulo 37; i.e., $4H = \{4, 3, -7\}$ and $11H = \{11, -1, -10\}$. Thus, we have

$$\begin{aligned}
4H \cdot 11H &= \{4, 3, -7\} \cdot \{11, -1, -10\} \\
&= \{4 \cdot 11,\ 4 \cdot (-1),\ 4 \cdot (-10),\ 3 \cdot 11,\ 3 \cdot (-1),\ 3 \cdot (-10),\ -7 \cdot 11,\ -7 \cdot (-1),\ -7 \cdot (-10)\} \\
&= \{44, -4, -40, 33, -3, -30, -77, 7, 70\} \\
&= \{7, 33, 34, 33, 34, 7, 34, 7, 33\} \\
&= \{7, 33, 34\} \\
&= 7H.
\end{aligned}$$

Therefore, $4H \cdot 11H = 7H$. Using the shortcut, we have $4H \cdot 11H = (4 \cdot 11)H = 44H = 7H$, which matches the result obtained by a more tedious calculation.

Example 22.2. We compute the set product once more, in order to find $1H \cdot 11H$. To ease the calculation somewhat, we'll write $1H = \{1, 10, -11\}$ and $11H = \{11, -1, -10\}$.

$$\begin{aligned}
1H \cdot 11H &= \{1, 10, -11\} \cdot \{11, -1, -10\} \\
&= \{1 \cdot 11,\ 1 \cdot (-1),\ 1 \cdot (-10),\ 10 \cdot 11,\ 10 \cdot (-1),\ 10 \cdot (-10),\ -11 \cdot 11,\ -11 \cdot (-1),\ -11 \cdot (-10)\} \\
&= \{11, -1, -10, 110, -10, -100, -121, 11, 110\} \\
&= \{11, 36, 27, 36, 27, 11, 27, 11, 36\} \\
&= \{11, 27, 36\} \\
&= 11H.
\end{aligned}$$

Therefore, $1H \cdot 11H = 11H$. The shortcut confirms that $1H \cdot 11H = (1 \cdot 11)H = 11H$.

As Example 22.2 suggests, $1H$ is the multiplicative identity element of U_{37}/H. Next, let's consider multiplicative inverses in U_{37}/H and how they relate to multiplicative inverses in U_{37}.

Example 22.3. The coset multiplication shortcut implies $2H \cdot 19H = (2 \cdot 19)H = 1H$, since $2 \cdot 19 = 1$ in U_{37}. Thus, $2H \cdot 19H = 1H$. Symbolically, we write the following:

- In U_{37}, $2 \cdot 19 = 1$ implies $2^{-1} = 19$; i.e., the multiplicative inverse of 2 is 19.
- In U_{37}/H, $2H \cdot 19H = 1H$ implies $(2H)^{-1} = 19H$; i.e., the multiplicative inverse of $2H$ is $19H$.

Combining these two, we obtain $(2H)^{-1} = 19H = 2^{-1}H$, so that $(2H)^{-1} = 2^{-1}H$.

We use the coset multiplication shortcut to "verify" the group properties for U_{37}/H. Here, we write "verify" (in quotes), since exhibiting a few examples as we do below is not enough to justify these group properties. For a complete justification, see Section 22.3 below.

(1) U_{37}/H is closed under coset multiplication. In Example 22.1 above, we saw that

$$4H \cdot 11H = (4 \cdot 11)H = 44H = 7H$$

so that the product of $4H$ and $11H$ is another element of U_{37}/H.

(2) Coset multiplication is associative. This was proved in Theorem 21.6.

(3) U_{37}/H contains the identity element $1H$. In Example 22.2, we saw that $1H \cdot 11H = (1 \cdot 11)H = 11H$.

(4) Every element in U_{37}/H has an inverse. In Example 22.3, we found that $2H \cdot 19H = (2 \cdot 19)H = 1H$, and so the multiplicative inverse of $2H$ is $19H$; i.e., $(2H)^{-1} = 19H$.

22.3 Quotient group G/H (generalization)

Let G be a group, and let H be a subgroup of G. Suppose that G/H (the set of distinct cosets of H) satisfies the coset multiplication shortcut; i.e., $aH \cdot bH = (ab)H$ for all aH, $bH \in G/H$. Then we'll show that G/H is a group under coset multiplication.

(1) G/H **is closed.** Let $aH, bH \in G/H$, where $a, b \in G$. Then $aH \cdot bH = (ab)H$. Since G is closed, we have $ab \in G$. Thus $(ab)H$ is the coset of H that is generated by the element $ab \in G$. Therefore, $(ab)H \in G/H$, which implies that $aH \cdot bH \in G/H$. Hence, G/H is closed.

(2) **Coset multiplication is associative.** This was proved in Theorem 21.6.

(3) G/H **contains an identity.** Consider $\varepsilon H \in G/H$, where ε is the identity of G. We have

$$\varepsilon H \cdot aH = (\varepsilon a)H = aH \quad \text{and} \quad aH \cdot \varepsilon H = (a\varepsilon)H = aH$$

for all $aH \in G/H$. Thus, εH is a multiplicative identity of G/H.

(4) G/H **contains inverses of its elements.** Let $aH \in G/H$, where $a \in G$. Since G is a group, there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = \varepsilon$ and $a^{-1} \cdot a = \varepsilon$. Thus $a^{-1}H \in G/H$, and

$$aH \cdot a^{-1}H = (a \cdot a^{-1})H = \varepsilon H \quad \text{and} \quad a^{-1}H \cdot aH = (a^{-1}a)H = \varepsilon H.$$

Thus, the multiplicative inverse of aH is $a^{-1}H$. Symbolically, we write $(aH)^{-1} = a^{-1}H$.

Proof know-how. An element of G/H has the form aH where $a \in G$ is the coset representative. Often, a proof about G/H involves working with these coset representatives (which are elements of G) and relies on what we know about the group G. For instance, when showing that G/H is closed, we used the closure of G to conclude that $ab \in G$, and hence $(ab)H \in G/H$.

We have just proved the following theorem.

Theorem 22.4. *Let G be a group, and let H be a subgroup of G. If G/H satisfies the coset multiplication shortcut, then G/H is a group under coset multiplication.*

Definition 22.5 (Quotient group)**.** The group G/H in Theorem 22.4 is called a *quotient group*.

Example 22.6. Let $H = \{1, 10, 26\}$ be a subgroup of U_{37}. Then U_{37}/H is a quotient group containing $\frac{36}{3} = 12$ elements. The key here is that we treat each coset aH as an *element* of U_{37}/H.

Example 22.7. Let G be a group, and let H be its subgroup. Assume G/H satisfies the coset multiplication shortcut. Given $aH \in G/H$ where $a \in G$, we'll compute $(aH)^n$ for integer exponents n.

First, consider a positive value of n, say $n = 3$. The shortcut implies

$$(aH)^3 = aH \cdot aH \cdot aH = (a \cdot a \cdot a)H = a^3H,$$

so that $(aH)^3 = a^3H$.

In any group, we define an element raised to the 0^{th} power to be the identity element. In G/H, this implies $(aH)^0 = \varepsilon H$, since εH is the multiplicative identity of G/H. In G, we have $a^0 = \varepsilon$. Thus, $(aH)^0 = \varepsilon H = a^0H$, so that $(aH)^0 = a^0H$.

With a negative exponent, say $n = -3$, we have

$$\begin{aligned}(aH)^{-3} &= ((aH)^{-1})^3 \\ &= (aH)^{-1} \cdot (aH)^{-1} \cdot (aH)^{-1} \\ &= a^{-1}H \cdot a^{-1}H \cdot a^{-1}H \\ &= (a^{-1})^3H \\ &= a^{-3}H.\end{aligned}$$

Therefore, $(aH)^{-3} = a^{-3}H$.

Example 22.7 suggests the following theorem, whose proof is left for you as an exercise.

Theorem 22.8. *Let G be a group, and let H be its subgroup. Assume G/H satisfies the coset multiplication shortcut. Given $aH \in G/H$, we have $(aH)^n = a^nH$ for all integer exponents n.*

Here is an example of an additive group. Recall that Theorem 21.11 (i.e., the coset addition shortcut), which states $(a+H)+(b+H) = (a+b)+H$, holds for any additive group G and its subgroup H, since additive groups are always commutative.

Example 22.9. Consider the subgroup $H = \{0, 4, 8\}$ of the additive group $\mathbb{Z}_{12}$. In Example 19.5, we found the following distinct cosets of H:

- $0 + H = 4 + H = 8 + H = \{0, 4, 8\}$ (original subgroup).
- $1 + H = 5 + H = 9 + H = \{1, 5, 9\}$.
- $2 + H = 6 + H = 10 + H = \{2, 6, 10\}$.
- $3 + H = 7 + H = 11 + H = \{3, 7, 11\}$.

Thus, $\mathbb{Z}_{12}/H = \{0 + H, 1 + H, 2 + H, 3 + H\}$. For instance, $(2 + H) + (3 + H) = (2 + 3) + H = 1 + H$, since $5 + H = 1 + H$. Hence, the coset sum $(2 + H) + (3 + H)$ is contained in $\mathbb{Z}_{12}/H$.

The additive identity element of $\mathbb{Z}_{12}/H$ is $0 + H$. For all $a + H \in \mathbb{Z}_{12}/H$, we have

$$(0 + H) + (a + H) = (0 + a) + H = a + H$$

and

$$(a + H) + (0 + H) = (a + 0) + H = a + H.$$

For additive inverses in $\mathbb{Z}_{12}/H$, consider the following. We have $(3+H)+(9+H) = (3 + 9) + H = 0 + H$, since $3 + 9 = 0$ in $\mathbb{Z}_{12}$. Symbolically, we write the following:

- In $\mathbb{Z}_{12}$, $3 + 9 = 0$ implies $-3 = 9$; i.e., the additive inverse of 3 is 9.
- In $\mathbb{Z}_{12}/H$, $(3 + H) + (9 + H) = 0 + H$ says $-(3 + H) = 9 + H$; the additive inverse of $3 + H$ is $9 + H$.

Combining these two, we obtain $-(3+H) = 9+H = -3+H$, so that $-(3+H) = -3+H$.

Example 22.9 above can be generalized as follows, whose proof is left for you as an exercise.

Theorem 22.10. *Let G be an additive group, and let H be a subgroup of G. Then G/H satisfies the coset addition shortcut and is a group under coset addition.*

We end the chapter with the following observation. Theorem 22.4 says:

$$\boxed{\text{The coset multiplication shortcut holds in } G/H} \implies \boxed{G/H \text{ is a quotient group}}.$$

We might ask, "When does the shortcut hold?" A possible answer is, "When G is commutative," as proved in Theorem 21.8. But in Chapter 21, Exercises #8 and #9, we saw that the shortcut also holds in D_4/Z, where $Z = \{\varepsilon, r_{180}\}$, even though D_4 is *non-commutative*. So the story is a bit more complicated!

Exercises

1. Consider the subgroup $H = \{0, 4, 8\}$ of the additive group $\mathbb{Z}_{12}$. For each $a \in \mathbb{Z}_{12}$, find and compare the orders of $a \in \mathbb{Z}_{12}$ and $a + H \in \mathbb{Z}_{12}/H$. What conjecture do you have?
2. Consider the subgroup $H = \{1, 3, 9\}$ of the multiplicative group U_{13}. For each $a \in U_{13}$, find and compare the orders of $a \in U_{13}$ and $aH \in U_{13}/H$. What conjecture do you have?

3. Consider the subgroup $H = \{1, 10, 26\}$ of U_{37}. Verify that H is equal to $\langle 10 \rangle = \{10^k \mid k \in \mathbb{Z}\}$, i.e., the cyclic subgroup generated by 10. (This computation is referred to in Section 22.2.)

4. Consider the subgroup $H = \{1, 10, 26\}$ of U_{37}.
   (a) Find and compare the orders of $6 \in U_{37}$ and $6H \in U_{37}/H$.
   (b) Repeat part (a) with $34 \in U_{37}$ and $34H \in U_{37}/H$. It might help to write 34 as -3 modulo 37.
   (c) Repeat part (a) with $4 \in U_{37}$ and $4H \in U_{37}/H$.
   (d) What conjecture do you have?

5. Consider the subgroup $H = \{1, 10, 26\}$ of U_{37}. Let $a \in U_{37}$ with $\text{ord}(a) = 18$. Show that $(aH)^{18} = 1H$. What does this say about the order of aH in U_{37}/H? Explain.

6. Our friends are working on Exercise #4:

   Elizabeth: Phew! I just found that $\text{ord}(4) = 18$ in U_{37}.

   Anita: Great! So $\text{ord}(4H)$ must be 18 as well, since $(4H)^{18} = 4^{18}H = 1H$.

   Elizabeth: But do we know if 18 is the *smallest* positive exponent for $4H$?

   Anita: Sure. If n is less than 18, then $(4H)^n = 4^nH$ can't equal $1H$, because $4^n \neq 1$.

   How would you respond to Anita?

7. **Prove:** Consider the subgroup $H = \{1, 10, 26\}$ of U_{37}. Let $aH \in U_{37}/H$ where $a \in U_{37}$. Then $\text{ord}(aH)$ in U_{37}/H is a divisor of $\text{ord}(a)$ in U_{37}.

8. Consider the subgroup $H = \{1, 10, 26\}$ of U_{37}.
   (a) Find $(15H)^{-1}$, i.e., the multiplicative inverse of $15H$ in U_{37}/H.
   (b) Find $(28H)^{-1}$.
   (c) Find $(3H)^{-1}$.

9. Consider the subgroup $H = \{1, 7\}$ of U_{16}.
   (a) Quick! How many distinct left (or right) cosets of H are there? Explain how you know.
   (b) Find the quotient group U_{16}/H.
   (c) Create the group table for U_{16}/H and verify that it's a group under coset multiplication.
   (d) Find the order of each $aH \in U_{16}/H$. Is the group cyclic?

10. Consider the subgroup $H = \{1, 9\}$ of U_{16}.
    (a) Find the quotient group U_{16}/H and determine if it's cyclic.
    (b) Compare your work in part (a) with Exercise #9. Are you surprised by the results?
    (c) U_{16} has another subgroup K with two elements. Find it and determine if U_{16}/K is cyclic.

11. Consider the additive group $\mathbb{Q}$ of rational numbers and its subgroup $\mathbb{Z}$.
    (a) Describe the elements of $\mathbb{Q}$ that are contained in the coset $\frac{1}{5} + \mathbb{Z}$.
    (b) Find all $\alpha \in \mathbb{Q}$ such that $\alpha + \mathbb{Z} = \frac{1}{5} + \mathbb{Z}$.
    (c) Find 10 distinct cosets in $\mathbb{Q}/\mathbb{Z}$.
    (d) Explain why $\mathbb{Q}/\mathbb{Z}$ contains infinitely many cosets.

12. (a) Find the order of $\frac{2}{5} + \mathbb{Z}$ in $\mathbb{Q}/\mathbb{Z}$.
    (b) Find the order of $\frac{6}{11} + \mathbb{Z}$ in $\mathbb{Q}/\mathbb{Z}$.
    (c) Find the order of $-\frac{3}{4} + \mathbb{Z}$ in $\mathbb{Q}/\mathbb{Z}$.
    (d) Explain why every element of $\mathbb{Q}/\mathbb{Z}$ has finite order.

13. Consider the additive group $\mathbb{R}$ of real numbers and its subgroup $\mathbb{Z}$. Does every element of $\mathbb{R}/\mathbb{Z}$ have finite order? Explain why or why not.

14. Let G be a (multiplicative) group, and let H be its subgroup. Suppose G/H satisfies the coset multiplication shortcut. Consider a function $\theta : G \to G/H$ where $\theta(a) = aH$ for all $a \in G$. Prove that θ is a homomorphism.

    Remark. Recall from Section 17.1 that homomorphisms provide a *unifying language* to talk about familiar algebraic properties (e.g., exponent rules, distributive law, etc.). Here's another such instance. This time, we described the coset multiplication shortcut using the language of homomorphisms.

15. Let θ be the homomorphism from Exercise #14; i.e., $\theta : G \to G/H$ where $\theta(a) = aH$ for all $a \in G$.
    (a) Explain why θ is onto.
    (b) Is θ necessarily one-to-one? If so, prove it. If not, provide a counterexample.
    (c) **Prove:** θ is one-to-one if and only if $H = \{\varepsilon\}$.

16. Let G be a cyclic group, and let H be a subgroup of G. Explain why G/H satisfies the coset multiplication shortcut, which in turn implies that G/H is a quotient group.

17. Recall from Example 13.4 that U_{13} is cyclic with generator 2; i.e., $U_{13} = \langle 2 \rangle$. With the subgroup $H = \{1, 3, 9\}$ of U_{13}, verify that U_{13}/H is cyclic with generator $2H$. (This exercise and Exercises #18 and #19 below are referenced in Chapter 23, Exercise #9.)

18. Observe that $\mathbb{Z}_{12}$ is cyclic with generator 1; i.e., $\mathbb{Z}_{12} = \langle 1 \rangle$. With the subgroup $H = \{0, 4, 8\}$ of $\mathbb{Z}_{12}$, verify that $\mathbb{Z}_{12}/H$ is cyclic with generator $1 + H$.

19. It turns out that U_{37} is cyclic with generator 2; i.e., $U_{37} = \langle 2 \rangle$. With the subgroup $H = \{1, 10, 26\}$ of U_{37}, is the quotient group U_{37}/H cyclic? Explain your reasoning.

20. In Example 22.7, we used the interpretation $(aH)^{-3} = ((aH)^{-1})^3$ to show that $(aH)^{-3} = a^{-3}H$. This time, use the interpretation $(aH)^{-3} = ((aH)^3)^{-1}$ to obtain the same result.

21. Prove Theorem 22.8.

    Hint: When n is negative, write it as $n = -(-n)$ where $-n$ is positive.

22. Prove Theorem 22.10.

23

Quotient Group Proofs

Much of this chapter, including the exercises, focus on proofs involving quotient groups. These proofs typically involve two groups: (1) group G whose elements have the form $a \in G$ and (2) quotient group G/H whose elements are *cosets* of the form $aH \in G/H$. (Sometimes, the subgroup H will play a role as well.) Carefully navigating between the two groups is the key to these proofs. Proof know-hows will provide in-depth analyses of the proofs, highlighting techniques and tips that can be applied to other proofs.

At the end of the chapter, we will examine how the quotient group G/H may be viewed as a simplified (or *collapsed*) version of the group G. The simplification process removes some "clutter" from G, while still maintaining its essential properties. This theme will be revisited in Chapter 25, when we study the *First Isomorphism Theorem*.

23.1 Sample quotient group proofs

The theorems and proofs in this section will involve multiplicative groups. Let G be a group, and let H be a subgroup of G. The coset multiplication shortcut says:

$$aH \cdot bH = (ab)H \text{ for all } aH, bH \in G/H.$$

Unless specified otherwise, assume that the shortcut holds in G/H so that G/H is a quotient group.

Theorem 23.1. *Let G be a group, and let H be a subgroup of G. If G is commutative, then G/H is commutative.*

PROOF. Let $aH, bH \in G/H$, where $a, b \in G$. We must show that $aH \cdot bH = bH \cdot aH$. Since G is commutative, we have $ab = ba$ in G. Therefore,

$$aH \cdot bH = (ab)H = (ba)H = bH \cdot aH.$$

Thus, $aH \cdot bH = bH \cdot aH$ as desired. ■

Proof know-how. A recurring theme of this chapter is that *a quotient group is a group*. To prove that a group is commutative, we (arbitrarily) choose two elements from the

group, say α and β, and show that $\alpha\beta = \beta\alpha$. This same approach applies to G/H as well, since G/H is a group after all. Thus, we begin with the elements $aH, bH \in G/H$ and show that $aH \cdot bH = bH \cdot aH$.

A proof about G/H, whose elements are cosets, often involves working with coset representatives, which are elements of G. The key to the above proof is the equality $ab = ba$ in G (as G is commutative), which implies the coset equality $(ab)H = (ba)H$ in G/H. This leads to our desired goal of $aH \cdot bH = bH \cdot aH$. (Compare with the proof of the closure of G/H, shown in Section 22.3.)

Example 23.2. The converse of Theorem 23.1 is: If G/H is commutative, then G is commutative. This is false. For a counterexample, consider D_4 and its center $Z = \{\varepsilon, r_{180}\}$. In Chapter 21, Exercise #10, we saw that D_4/Z is a commutative group with 4 elements. However, D_4 is non-commutative.

Example 23.3. Consider the subgroup $H = \{1, 3, 9\}$ of U_{13}. We first find the order of 4:

$$4^1 = 4, \;\; 4^2 = 3, \;\; 4^3 = 12, \;\; 4^4 = 9, \;\; 4^5 = 10, \;\; 4^6 = 1,$$

so that $\text{ord}(4) = 6$ in U_{13}. Next, we find the order of $4H$:

$$(4H)^1 = 4^1H = 4H,$$
$$(4H)^2 = 4^2H = 3H = 1H.$$

Therefore, $\text{ord}(4H) = 2$ in U_{13}/H. The key to the above computation is $3H = 1H$. Even though 3 is not the identity element of U_{13}, the coset that it generates, namely $3H$, is the identity element of U_{13}/H. Indeed, Theorem 19.14 tells us that $3H$ and $9H$ are both equal to $1H$, since $3, 9 \in H$.

Since $4^6 = 1$ in U_{13}, we have $(4H)^6 = 4^6H = 1H$ in U_{13}/H. But this does *not* necessarily imply that $\text{ord}(4H) = 6$. Instead, $(4H)^6 = 1H$ implies that the order of $4H$ is a divisor of 6 (Theorem 12.18). We saw above that $\text{ord}(4H) = 2$, which is indeed a divisor of 6.

The above example motivates the following theorem.

Theorem 23.4. *Let $a \in G$ with finite order. Then* $\text{ord}(aH)$ *in G/H is a divisor of* $\text{ord}(a)$ *in G.*

PROOF. Let $n = \text{ord}(a)$ so that $a^n = \varepsilon$. Then $(aH)^n = a^nH = \varepsilon H$. Since $(aH)^n = \varepsilon H$, Theorem 12.18 implies that $\text{ord}(aH)$ is a divisor of n. ■

Proof know-how. We again highlight the notion that *a quotient group is a group*. Thus, any theorem that we've proven about a group applies to the quotient group G/H as well. In particular, Theorem 12.18 says that if a group element α raised to the n^{th} power equals the identity element, then $\text{ord}(\alpha)$ is a divisor of n. Applying this to G/H, we showed that $(aH)^n$ equals εH, which is the identity element of G/H. Thus, we concluded that $\text{ord}(aH)$ is a divisor of n.

While this proof is about the element $aH \in G/H$, we again worked with the coset representative $a \in G$ and used the knowledge that $a^n = \varepsilon$ in G. This, in turn, led to the conclusion that $(aH)^n = \varepsilon H$ in G/H.

The next theorem isn't necessarily about quotient groups, but it foreshadows the work that we'll do in the next chapter when we (finally!) address the question: "When does the coset multiplication shortcut hold?" Let G be a group, and let H be its subgroup. Theorem 21.8 states that if G is commutative, then the coset multiplication shortcut holds in G/H. But in Chapter 21, Exercises #8 and #9, we saw that the shortcut also holds in D_4/Z, where $Z = \{\varepsilon, r_{180}\}$, even though D_4 is *non-commutative.* Here is a generalization.

Theorem 23.5. *Let G be a group, $Z = \{z \in G \mid zg = gz \text{ for all } g \in G\}$ the center of G, and $a, b \in G$. Define the coset product by $aZ \cdot bZ = \{\alpha \cdot \beta \mid \alpha \in aZ,\ \beta \in bZ\}$. Then $aZ \cdot bZ = (ab)Z$.*

PROOF. To prove the set equality $aZ \cdot bZ = (ab)Z$, we must show $aZ \cdot bZ \subseteq (ab)Z$ and $(ab)Z \subseteq aZ \cdot bZ$. Note that $(ab)Z \subseteq aZ \cdot bZ$ follows immediately from Chapter 21, Exercise #19. Thus, it suffices to show that $aZ \cdot bZ \subseteq (ab)Z$. Let $\alpha \cdot \beta \in aZ \cdot bZ$, where $\alpha \in aZ$ and $\beta \in bZ$. Thus, $\alpha = ax$ and $\beta = by$ where $x, y \in Z$. Since $x \in Z$, we have $xb = bx$, as x commutes with every element of G. Therefore,

$$\alpha \cdot \beta = (ax)(by) = a(xb)y = a(bx)y = (ab)(xy),$$

so that $\alpha \cdot \beta = (ab)(xy)$. Moreover, $xy \in Z$ by the closure of Z and so $(ab)(xy) \in (ab)Z$. Thus $\alpha \cdot \beta \in (ab)Z$, which implies $aZ \cdot bZ \subseteq (ab)Z$, as desired. ∎

In the next theorem and its proof, we work with *three* groups: the group G, its subgroup H, and the quotient group G/H. Note that H is a subset of G, so that the elements of H (including its identity) are also elements of G, and the two groups share the same operation. The quotient group G/H, however, is a separate group (though closely related), with different elements and a different operation.

Theorem 23.6. *Let G be a group, and let H be a subgroup of G. If every element of H and G/H has finite order, then every element of G also has finite order.*

PROOF. Let $g \in G$. We will show that $\text{ord}(g)$ is finite by finding a positive integer k such that $g^k = \varepsilon$. Consider the element $gH \in G/H$, which has finite order. With $n = \text{ord}(gH)$, we have $(gH)^n = \varepsilon H$, which implies $g^n H = H$. Hence, $g^n \in H$ by Theorem 19.14. Noting that $g^n \in H$ has finite order, let $m = \text{ord}(g^n)$. Thus, $(g^n)^m = \varepsilon$ so that $g^{nm} = \varepsilon$. Therefore, $\text{ord}(g)$ is finite. ∎

Proof know-how. Here, a proof about an element $g \in G$ involves working with the coset it generates, i.e., $gH \in G/H$, and using what we know about the quotient group G/H. It is indeed a role reversal between G and G/H when compared to the proofs of Theorems 23.1 and 23.4.

Also note that the identity element of G/H is εH (which equals the subgroup H), while the identity element of H (and G) is ε. Observe that $\varepsilon \in G$ generates the coset $\varepsilon H \in G/H$.

To close this section, we will present one more proof. Note that by Theorem 23.5, the coset multiplication shortcut holds in G/Z, and thus G/Z is a quotient group.

Theorem 23.7. *Let G be a group, and let Z be the center of G. If G/Z is cyclic, then G is commutative.*

PROOF. Assume G/Z is cyclic with a generator gZ. Let $a, b \in G$. (We must show that $ab = ba$.) Then the elements $aZ, bZ \in G/Z$ are integer powers of gZ; i.e., $aZ = (gZ)^i = g^iZ$ and $bZ = (gZ)^j = g^jZ$ for some integers i and j. Since $a \in aZ$ and $aZ = g^iZ$, we have $a \in g^iZ$. Thus $a = g^ix$ for some $x \in Z$. Similarly, $b = g^jy$ for some $y \in Z$. Therefore,

$$ab = \text{(details are left for you as an exercise)} = ba,$$

so that $ab = ba$. Hence, G is commutative. ■

Proof know-how. Similar to the proof of Theorem 23.6, this is a proof about elements $a, b \in G$ that involves working with the cosets they generate, i.e., $aZ, bZ \in G/Z$. Then we use what we know about the quotient group G/Z, namely that it's cyclic. Once we establish the coset equality $aZ = g^iZ$, we use the fact that $a \in aZ$ (Theorem 19.12) to deduce that $a = g^ix$ for some $x \in Z$.

23.2 Collapsing G into G/H

Example 23.8. Consider the following composition table for D_4:

$\circ$	ε	r_{90}	r_{180}	r_{270}	h	v	d	d'
ε	ε	r_{90}	r_{180}	r_{270}	h	v	d	d'
r_{90}	r_{90}	r_{180}	r_{270}	ε	d'	d	h	v
r_{180}	r_{180}	r_{270}	ε	r_{90}	v	h	d'	d
r_{270}	r_{270}	ε	r_{90}	r_{180}	d	d'	v	h
h	h	d	v	d'	ε	r_{180}	r_{90}	r_{270}
v	v	d'	h	d	r_{180}	ε	r_{270}	r_{90}
d	d	v	d'	h	r_{270}	r_{90}	ε	r_{180}
d'	d'	h	d	v	r_{90}	r_{270}	r_{180}	ε

The thick lines divide the table into four quadrants:

- The top left quadrant shows that a rotation composed with a rotation is a rotation.
- The top right quadrant shows that a rotation composed with a reflection (in that order) is a reflection.
- The bottom left quadrant shows that a reflection composed with a rotation (in that order) is a reflection.
- The bottom right quadrant shows that a reflection composed with a reflection is a rotation.

In fact, we observed that this is true in all groups D_n and not just in D_4. (See Chapter 5, Exercise #13.)

Let's consider the subgroup $H = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$ of D_4. Then the distinct cosets of H are as follows:

- $\varepsilon H = r_{90}H = r_{180}H = r_{270}H = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$.
- $hH = vH = dH = d'H = \{h, v, d, d'\}$.

Hence, $D_4/H = \{\varepsilon H,\ hH\}$ where εH is the coset containing all the rotations of D_4, while hH contains all the reflections. Moreover, the table for D_4 illustrates how the coset multiplication shortcut holds in D_4/H. For instance, the top right quadrant contains the elements in the set product $\varepsilon H \cdot hH$, where we multiply every element of εH by every element of hH. This quadrant contains all reflections, i.e., the elements of the coset hH, and thus $\varepsilon H \cdot hH = hH$. This equality respects the shortcut, which says $\varepsilon H \cdot hH = (\varepsilon \cdot h)H = hH$.

Arguing similarly with the other quadrants, the composition table for D_4 can be collapsed into the following table for the quotient group D_4/H:

$\cdot$	εH	hH
εH	εH	hH
hH	hH	εH

In this table for D_4/H, the four rotations of D_4 have been collapsed into a single element $\varepsilon H \in D_4/H$. Similarly, the four reflections have been collapsed into a single element $hH \in D_4/H$. The quotient group D_4/H does not keep track of the individual rotations and reflections from D_4. But the table for D_4/H still captures the fact that a rotation composed with a rotation is a rotation, a reflection composed with a reflection is a rotation, and a rotation composed with a reflection (in either order) is a reflection.

Remark. For the subgroup $H = \{\varepsilon,\ r_{90},\ r_{180},\ r_{270}\}$ of D_4, we have $[D_4 : H] = 2$, since there are two distinct left (and right) cosets of H. We used the group table of D_4 to observe that D_4/H satisfies the coset multiplication shortcut. In the next chapter, we'll prove something more general; namely: If $[G : H] = 2$, then G/H satisfies the shortcut so that G/H is a quotient group.

Example 23.9. Let $H = \{\varepsilon,\ r_{180},\ h,\ \upsilon\}$ be a subgroup of D_4. We have $H = C(h) = \{\sigma \in D_4 \mid \sigma \circ h = h \circ \sigma\}$, i.e., the *centralizer* of h containing the elements of D_4 that commute with h. (See Section 5.3.) Hence, H is indeed a subgroup. The distinct cosets of H are as follows:

- $\varepsilon H = r_{180}H = hH = \upsilon H = \{\varepsilon,\ r_{180},\ h,\ \upsilon\}$.
- $r_{90}H = r_{270}H = d'H = dH = \{r_{90},\ r_{270},\ d',\ d\}$.

Thus, $D_4/H = \{\varepsilon H,\ r_{90}H\}$ where the coset εH contains the elements of D_4 that commute with h, while $r_{90}H$ contains those that don't. Here is the group table for D_4, with its elements rearranged to highlight the elements in the cosets εH and $r_{90}H$:

$\circ$	ε	r_{180}	h	υ	r_{90}	r_{270}	d'	d
ε	ε	r_{180}	h	υ	r_{90}	r_{270}	d'	d
r_{180}	r_{180}	ε	υ	h	r_{270}	r_{90}	d	d'
h	h	υ	ε	r_{180}	d	d'	r_{270}	r_{90}
υ	υ	h	r_{180}	ε	d'	d	r_{90}	r_{270}
r_{90}	r_{90}	r_{270}	d'	d	r_{180}	ε	υ	h
r_{270}	r_{270}	r_{90}	d	d'	ε	r_{180}	h	υ
d'	d'	d	r_{90}	r_{270}	h	υ	ε	r_{180}
d	d	d'	r_{270}	r_{90}	υ	h	r_{180}	ε

Again, this table shows how the coset multiplication shortcut holds in D_4/H. For instance, the bottom left quadrant contains the elements in the set product $r_{90}H \cdot \varepsilon H$, where we multiply every element of $r_{90}H$ by every element of εH. This quadrant contains all the elements of the coset $r_{90}H$, and thus $r_{90}H \cdot \varepsilon H = r_{90}H$. This equality respects the shortcut, which says $r_{90}H \cdot \varepsilon H = (r_{90} \cdot \varepsilon)H = r_{90}H$.

Arguing similarly with the other quadrants, the composition table for D_4 can be collapsed into the following table for the quotient group D_4/H:

$\cdot$	εH	$r_{90}H$
εH	εH	$r_{90}H$
$r_{90}H$	$r_{90}H$	εH

In this table for D_4/H, the four elements of $H = C(h)$ are collapsed into a single element $\varepsilon H \in D_4/H$. Similarly, the four elements of D_4 that don't commute with h are collapsed into an element $r_{90}H \in D_4/H$. But the table for D_4/H still captures the following facts:

- An element of D_4 that commutes with h composed with another element that commutes with h is an element that commutes with h. (Also see Chapter 5, Exercise #15.) This is captured by $\varepsilon H \cdot \varepsilon H = \varepsilon H$.
- An element of D_4 that commutes with h composed with an element that doesn't (in either order) is an element that doesn't commute with h. This is captured by $\varepsilon H \cdot r_{90}H = r_{90}H \cdot \varepsilon H = r_{90}H$.
- An element of D_4 that doesn't commute with h composed with another element that doesn't is an element that commutes with h. This is captured by $r_{90}H \cdot r_{90}H = \varepsilon H$.

Exercises

Unless specified otherwise, assume that the coset multiplication shortcut holds in G/H.

1. Since *a quotient group is a group*, any property that we know about groups applies to G/H as well. Let's consider the "socks-shoes," for example. What goes into the empty boxes?
$$(aH \cdot bH)^{-1} = \square H \cdot \square H.$$
2. Consider the following statement:

   If $aH = bH$ in G/H, then $a = b$ in G.

   Is it true or false? If it's true, prove it. If it's false, give a counterexample.
3. Let $H = \{1, 3, 9\}$ be a subgroup of U_{13}, and consider $U_{13}/H = \{1H, 2H, 4H, 7H\}$.
   (a) Find all integers n for which $(2H)^n = 1H$ in U_{13}/H.
   (b) Find all integers n for which $2^n \in H$.
   (c) What conjecture do you have?
4. **Prove:** Let $gH \in G/H$ and $n \in \mathbb{Z}$. Then $(gH)^n = \varepsilon H$ if and only if $g^n \in H$.
5. Suppose $[G : H] = n$. Show that $g^n \in H$ for all $g \in G$.

   Recall: $[G : H]$ is the number of (left) cosets of H, which is also the size of G/H.

6. Find an additive group G, a subgroup H, and an element $a \in G$ such that

$$a + H \neq 0 + H, \text{ ord}(a + H) \text{ in } G/H \text{ is finite, and ord}(a) \text{ in } G \text{ is infinite.}$$

7. Same as Exercise #6, but with a multiplicative group G.

8. Let G be a group, and let H be a subgroup of G. Determine if each statement is true or false. If it's true, prove it. If it's false, give a counterexample.

   (a) If G and H are finite, then G/H is finite.

   (b) If G/H is finite, then G and H are finite.

   (c) If G is infinite and H is finite, then G/H is infinite.

9. Let $G = \langle g \rangle$ be a cyclic group, and let H be a subgroup of G. Prove that $G/H = \langle gH \rangle$, i.e., that G/H is cyclic with generator gH. (See Chapter 22, Exercises #17, #18, and #19. Also, this exercise is referenced in Example 25.7.)

   Hint: We have $\langle gH \rangle \subseteq G/H$ by definition. Thus, it suffices to show the other inclusion $G/H \subseteq \langle gH \rangle$.

10. Is the converse of Exercise #9 true? In other words, if G/H is cyclic, then must G also be cyclic? If so, prove it. If not, provide a counterexample.

11. Complete the proof of Theorem 23.7 by filling in the details to establish the equality $ab = ba$.

12. Let G be a non-commutative group with 343 elements. Let Z be the center of G and assume $Z \neq \{\varepsilon\}$. How many elements must Z contain? Explain your reasoning.

13. Let G be a group, and let Z be the center of G. Explain why $[G : Z]$ is *not* a prime number.

14. **Prove:** Let G be a finite group (so each $g \in G$ has finite order), and let H be a subgroup of G. If G/H has an element of order n, then G has an element of order n.

15. (a) Write the contrapositive of the implication in Theorem 23.7.

    (b) Explain how $G = D_4$ serves as an example of the contrapositive from part (a).

16. **Prove:** Let G be a commutative group, and let H be its subgroup. If every element $h \in H$ is a square in H (i.e., $h = k^2$ for some $k \in H$) and every element $aH \in G/H$ is a square in G/H (i.e., $aH = (bH)^2$ for some $bH \in G/H$), then every element of G is a square in G.

17. **Prove:** Let G be a group, and let H be its subgroup. Suppose every element of H and G/H has order 3^n where n is a non-negative integer. Then every element of G also has order 3^n where $n \geq 0$.

18. Let H be a subgroup of G with $[G : H] = 2$. Prove that $aH = Ha$ for all $a \in G$. (This is the statement of Theorem 24.10.)

19. Let G be a group, and let H be a subgroup with $[G : H] = 8$. If $g \in G$ has odd order (i.e., ord(g) is odd), then $g \in H$.

20. Let G be a group, and let H be a subgroup with $[G : H] = m$. If $g \in G$ has order n where $\gcd(m, n) = 1$, then $g \in H$. (Compare with Exercise #19.)

21. Let $H = \{\varepsilon, r_{180}, d, d'\}$ be a subgroup of D_4. Note that $H = C(d)$, i.e., the centralizer of d in D_4. Thus, H is indeed a subgroup. Proceed as in Examples 23.8 and 23.9 as follows:

    (a) Find the distinct cosets of H.

    (b) Create a group table for D_4, with its elements rearranged to highlight the elements in the distinct cosets that you found in part (a).

    (c) Explain how your table for D_4 shows that the coset multiplication shortcut holds in D_4/H.

    (d) Describe how the table for D_4 can be collapsed into the table for D_4/H. In particular, what feature of the original group D_4 is still captured by the table for D_4/H?

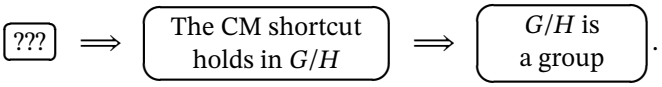
24

Normal Subgroups

Let G be a group, and let H be a subgroup of G. The coset multiplication shortcut says:

$$aH \cdot bH = (ab)H \text{ for } aH,\ bH \in G/H.$$

We proved (in Theorem 22.4) that if the shortcut holds, then G/H is a quotient group under coset multiplication. Thus, the following diagram illustrates what we know so far:



As we'll learn shortly, the missing piece ??? turns out to be: "H is a normal subgroup." The goal of this chapter is to understand what it means for a subgroup to be *normal* and how that relates to the coset multiplication shortcut. While most of our work with cosets has involved left cosets, the *right* cosets will play a prominent role when dealing with normal subgroups. We'll also revisit the notion of *conjugation*, which was first introduced in Section 12.3.

24.1 How does the shortcut fail and work?

We revisit earlier examples of coset multiplication to see how the shortcut fails and how it works.

Example 24.1 (Example 21.10 revisited)**.** Let $H = \{\varepsilon, \upsilon\}$ be a subgroup of D_4. We have $r_{90}H = \{r_{90}, d\}$ and $d'H = \{d', r_{270}\}$, and we compute the set product $r_{90}H \cdot d'H$ as shown:

$$\begin{aligned} r_{90}H \cdot d'H &= \{r_{90}, d\} \cdot \{d', r_{270}\} \\ &= \{r_{90} \cdot d',\ r_{90} \cdot r_{270},\ d \cdot d',\ d \cdot r_{270}\} \\ &= \{\upsilon, \varepsilon, r_{180}, h\}. \end{aligned}$$

Therefore, $r_{90}H \cdot d'H = \{\upsilon, \varepsilon, r_{180}, h\}$, which is *not* a coset of H as it has too many elements. Since $(r_{90} \cdot d')H = \upsilon H = \{\upsilon, \varepsilon\}$, we have $r_{90}H \cdot d'H \neq (r_{90} \cdot d')H$. The coset

multiplication shortcut fails! But all is not lost. Even though $r_{90}H \cdot d'H \neq (r_{90} \cdot d')H$, we notice that the elements of $(r_{90} \cdot d')H$, namely v and ε, are contained in $r_{90}H \cdot d'H = \{v, \varepsilon, r_{180}, h\}$. Hence, we do have a set inclusion $(r_{90} \cdot d')H \subseteq r_{90}H \cdot d'H$. In fact, this inclusion *always* holds, as shown in the next theorem.

The following theorem is left for you to prove in Chapter 21, Exercise #19.

Theorem 24.2. *Let G be a group, and let H be a subgroup of G. For $a, b \in G$, define the coset product by $aH \cdot bH = \{\alpha \cdot \beta \mid \alpha \in aH,\ \beta \in bH\}$. Then $(ab)H \subseteq aH \cdot bH$.*

Next, we analyze situations where coset multiplication does hold.

Example 24.3 (Chapter 21, Exercise #20 revisited)**.** Let G be a *commutative* group, H a subgroup of G, and $a, b \in G$. We will show that $aH \cdot bH = (ab)H$. Theorem 24.2 ensures $(ab)H \subseteq aH \cdot bH$. Let's see why the other inclusion $aH \cdot bH \subseteq (ab)H$ is also true. Suppose $\alpha \cdot \beta \in aH \cdot bH$, where $\alpha \in aH$ and $\beta \in bH$. Thus, $\alpha = ah$ and $\beta = bk$ for some $h, k \in H$. Since G is commutative, we have

$$\alpha \cdot \beta = (ah)(bk) = a(\boldsymbol{hb})k = a(\boldsymbol{bh})k = (ab)(hk) \in (ab)H.$$

Thus, $\alpha \cdot \beta \in (ab)H$ so that $aH \cdot bH \subseteq (ab)H$.

In Example 24.3 above, the key step that allowed us to conclude that $\alpha \cdot \beta \in (ab)H$ is the equality $bh = hb$, where $b \in G$ and $h \in H$. This requirement can be relaxed a bit, as shown in the next example.

Example 24.4 (Example 19.9 revisited)**.** Let $H = \{\varepsilon, r_{180}, h, v\}$ be a subgroup of D_4. We compute and compare the left coset dH and the right coset Hd.

- $dH = \{d \cdot \varepsilon,\ d \cdot r_{180},\ \boldsymbol{d} \cdot \boldsymbol{h},\ d \cdot v\} = \{d,\ d',\ \boldsymbol{r_{270}},\ r_{90}\}$.
- $Hd = \{\varepsilon \cdot d,\ r_{180} \cdot d,\ h \cdot d,\ \boldsymbol{v} \cdot \boldsymbol{d}\} = \{d,\ d',\ r_{90},\ \boldsymbol{r_{270}}\}$.

Thus we have a *coset equality* $dH = Hd$, because these sets contain the same four elements. This does not imply that we have an *element-by-element equality*, i.e., $dk = kd$ for all $k \in H$. What we *can* say, which we'll find useful next, is that an element such as $dh \in dH$ is also contained in the set Hd, since $dH = Hd$. Thus $dh \in Hd$, which implies that $dh = kd$ for some $k \in H$. In fact, we have $dh = vd$ where $v \in H$.

Example 24.5. Let $H = \{\varepsilon, r_{180}, h, v\}$ be a subgroup of D_4, and consider the cosets $r_{90}H$ and dH. We'll leave it to you to compute the coset product $r_{90}H \cdot dH$ and verify that it equals $(r_{90} \cdot d)H$. Here, we will analyze why the set inclusion $r_{90}H \cdot dH \subseteq (r_{90} \cdot d)H$ must hold. Consider $\alpha = r_{90}v \in r_{90}H$ and $\beta = dr_{180} \in dH$. We then have

$$\alpha\beta = (r_{90}v)(dr_{180}) = r_{90}(\boldsymbol{vd})r_{180} = r_{90}(\boldsymbol{dh})r_{180} = (r_{90}d)(hr_{180}) \in (r_{90} \cdot d)H,$$

where $hr_{180} \in H$ by the closure of H. The key step here is $dh = vd$, where $d \in D_4$ and $h, v \in H$. As we saw in Example 24.4 above, a relationship such as $dh = vd$ holds because $dH = Hd$; i.e., the left and right cosets of H generated by d are equal. Indeed, we saw in Chapter 19, Exercise #15 that $gH = Hg$ for *all* $g \in D_4$.

24.2 Normal subgroups: What and why

The examples from Section 24.1 motivate the following definition and the subsequent theorem.

Definition 24.6 (Normal subgroup)**.** Let H be a subgroup of a group G. Then H is called a *normal subgroup* of G if $gH = Hg$ for all $g \in G$.

Remark. We often say, "H is normal in G," to mean, "H is a normal subgroup of G."

Why should we care about normal subgroups? Here is the answer:

$$\boxed{H \text{ is a normal subgroup}} \implies \boxed{\text{The CM shortcut holds in } G/H} \implies \boxed{G/H \text{ is a group}}.$$

In other words, if H is a normal subgroup of G, then the coset multiplication shortcut holds in G/H, so that G/H is a (quotient) group under coset multiplication. Here is the theorem, whose proof resembles the calculations we did in Example 24.5.

Theorem 24.7. *Let G be a group, and let H be a subgroup of G. For $a, b \in G$, define the coset product by $aH \cdot bH = \{\alpha \cdot \beta \mid \alpha \in aH,\ \beta \in bH\}$. If H is normal in G, then $aH \cdot bH = (ab)H$.*

PROOF. Assume H is normal in G. We will show the set equality $aH \cdot bH = (ab)H$. First, we know from Theorem 24.2 that $(ab)H \subseteq aH \cdot bH$. (This is *always* true, whether or not H is a normal subgroup.)

Next, we will show that $aH \cdot bH \subseteq (ab)H$. Suppose $x \in aH \cdot bH$ so that $x = ah \cdot bk$ for some $h, k \in H$. Since H is a normal subgroup, we have $bH = Hb$. And since $hb \in Hb$, we have $hb \in bH$ so that $hb = bj$ for some $j \in H$. Therefore,

$$x = ah \cdot bk = a(\boldsymbol{hb})k = a(\boldsymbol{bj})k = ab \cdot jk$$

where $jk \in H$ by the closure of H. Thus, $x = ab \cdot jk \in (ab)H$ so that $aH \cdot bH \subseteq (ab)H$. Therefore, $aH \cdot bH = (ab)H$ as desired. ∎

Proof know-how. Remember that $bH = Hb$ does *not* mean $bh = hb$ for all $h \in H$. Here's what we say instead: Given $hb \in Hb$, we have $hb \in bH$ (since $bH = Hb$). Thus, we can write $hb = bj$ for some $j \in H$.

24.3 Examples of normal subgroups

Example 24.8 (G and $\{\varepsilon\}$ are normal subgroups)**.** Let G be a group. Then G is its own subgroup, and you'll show in an exercise at the end of the chapter that G is a *normal* subgroup of G. Moreover, the trivial subgroup $\{\varepsilon\}$ is normal in G. For all $g \in G$, we have $g\{\varepsilon\} = \{g\}$ and $\{\varepsilon\}g = \{g\}$, so that $g\{\varepsilon\} = \{\varepsilon\}g$.

Example 24.9 (Chapter 19, Exercise #16 revisited)**.** Let $K = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$ be a subgroup of D_4. We found that $aK = Ka$ for all $a \in D_4$; i.e., all left and right cosets of K are equal. More specifically,

$$aK = Ka = \begin{cases} K & \text{if } a \in K \text{ (i.e., } a = \text{rotation)}, \\ D_4 - K & \text{if } a \notin K \text{ (i.e., } a = \text{reflection)}. \end{cases}$$

Here, $D_4 - K$ is the set of elements in D_4 that are *not* in K. In other words, $D_4 - K = \{h, v, d, d'\}$, the subset of D_4 containing all reflections.

In Example 24.9 above, K and D_4 contain 4 and 8 elements, respectively. Therefore, $[D_4 : K] = 2$. It turns out that a subgroup of index 2 is always a normal subgroup. The following theorem is left for you to prove in Chapter 23, Exercise #18.

Theorem 24.10 (Index 2 subgroups are normal)**.** *Let G be a group, and let H be a subgroup of G with $[G : H] = 2$. Then H is normal in G.*

Example 24.11 (Commutative groups have normal subgroups)**.** Let G be a *commutative* group, and let H be a subgroup of G. Then for each $g \in G$, we have $gH = Hg$, because we have element-by-element equality; i.e., $gh = hg$ for all $h \in H$. Thus, H is normal in G.

Example 24.12 (Center is a normal subgroup)**.** Let G be a group, and let $Z = \{z \in G \mid zg = gz \text{ for all } g \in G\}$; i.e., Z is the *center* of G. (Example: If $G = D_4$, then $Z = \{\varepsilon, r_{180}\}$.) Then for each $g \in G$, we have $gZ = Zg$, because we have element-by-element equality; i.e., $gz = zg$ for all $z \in Z$. Thus, Z is normal in G. For this reason, we often think of the center Z as the "commutative part" of group G.

Example 24.13 (Kernel is a normal subgroup)**.** In Chapter 21, Exercise #22, we proved the following:

> Let $\theta : G \to H$ be a group homomorphism, and let $K = \ker\theta = \{a \in G \mid \theta(a) = \varepsilon_H\}$. Then $gK = Kg$ for all $g \in G$.

Thus, the kernel of a homomorphism is a normal subgroup of the domain G.

Example 24.14 (Non-example)**.** Consider again the subgroup $H = \{\varepsilon, v\}$ of D_4. We saw in Example 24.1 that the coset multiplication shortcut fails in D_4/H. Thus, by the contrapositive of Theorem 24.7, H is *not* a normal subgroup of D_4. And, in fact, we saw in Chapter 19, Exercise #14 that $r_{90}H = \{r_{90}, d\}$ and $Hr_{90} = \{r_{90}, d'\}$, so that $r_{90}H \neq Hr_{90}$.

24.4 Normal subgroup test

Directly showing that a subgroup is normal (i.e., by showing $gH = Hg$ for all $g \in G$) can be a tedious task. Fortunately, there's an "easier" way—in quotes, since nothing is easy in abstract algebra! The theorem below is often called the *normal subgroup test*.

Theorem 24.15 (Normal subgroup test)**.** *Let G be a group, and let H be a subgroup of G. Then H is normal in G if and only if $gHg^{-1} \subseteq H$ for all $g \in G$.*

Remark. Recall from Chapter 12, Exercise #23 that for a fixed $g \in G$, we define $gHg^{-1} = \{ghg^{-1} \mid h \in H\}$. Then gHg^{-1} is a subgroup of G and is called a *conjugate* of H.

PROOF. We must prove two implications:

- If H is normal in G, then $gHg^{-1} \subseteq H$ for all $g \in G$.
- If $gHg^{-1} \subseteq H$ for all $g \in G$, then H is normal in G.

First, assume H is normal in G. Let $g \in G$. We must show that $gHg^{-1} \subseteq H$. Let $x \in gHg^{-1}$ so that $x = ghg^{-1}$ for some $h \in H$. Since H is normal in G, we have $gH = Hg$. And as $gh \in gH$, we have $gh \in Hg$ so that $gh = kg$ for some $k \in H$. Then $x = (gh)g^{-1} = (kg)g^{-1} = k(gg^{-1}) = k\varepsilon = k \in H$. Hence, $x \in H$ so that $gHg^{-1} \subseteq H$.

Next, assume $gHg^{-1} \subseteq H$ for all $g \in G$. Let $a \in G$. We will show that $aH = Ha$ by showing $aH \subseteq Ha$ and $Ha \subseteq aH$. For $aH \subseteq Ha$, suppose $ah \in aH$ for some $h \in H$. Then $aha^{-1} \in aHa^{-1}$. But we know that $aHa^{-1} \subseteq H$, so that $aha^{-1} \in H$. Thus, $aha^{-1} = k$ for some $k \in H$. Right multiplication by a yields $ah = ka \in Ha$, so that $aH \subseteq Ha$. A similar argument, whose details we'll leave for you as an exercise at the end of the chapter, shows that $Ha \subseteq aH$. Therefore, $aH = Ha$ as desired. ∎

Proof know-how. In the first part of the proof, the coset equality $gH = Hg$ allows us to conclude that $gh = kg$ for some $k \in H$. This is similar to the proof of Theorem 24.7.

In the second part of the proof, we begin with an element $ah \in aH$, with an aim of showing $ah \in Ha$. We proceed to consider the element aha^{-1}, but how did we know to take that step? We used the "working backwards" technique again. Our goal was to show that $ah \in Ha$, i.e., to find some $k \in H$ such that $ah = ka$. We worked backwards and solved this equation for k by right-multiplying both sides by a^{-1}. We found $aha^{-1} = k$, which suggested that we work with the element aha^{-1} and show that it's in H. As usual, the "working backwards" process of solving for k is scratch work and must not be included in the proof itself.

Remark. The normal subgroup test is an "if and only if" statement, with two implications:

(1) If H is normal in G, then $gHg^{-1} \subseteq H$ for all $g \in G$.

(2) If $gHg^{-1} \subseteq H$ for all $g \in G$, then H is normal in G.

We typically use implication (2): To conclude that H is normal in G, we show that $gHg^{-1} \subseteq H$ for all $g \in G$. But implication (1) can be used as well: If we already *know* that H is a normal subgroup, then we can conclude and use the fact that $gHg^{-1} \subseteq H$ for all $g \in G$. The use of both implications will be demonstrated in the proofs below.

Let $G = G(\mathbb{Z}_{10})$, the multiplicative group of invertible 2×2 matrices with entries in $\mathbb{Z}_{10}$. Consider its subgroup $H = S(\mathbb{Z}_{10}) = \{\alpha \in G(\mathbb{Z}_{10}) \mid \det \alpha = 1\}$. Theorem 11.9 states that H is a subgroup of G. We will now show that H is normal in G using implication (2) of the normal subgroup test.

Theorem 24.16. *$H = S(\mathbb{Z}_{10})$ is normal in $G = G(\mathbb{Z}_{10})$.*

PROOF. Let $g \in G$. We will show that $gHg^{-1} \subseteq H$. Let $x \in gHg^{-1}$ so that $x = ghg^{-1}$ for some $h \in H$. Noting that $\det h = 1$, we have

$$\det x = \det(ghg^{-1}) = \det g \cdot \det h \cdot (\det g)^{-1} = \det g \cdot 1 \cdot (\det g)^{-1} = 1.$$

Thus, $\det x = 1$ so that $x \in H$. Hence, $gHg^{-1} \subseteq H$, which implies that H is normal in G. ∎

Proof know-how. In our proofs, we do *not* need to specify whether we used implication (1) or implication (2) of the normal subgroup test.

Let $\theta : G \to H$ be a group homomorphism, with $K = \ker\theta = \{a \in G \mid \theta(a) = \varepsilon_H\}$. We've already seen that the kernel K is a subgroup of the domain G. (See Theorem 18.6.) In fact, we've also shown that K is normal in G using the *definition* of a normal subgroup. (See Example 24.13.) In an exercise at the end of the chapter, you will reprove that K is normal in G, but this time using the normal subgroup test. It is an important result, and thus we will state it as a theorem here.

Theorem 24.17 (Kernel is a normal subgroup)**.** *Let $\theta : G \to H$ be a group homomorphism, with $K = \ker\theta = \{a \in G \mid \theta(a) = \varepsilon_H\}$. Then K is normal in G.*

Example 24.18. Consider the function $f : D_4 \to \mathbb{R}^*$ where

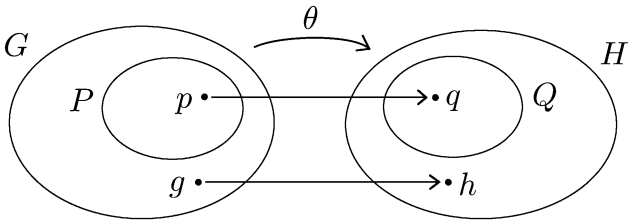
$$f(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is a rotation,} \\ -1 & \text{if } \sigma \text{ is a reflection.} \end{cases}$$

As shown in Chapter 17, Exercise #3, f is a homomorphism with kernel $K = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$, which gives yet another reason why this subgroup is normal in D_4.

Let $\theta : G \to H$ be a group homomorphism, and let Q be a subgroup of H. Define the set

$$P = \{a \in G \mid \theta(a) \in Q\}.$$

This set P is called the *preimage* of Q and contains the elements of the domain G that map to the elements of Q. In the figure below, we have $p \in P$, because $\theta(p) \in Q$; but $g \notin P$, since $\theta(g) \notin Q$. In Chapter 18, Exercise #17, you showed that P is a subgroup of G. Also note that if $Q = \{\varepsilon_H\}$, then P is the kernel of θ. Thus we may view the preimage P as a generalization of the kernel.



Below, we will prove that if Q is normal in H, then P is normal in G. The proof uses both implications (1) and (2) of the normal subgroup test. To prove that P is normal in G, we show $gPg^{-1} \subseteq P$, just as we did in the proof of Theorem 24.16. This relies on implication (2). But we also *know* that Q is normal in H so that implication (1) ensures $hQh^{-1} \subseteq Q$ for all $h \in H$. We will use this fact to show that $gPg^{-1} \subseteq P$.

Theorem 24.19. *Let $\theta : G \to H$ be a group homomorphism, and let Q be a subgroup of H. Define the set*

$$P = \{a \in G \mid \theta(a) \in Q\}.$$

If Q is normal in H, then P is normal in G.

PROOF. Assume Q is normal in H. Let $g \in G$. We will show that $gPg^{-1} \subseteq P$. Let $x \in gPg^{-1}$ so that $x = gpg^{-1}$ for some $p \in P$. We will show that $\theta(x) \in Q$, which will imply that $x \in P$.

Let $h = \theta(g) \in H$ and $q = \theta(p) \in Q$. Then, $\theta(x) = \theta(gpg^{-1}) = \theta(g) \cdot \theta(p) \cdot \theta(g)^{-1} = h \cdot q \cdot h^{-1}$, so that $\theta(x) = h \cdot q \cdot h^{-1}$. Since Q is normal in H, we have $hQh^{-1} \subseteq Q$. Thus $h \cdot q \cdot h^{-1} \in hQh^{-1}$ is also an element of Q. Hence $\theta(x) \in Q$ so that $x \in P$. Therefore $gPg^{-1} \subseteq P$ and P is normal in G, as desired. ■

Exercises

When working with the group D_4, refer to Appendix B for its group table.

1. Let $H = \{\varepsilon, d\}$ be a subgroup of D_4.

   (a) Give an example which shows that the coset multiplication shortcut does *not* hold in D_4/H.

   (b) Give an example which shows H is *not* normal in D_4.

2. Let $H = \{\varepsilon, d\}$, $K = \{\varepsilon, v\}$, $Z = \{\varepsilon, r_{180}\}$ be subgroups of D_4.

   Note: By Exercise #1(b) and Example 24.14, respectively, H and K are *not* normal in D_4.

   (a) Explain why Z is normal in D_4.

   (b) Compute the set product $HK = \{\alpha\beta \mid \alpha \in H,\ \beta \in K\}$. Is HK a subgroup of D_4? Explain.

   (c) Compute the set product ZK, defined similarly. Is ZK a subgroup of D_4? Explain.

3. Determine if each statement is true or false. If it's true, prove it. If it's false, give a counterexample.

   (a) Let H be a normal subgroup of G. If H and G/H are commutative, then G is commutative.

   (b) If H is a normal subgroup of G and K is a normal subgroup of H, then K is normal in G.

4. Let N be a normal subgroup of G and let H be a subgroup of G (but not necessarily normal in G). Define the set product $NH = \{nh \mid n \in N,\ h \in H\}$. Prove that NH is a subgroup of G.

5. Consider the subgroup $K = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$ of D_4. In this problem, we'll analyze why the set inclusion $dK \cdot vK \subseteq (dv)K$ must hold.

   Note: Your work from Chapter 19, Exercise #16 will come in handy here.

   (a) Find the element $\boxed{?} \in K$ such that $r_{90}v = v\boxed{?}$.

   (b) Let $\alpha\beta \in dK \cdot vK$, where $\alpha = dr_{90} \in dK$ and $\beta = vr_{180} \in vK$. Using only your result from part (a), explain why $\alpha\beta \in (dv)K$.

   (c) Explain why $dK \cdot vK \subseteq (dv)K$.

6. Consider again the subgroup $K = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$ of D_4. In this problem, we'll analyze why the set inclusion $hK \cdot d'K \subseteq (hd')K$ must hold.

   (a) Find the element $\boxed{?} \in K$ such that $r_{270}d' = d'\boxed{?}$.

   (b) Let $\alpha\beta \in hK \cdot d'K$, where $\alpha = hr_{270} \in hK$ and $\beta = d'r_{90} \in d'K$. Using only your result from part (a), explain why $\alpha\beta \in (hd')K$. Don't look at the D_4 table again!

   (c) Explain why $hK \cdot d'K \subseteq (hd')K$.

7. Let G be a group, and let $H = \{\varepsilon, h\}$ be a normal subgroup of G with two elements. Prove that H is contained in the center of G, i.e., that $H \subseteq Z$, where $Z = \{z \in G \mid zg = gz \text{ for all } g \in G\}$.

8. Complete the proof of Theorem 24.15 by showing the set inclusion $Ha \subseteq aH$.

   Hint: Use $gHg^{-1} \subseteq H$ with $g = a^{-1}$.

9. Prove Theorem 24.17. You must use the normal subgroup test.

10. Let $\theta : G \to H$ be an *onto* group homomorphism. Let K be a subgroup of G, and define

$$J = \{\theta(k) \mid k \in K\}.$$

    (a) **Prove:** If K is normal in G, then J is normal in H.

    (b) Where in your proof did you use the fact that θ is onto? Explain.

11. **Prove:** Let G be a group, H a *commutative* group, and $\theta : G \to H$ a group homomorphism. Suppose N is a subgroup of G with $\ker\theta \subseteq N$. Then N is normal in G.

12. Let $H = \{\varepsilon, r_{180}, h, v\}$ be a subgroup of D_4. Compute the coset product $r_{90}H \cdot dH$ and verify that it equals $(r_{90} \cdot d)H$. (See Example 24.5.)

13. Let G be a group. Explain why G is a normal subgroup of G. (See Example 24.8.)

14. (a) Let G be a group containing 10 elements. Suppose H is a subgroup of G that's *not* normal in G. How many elements does H contain?

    (b) Repeat part (a) but with G containing 14 elements.

    (c) Repeat part (a) but with G containing 22 elements.

    (d) Generalize your results from parts (a), (b), and (c).

15. **Prove:** Let H and K be normal subgroups of a group G. If $H \cap K = \{\varepsilon\}$, then $hk = kh$ for all $h \in H$ and $k \in K$.

16. **Prove:** Let G be a group, and let H be a normal subgroup of G. If H is cyclic, then every subgroup of H is normal in G.

17. **Prove:** Let G be a group, and let H be a subgroup of G with $[G : H] = 2$. If a, $b \notin H$, then $ab \in H$. (See Examples 23.8 and 23.9.)

18. Define $\varepsilon, \sigma, \gamma, \tau, \mu, \delta \in S_3$, respectively, by
$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Recall: You constructed the group table for S_3 in Chapter 6, Exercise #2.

(a) Verify that $H = \{\varepsilon, \sigma, \gamma\}$ is a subgroup of S_3 (or see Chapter 6, Exercise #4).

(b) Compute the product $\varepsilon \cdot \sigma \cdot \gamma \cdot \tau \cdot \mu \cdot \delta$, using any set of parentheses to pair the elements. Then verify that the product is *not* in H.

(c) Compute the product $\gamma \cdot \mu \cdot \varepsilon \cdot \delta \cdot \tau \cdot \sigma$, again using any set of parentheses. Then verify that the product is *not* in H.

(d) Multiply all the elements of S_3, using a different order from the products in parts (b) and (c). Then verify that the product is *not* in H.

(e) What conjecture do you have?

19. Let $H = \{\varepsilon, r_{90}, r_{180}, r_{270}\}$ be a subgroup of D_4.

(a) Compute the product $\varepsilon \cdot r_{90} \cdot r_{180} \cdot r_{270} \cdot h \cdot v \cdot d \cdot d'$, using any set of parentheses to pair the elements. Then verify that the product is in H.

(b) Multiply all the elements of D_4, taken in the order of your choice. Verify that the product is in H.

(c) Repeat part (b) a couple more times, each time multiplying the elements in different order.

(d) What conjecture do you have?

20. Repeat Exercise #19 using each of these subgroups of D_4. What conjecture do you have?

(a) $H = \{\varepsilon, r_{180}, h, v\}$.

(b) $H = \{\varepsilon, r_{180}, d, d'\}$.

21. Let G be a group with $2n$ elements, and let H be a subgroup of G with n elements. Let α be the product of all the elements of G, taken in any order. Prove the following:

- If n is odd, then $\alpha \notin H$.
- If n is even, then $\alpha \in H$.

22. **Prove:** Let G be a group. Consider the direct product $G \times G$ and its subset $\triangle G = \{(g, g) \mid g \in G\}$. Then G is commutative if and only if $\triangle G$ is normal in $G \times G$.

Note: In Chapter 11, Exercise #26, you showed that $\triangle G$ is a subgroup of $G \times G$.

25

First Isomorphism Theorem

In this final chapter on group theory, we will study the *First Isomorphism Theorem*, a culminating theorem that brings together many concepts that we've explored, including isomorphism, homomorphism, kernel, image, cosets, quotient group, just to name a few. Along the way, we'll (finally!) answer the question first posed in Section 18.3; namely: Why does a homomorphism partition the domain into *equal-sized* subsets? We will *not* prove the First Isomorphism Theorem in this textbook, although you're encouraged to try. Instead, we will make sense of the theorem using motivating examples.

To illustrate the power of the First Isomorphism Theorem, we will build homomorphisms (almost) from scratch. For instance, we'll find all homomorphisms $\theta : U_{13} \to U_{13}$ whose kernel contains four elements. Using the First Isomorphism Theorem, we can show that there are only two such homomorphisms.

25.1 Familiar homomorphism

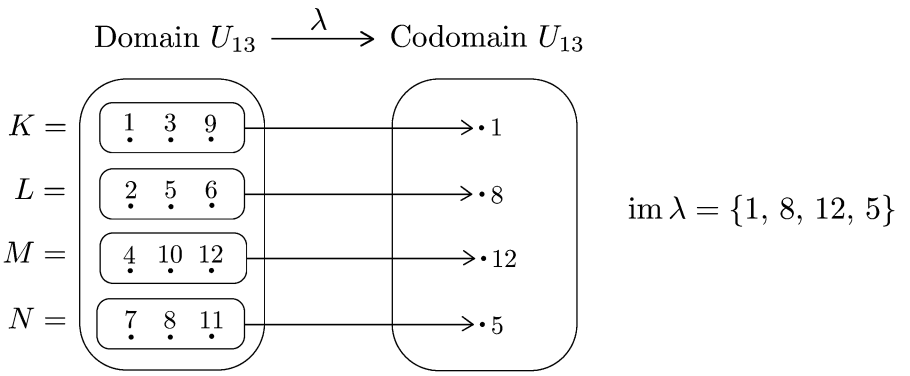
We revisit the homomorphism from Example 18.1. Consider $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. Here are the values of $\lambda(a)$ for each input a in the domain U_{13}:

$$\begin{array}{llll} \lambda(1) = 1, & \lambda(4) = 12, & \lambda(7) = 5, & \lambda(10) = 12, \\ \lambda(2) = 8, & \lambda(5) = 8, & \lambda(8) = 5, & \lambda(11) = 5, \\ \lambda(3) = 1, & \lambda(6) = 8, & \lambda(9) = 1, & \lambda(12) = 12. \end{array}$$

Recall that the homomorphism λ partitions the elements of the domain into subsets, according to the elements in the codomain to which they are mapped:

$$\begin{aligned} \ker\lambda = K &= \{a \in U_{13} \mid \lambda(a) = \ 1\} = \{1, \ 3, \ 9\}, \\ L &= \{a \in U_{13} \mid \lambda(a) = \ 8\} = \{2, \ 5, \ 6\}, \\ M &= \{a \in U_{13} \mid \lambda(a) = 12\} = \{4, 10, 12\}, \\ N &= \{a \in U_{13} \mid \lambda(a) = \ 5\} = \{7, \ 8, 11\}. \end{aligned}$$

For instance, the subset L contains elements of the domain U_{13} that map to 8 in the codomain U_{13}. We have $\lambda(2) = \lambda(5) = \lambda(6) = 8$, so that $L = \{2, 5, 6\}$. Note how the domain U_{13} has been divided into 4 equal-sized subsets, namely K, L, M, and N. The diagram below illustrates this scenario:



The fact that λ partitions the domain is not a big deal, since any function, homomorphism or not, does the same. (See Chapter 18, Exercise #10.) But the question remains: Why are these subsets *equal-sized*?

To answer this question, we consider the kernel of λ; namely $K = \ker\lambda = \{1, 3, 9\}$. As we computed in Example 19.1, the distinct cosets of K in U_{13} are as follows:

- $1K = 3K = 9K = \{1, 3, 9\}$ (original subgroup).
- $2K = 5K = 6K = \{2, 5, 6\}$.
- $4K = 10K = 12K = \{4, 10, 12\}$.
- $7K = 8K = 11K = \{7, 8, 11\}$.

These cosets are precisely the sets K, L, M, and N created by the homomorphism λ. Let's dig deeper and make sense of this observation. For example, consider $2 \in U_{13}$ (the domain), and note that $\lambda(2) = 8$. Then every element of the coset $2K = \{2, 5, 6\}$ also maps to 8. In other words, if $a \in 2K$, then $\lambda(a) = 8$. Furthermore, *only* the elements of $2K$ map to 8; i.e., if $a \notin 2K$, then $\lambda(a) \neq 8$. This is equivalent to its contrapositive: If $\lambda(a) = 8$, then $a \in 2K$.

Example 25.1. Again, consider the homomorphism $\lambda : U_{13} \to U_{13}$ where $\lambda(a) = a^3$ for all $a \in U_{13}$. Let $K = \ker\lambda = \{1, 3, 9\}$. We have $\lambda(7) = 5$. Then the elements of the coset $7K = \{7, 8, 11\}$ also map to 5, and those are the only elements in U_{13} (the domain) that map to 5.

Here is the generalization of the above observation.

Theorem 25.2. *Let $\theta : G \to H$ be a group homomorphism with $K = \ker\theta$. Let $g \in G$ such that $\theta(g) = h$ where $h \in H$. Given $a \in G$, $a \in gK$ if and only if $\theta(a) = h$.*

PROOF. We must prove two implications:

- If $a \in gK$, then $\theta(a) = h$.
- If $\theta(a) = h$, then $a \in gK$.

We will prove the first implication. The proof of the second implication is left for you as an exercise.

Assume $a \in gK$ so that $a = gk$ for some $k \in K$. Note that $\theta(k) = \varepsilon_H$, since k is in the kernel of θ. Thus, $\theta(a) = \theta(gk) = \theta(g) \cdot \theta(k) = h \cdot \varepsilon_H = h$, as desired. ■

Remark. Theorem 25.2 implies that the cosets of $K = \ker\theta$ partition the domain in the same way that the homomorphism θ does. This explains why the subsets created by the homomorphism are *equal-sized*, since we know that all cosets of K have the same size (see Theorem 19.15).

But there's more! Shown below are the group tables for the quotient group $U_{13}/K = \{1K, 2K, 4K, 7K\}$ and the image of the homomorphism $\operatorname{im}\lambda = \{\lambda(1), \lambda(2), \lambda(4), \lambda(7)\} = \{1, 8, 12, 5\}$.

Table for U_{13}/K:

$\cdot$	$1K$	$2K$	$4K$	$7K$
$1K$	$1K$	$2K$	$4K$	$7K$
$2K$	$2K$	$4K$	$7K$	$1K$
$4K$	$4K$	$7K$	$1K$	$2K$
$7K$	$7K$	$1K$	$2K$	$4K$

Table for $\operatorname{im}\lambda$:

$\cdot$	1	8	12	5
1	1	8	12	5
8	8	12	5	1
12	12	5	1	8
5	5	1	8	12

The two tables are essentially the same, with the following correspondences:

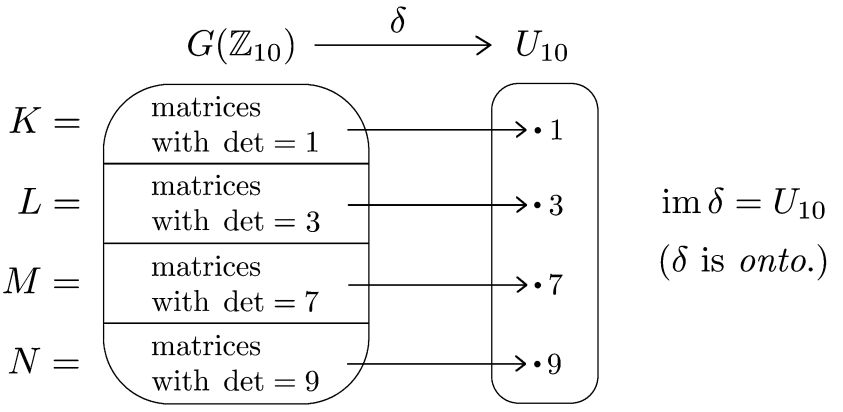
- $1K \in U_{13}/K$ corresponds to $\lambda(1) = 1 \in \operatorname{im}\lambda$.
- $2K \in U_{13}/K$ corresponds to $\lambda(2) = 8 \in \operatorname{im}\lambda$.
- $4K \in U_{13}/K$ corresponds to $\lambda(4) = 12 \in \operatorname{im}\lambda$.
- $7K \in U_{13}/K$ corresponds to $\lambda(7) = 5 \in \operatorname{im}\lambda$.

Hence, the two groups are isomorphic; i.e., $U_{13}/K \cong \operatorname{im}\lambda$, where $gK \in U_{13}/K$ corresponds to $\lambda(g) \in \operatorname{im}\lambda$.

25.2 Another homomorphism

Consider the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det\alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. Recall that $G(\mathbb{Z}_{10})$ refers to the multiplicative group of invertible 2×2 matrices with entries in $\mathbb{Z}_{10}$.

This homomorphism partitions the domain $G(\mathbb{Z}_{10})$ into four subsets, according to the elements in the codomain to which they are mapped (i.e., according to their determinants):



More precisely, the domain $G(\mathbb{Z}_{10})$ has been partitioned into these 4 subsets:

$$\begin{aligned} K &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 1\} \quad \longleftarrow \text{ this is } \ker\delta, \\ L &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 3\}, \\ M &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 7\}, \\ N &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 9\}. \end{aligned}$$

The kernel of this homomorphism is $K = \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 1\}$. Recall that this set is often called $S(\mathbb{Z}_{10})$. To find the distinct cosets of K, define the matrices

$$\sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_7 = \begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_9 = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix},$$

with determinants 1, 3, 7, and 9, respectively. (Note that $\sigma_1 = \varepsilon$.) We showed in Chapter 19, Exercise #21 that $\sigma_3 K = \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 3\}$; i.e., the coset $\sigma_3 K$ where $K = S(\mathbb{Z}_{10})$ and σ_3 is a fixed element of $G(\mathbb{Z}_{10})$ with $\det\sigma_3 = 3$ is equal to the set of all matrices in $G(\mathbb{Z}_{10})$ with determinant 3. We have similar equalities for the cosets $\sigma_7 K$ and $\sigma_9 K$. Therefore, the distinct cosets of K are as follows:

$$\begin{aligned} \sigma_1 K &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 1\}, \\ \sigma_3 K &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 3\}, \\ \sigma_7 K &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 7\}, \\ \sigma_9 K &= \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 9\}. \end{aligned}$$

Once again the cosets of $K = \ker\delta$ partition the domain $G(\mathbb{Z}_{10})$ in the same way that the homomorphism δ does. In particular, we verify that Theorem 25.2 is satisfied by the homomorphism δ:

- $\delta(\sigma_1) = 1$ and elements of the coset $\sigma_1 K$ are precisely those that map to 1.
- $\delta(\sigma_3) = 3$ and elements of the coset $\sigma_3 K$ are precisely those that map to 3.
- $\delta(\sigma_7) = 7$ and elements of the coset $\sigma_7 K$ are precisely those that map to 7.
- $\delta(\sigma_9) = 9$ and elements of the coset $\sigma_9 K$ are precisely those that map to 9.

In an exercise, you will create and compare the group tables for $G(\mathbb{Z}_{10})/K = \{\sigma_1 K,\ \sigma_3 K,\ \sigma_7 K,\ \sigma_9 K\}$ and $\operatorname{im}\delta = U_{10} = \{1, 3, 7, 9\}$. You'll see that the two tables are essentially the same, so that the two groups are isomorphic; i.e., $G(\mathbb{Z}_{10})/K \cong \operatorname{im}\delta$, where $gK \in G(\mathbb{Z}_{10})/K$ corresponds to $\delta(g) \in \operatorname{im}\delta$.

25.3 First Isomorphism Theorem

Finally, here is the First Isomorphism Theorem that generalizes the results of Sections 25.1 and 25.2.

Theorem 25.3 (First Isomorphism Theorem for groups)**.** *Let $\theta : G \to H$ be a group homomorphism with $K = \ker\theta$. Then $G/K \cong \operatorname{im}\theta$, where $gK \in G/K$ corresponds to $\theta(g) \in \operatorname{im}\theta$.*

Remark. There is the Second Isomorphism Theorem, and even third and fourth ones, depending on the algebraist you talk to. But we will cover only the First Isomorphism Theorem in this book.

25.4 Finding and building homomorphisms

Example 25.4 (Example 18.16 revisited)**.** Consider the group $U_{18} = \{1, 5, 7, 11, 13, 17\}$. Let $\theta : U_{18} \to U_{18}$ be a homomorphism with $K = \ker\theta = \{1, 7, 13\}$ and $\theta(5) = 17$. The distinct cosets of K are as follows:

- $1K = \ 7K = 13K = \{1,\ \ 7,\ 13\}$.
- $5K = 17K = 11K = \{5, 17, 11\}$.

Since the elements in the kernel map to the identity element of the codomain, $\theta(1) = \theta(7) = \theta(13) = 1$. Then Theorem 25.2 implies that the elements of the coset $5K$ map to $\theta(5) = 17$. Therefore we obtain $\theta(5) = \theta(17) = \theta(11) = 17$, and we've found the value of $\theta(a)$ for all $a \in U_{18}$. Note also the following:

- Since $17 = 5 \cdot 7$ modulo 18, we have $\theta(17) = \theta(5 \cdot 7) = \theta(5) \cdot \theta(7) = 17 \cdot 1 = 17$.
- Since $11 = 5 \cdot 13$ modulo 18, we have $\theta(11) = \theta(5 \cdot 13) = \theta(5) \cdot \theta(13) = 17 \cdot 1 = 17$.

Example 25.5. Consider the group $U_{22} = \{1, 3, 5, 7, 9, 13, 15, 17, 19, 21\}$. Let $\theta : U_{22} \to U_{22}$ be a homomorphism with $K = \ker\theta = \{1, 21\}$ and $\theta(3) = 15$. We will find the value of $\theta(a)$ for all $a \in U_{22}$. First, the distinct cosets of K are as follows:

- $1K = 21K = \{1, 21\}$.
- $3K = 19K = \{3, 19\}$.
- $5K = 17K = \{5, 17\}$.
- $7K = 15K = \{7, 15\}$.
- $9K = 13K = \{9, 13\}$.

Since $1, 21 \in \ker\theta$, we have $\theta(1) = \theta(21) = 1$. We're given that $\theta(3) = 15$. Since 3 and 19 are in the same coset of K, Theorem 25.2 ensures $\theta(3) = \theta(19) = 15$. Next, we use the fact that θ is operation preserving. As $9 = 3^2$ in U_{22}, we have $\theta(9) = \theta(3^2) = \theta(3)^2 = 15^2 = 5$. Thus $\theta(9) = 5$ and Theorem 25.2 implies that $\theta(9) = \theta(13) = 5$. In an exercise at the end of the chapter, you'll find the remaining values of $\theta(a)$.

Example 25.6. Let $\theta : U_{13} \to U_{13}$ be a homomorphism whose kernel $K = \ker\theta$ has 4 elements. Just this information is enough to completely determine θ.

Note that $\ker\theta$ is a subgroup of the domain U_{13} (Theorem 18.6). Recall that U_{13} is a cyclic group (with a generator 2). Then Theorem 14.15 implies that U_{13} has a unique subgroup of size 4. In Example 14.11, we found that subgroup to be $\{1, 5, 8, 12\}$. Thus, we have $K = \{1, 5, 8, 12\}$. The distinct cosets of K are as follows:

- $1K = \ 5K = 8K = 12K = \{1, 5, 8, 12\}$.
- $2K = 10K = 3K = 11K = \{2, 10, 3, 11\}$.
- $4K = \ 7K = 6K = \ 9K = \{4, 7, 6, 9\}$.

The elements of the kernel, namely 1, 5, 8, 12, all map to 1. Moreover, the First Isomorphism Theorem implies that U_{13}/K and $\operatorname{im}\theta$ are isomorphic. And since $U_{13}/K = \{1K, 2K, 4K\}$ contains 3 elements, $\operatorname{im}\theta$ must also contain 3 elements. But $\operatorname{im}\theta$ is a subgroup of the codomain U_{13} (Theorem 18.11), and U_{13} has a unique subgroup of size 3, namely $\{1, 3, 9\}$. Therefore, $\operatorname{im}\theta = \{1, 3, 9\}$.

In the isomorphism $U_{13}/K \cong \operatorname{im}\theta$, we have the correspondence $1K \leftrightarrow 1$ between the identity elements. And there are two ways in which $2K, 4K \in U_{13}/K$ correspond with $3, 9 \in \operatorname{im}\theta$.

- **Option 1:** $2K \leftrightarrow 3$ and $4K \leftrightarrow 9$. Therefore, we have

  $$\theta(1) = \theta(5) = \theta(8) = \theta(12) = 1,$$
  $$\theta(2) = \theta(10) = \theta(3) = \theta(11) = 3,$$
  $$\theta(4) = \theta(7) = \theta(6) = \theta(9) = 9.$$

- **Option 2:** $2K \leftrightarrow 9$ and $4K \leftrightarrow 3$. Therefore, we have

  $$\theta(1) = \theta(5) = \theta(8) = \theta(12) = 1,$$
  $$\theta(2) = \theta(10) = \theta(3) = \theta(11) = 9,$$
  $$\theta(4) = \theta(7) = \theta(6) = \theta(9) = 3.$$

Thus, there are two possible homomorphisms $\theta : U_{13} \to U_{13}$ whose kernel contains 4 elements. In an exercise, you'll find a formula for each of these options and verify that these functions are operation preserving, and hence honest-to-goodness homomorphisms.

Example 25.7. This time, let $\theta : U_{13} \to U_{13}$ be a homomorphism whose kernel $K = \ker\theta$ has 2 elements. Then $K = \{1, 12\}$, the unique subgroup of U_{13} of size 2. There

are 6 distinct cosets of K:

- $1K = 12K = \{1, 12\}$.
- $3K = 10K = \{3, 10\}$.
- $5K = 8K = \{5, 8\}$.
- $2K = 11K = \{2, 11\}$.
- $4K = 9K = \{4, 9\}$.
- $6K = 7K = \{6, 7\}$.

The First Isomorphism Theorem implies that $U_{13}/K \cong \operatorname{im}\theta$, and thus $\operatorname{im}\theta$ also contains 6 elements. Then $\operatorname{im}\theta = \{1, 3, 4, 9, 10, 12\}$, the unique subgroup of U_{13} of size 6. Let $\varphi : U_{13}/K \to \operatorname{im}\theta$ denote the isomorphism that's ensured by the First Isomorphism Theorem. To find the rule for φ (i.e., the correspondence between the elements of U_{13}/K and $\operatorname{im}\theta$), observe that U_{13} is cyclic and thus U_{13}/K is also cyclic. (See Chapter 23, Exercise #9.) In particular, noting that $U_{13} = \langle 2\rangle$, we obtain $U_{13}/K = \langle 2K\rangle$. Thus, to determine the isomorphism φ, it suffices to determine the value of $\varphi(2K)$.

Now, $\operatorname{ord}(2K) = 6$ in U_{13}/K and isomorphisms preserve order (Theorem 17.15). Hence, $\varphi(2K) \in \operatorname{im}\theta$ must also have order 6, and $\operatorname{im}\theta$ contains 2 elements of order 6, namely 4 and 10. Each of these choices determines the isomorphism φ, and hence the homomorphism θ.

- **Option 1:** $\varphi(2K) = 4$; i.e., $2K \leftrightarrow 4$. Since φ is operation preserving, we have the following:

  $\varphi(1K) = \varphi((2K)^0) = \varphi(2K)^0 = 4^0 = 1$. Thus, $\theta(1) = \theta(12) = 1$ (as expected).
  $\varphi(2K) = \varphi((2K)^1) = \varphi(2K)^1 = 4^1 = 4$. Thus, $\theta(2) = \theta(11) = 4$.
  $\varphi(3K) = \varphi((2K)^4) = \varphi(2K)^4 = 4^4 = 9$. Thus, $\theta(3) = \theta(10) = 9$.
  $\varphi(4K) = \varphi((2K)^2) = \varphi(2K)^2 = 4^2 = 3$. Thus, $\theta(4) = \theta(9) = 3$.
  $\varphi(5K) = \varphi((2K)^3) = \varphi(2K)^3 = 4^3 = 12$. Thus, $\theta(5) = \theta(8) = 12$.
  $\varphi(6K) = \varphi((2K)^5) = \varphi(2K)^5 = 4^5 = 10$. Thus, $\theta(6) = \theta(7) = 10$.

- **Option 2:** $\varphi(2K) = 10$; i.e., $2K \leftrightarrow 10$. In an exercise, you'll determine $\theta(a)$ for all $a \in U_{13}$.

Thus, there are two possible homomorphisms $\theta : U_{13} \to U_{13}$ whose kernel contains 2 elements. In an exercise, you'll find a formula for each of these options and verify that these functions are operation preserving, and hence honest-to-goodness homomorphisms.

Exercises

1. Consider the homomorphism $\delta : G(\mathbb{Z}_{10}) \to U_{10}$ where $\delta(\alpha) = \det\alpha$ for all $\alpha \in G(\mathbb{Z}_{10})$. In Section 25.2, we saw that $K = \ker\delta$ is equal to $S(\mathbb{Z}_{10}) = \{\alpha \in G(\mathbb{Z}_{10}) \mid \det\alpha = 1\}$.

   (a) Describe the distinct left cosets of K in $G(\mathbb{Z}_{10})$.

   (b) Find the image $\operatorname{im}\delta$.

   (c) The First Isomorphism Theorem says that $G(\mathbb{Z}_{10})/K \cong \operatorname{im}\delta$, where $gK \in G(\mathbb{Z}_{10})/K$ corresponds to $\delta(g) \in \operatorname{im}\delta$. Create group tables for $G(\mathbb{Z}_{10})/K$ and $\operatorname{im}\delta$ to verify this isomorphism.

2. Suppose $\varphi : U_{17} \to U_{17}$ is a homomorphism with kernel $K = \{1, 4, 13, 16\}$.
   (a) Verify that K is indeed a subgroup of the domain U_{17}.
   (b) Find all distinct cosets of K in U_{17}.
   (c) Suppose $\varphi(10) = 4$. Find all other $a \in U_{17}$ such that $\varphi(a) = 4$. How do you know that you've found *all* such elements?

3. Consider again the homomorphism $\varphi : U_{17} \to U_{17}$ with kernel $K = \{1, 4, 13, 16\}$. As in Exercise #2, suppose $\varphi(10) = 4$.
   (a) Find the value of $\varphi(a)$ for all $a \in U_{17}$.
   (b) Find the image $\operatorname{im}\varphi$.
   (c) Create the group tables for U_{17}/K and $\operatorname{im}\varphi$.
   (d) Verify that the two tables in part (c) are essentially the same. How are the groups U_{17}/K and $\operatorname{im}\varphi$ related?
   (e) Find a *formula* for the function φ and verify that it's operation preserving, and hence an honest-to-goodness homomorphism.

4. Consider the homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$.
   (a) Find the kernel $K = \ker\varphi$.
   (b) Find all distinct cosets of K in $\mathbb{Z}$.
   Note: Since $\mathbb{Z}$ is an additive group, these cosets have the form $a + K$ where $a \in \mathbb{Z}$.
   (c) Find the image $\operatorname{im}\varphi$.
   (d) Create and compare the group tables for $\mathbb{Z}/K$ and $\operatorname{im}\varphi$ to verify that they're isomorphic.

5. Use the First Isomorphism Theorem to prove that $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$. (This exercise is referenced in Chapter 32, Exercise #21.)

6. Consider the homomorphism $\gamma : \mathbb{Z}_{12} \to \mathbb{Z}_{18}$ where $\gamma(a) = 6a$ for all $a \in \mathbb{Z}_{12}$.
   (a) Find the kernel $K = \ker\gamma$.
   (b) Find all distinct cosets of K in $\mathbb{Z}_{12}$.
   Note: Since $\mathbb{Z}_{12}$ is an additive group, these cosets have the form $a + K$ where $a \in \mathbb{Z}_{12}$.
   (c) Find the image $\operatorname{im}\gamma$.
   (d) Create and compare the group tables for $\mathbb{Z}_{12}/K$ and $\operatorname{im}\gamma$ to verify that they're isomorphic.

7. Let $\theta : G \to H$ be a group homomorphism with $K = \ker\theta$. For $a, b \in G$, show that if $\theta(a) = \theta(b)$, then $aK = bK$. You may *not* use the First Isomorphism Theorem in your argument.

8. Complete the proof of Theorem 25.2 by proving its second implication.

9. Let $\theta : U_{20} \to U_{20}$ be a homomorphism with kernel $K = \{1, 11\}$ and $\theta(3) = 13$.
   (a) Find all distinct cosets of K in U_{20}.
   (b) Find the value of $\theta(a)$ for all $a \in U_{20}$.

(c) Create and compare the group tables for U_{20}/K and $\operatorname{im}\theta$ to verify that they're isomorphic.

10. Let $\theta : \mathbb{Z}_{40} \to \mathbb{Z}_{40}$ be a homomorphism with $K = \ker\theta = \langle 5\rangle$ and $\theta(17) = 24$.

(a) Find all elements in the domain that map to 24.

(b) Find the distinct cosets of K in $\mathbb{Z}_{40}$.
Note: Since $\mathbb{Z}_{40}$ is an additive group, these cosets have the form $a + K$ where $a \in \mathbb{Z}_{40}$.

(c) Find the value of $\theta(a)$ for all $a \in \mathbb{Z}_{40}$.

(d) Create and compare group tables for $\mathbb{Z}_{40}/K$ and $\operatorname{im}\theta$ to verify that they're isomorphic.

11. Let $\gamma : U_{31} \to U_{31}$ be a homomorphism with kernel $K = \{1, 5, 6, 25, 26, 30\}$.

(a) Find all distinct cosets of K in U_{31}.

(b) Suppose $\gamma(10) = 2$. Find all other $a \in U_{31}$ such that $\gamma(a) = 2$.

(c) Find the value of $\gamma(a)$ for all $a \in U_{31}$.

(d) Find the image $\operatorname{im}\gamma$.

(e) Create and compare the group tables for U_{31}/K and $\operatorname{im}\gamma$ to verify that they're isomorphic.

12. (a) Find the *ratio* of the number of elements in $G(\mathbb{Z}_{10})$ to the number of elements in $S(\mathbb{Z}_{10})$.

(b) **(Challenge)** Find the number of elements in $G(\mathbb{Z}_{10})$.

13. Complete Example 25.5 by finding the values of $\theta(a)$ for $a = 5, 7, 15, 17$.

14. In Example 25.6, we found two possible homomorphisms $\theta : U_{13} \to U_{13}$ whose kernel has 4 elements.

(a) Find a formula for each of these options.
Hint: Each formula has the form $\theta(a) = a^k$ where k is an integer.

(b) Using the formulas in part (a), verify that the functions that we found are operation preserving, and hence honest-to-goodness homomorphisms.

15. Complete Example 25.7 as follows:

(a) For Option 2, i.e., $\varphi(2K) = 10$, find the values of $\theta(a)$ for all $a \in U_{13}$.

(b) Find a formula for each of the two possible homomorphisms.

(c) Using the formulas in part (b), verify that the functions that we found are operation preserving, and hence honest-to-goodness homomorphisms.

16. (a) Find all possible homomorphisms $\theta : U_{13} \to U_{13}$ whose kernel has 3 elements.

(b) Find a formula for each of the possible homomorphisms you found in part (a).

(c) Using the formulas in part (b), verify that the functions that we found are operation preserving, and hence honest-to-goodness homomorphisms.

17. Repeat Exercise #16, but with a kernel containing the following:
    (a) 6 elements.
    (b) 12 elements.
    (c) 1 element. (Be careful! There is more than one possibility!)

18. In Example 25.6, we found two possible homomorphisms $\theta : U_{13} \to U_{13}$ whose kernel has 4 elements. For each of these, we found a formula (in Exercise #14), which we used to verify that the function we found is operation preserving. Explain why these functions are operation preserving *without* using the formulas. Verifying $\theta(ab) = \theta(a) \cdot \theta(b)$ for all $a, b \in U_{13}$ is not recommended!

19. Find all homomorphisms $\theta : U_{16} \to U_{16}$ with $\ker \theta = \{1, 7\}$.

20. (a) Find all homomorphisms $\theta : S_3 \to \mathbb{Z}_2$.
    (b) Find all homomorphisms $\theta : S_3 \to \mathbb{Z}_4$.
    (c) Find all homomorphisms $\theta : S_3 \to \mathbb{Z}_8$.
    (d) Find all homomorphisms $\theta : S_3 \to \mathbb{Z}_n$, where n is a power of 2.

21. **Prove:** Let $\theta : G \to H$ be a group homomorphism. Then the number of elements in $\operatorname{im} \theta$ is a divisor of the number of elements in G.

22. (a) Let $\theta : G \to H$ be a group homomorphism where G and H contain 8 and 15 elements, respectively. Explain why $\theta(g) = \varepsilon_H$ for all $g \in G$.
    (b) Repeat part (a), with G and H containing 10 and 27 elements, respectively.
    (c) Repeat part (a), with G and H containing 21 and 22 elements, respectively.
    (d) Repeat part (a), with G and H containing m and n elements, respectively, where $\gcd(m, n) = 1$.

23. **Prove:** Let K be a normal subgroup of a group G. Suppose $\varphi : G/K \to H$ is a group isomorphism. Then there exists an onto homomorphism $\theta : G \to H$ whose kernel is K.

Unit VI: Introduction to Rings

Notice that the set of integers $\mathbb{Z}$ comes with *two* operations, namely, addition and multiplication, and likewise for $\mathbb{Z}_{12}$ and $\mathbb{R}$. Thus, rather than focusing on one operation at a time (i.e., $\mathbb{Z}$, $\mathbb{Z}_{12}$, and $\mathbb{R}$ are all *additive* groups), we will consider both operations simultaneously. Hence, we arrive at the study of *rings*, i.e., sets like $\mathbb{Z}$, $\mathbb{Z}_{12}$, and $\mathbb{R}$ that have two operations satisfying familiar properties such as $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$. Chapter 27 introduces special kinds of rings called *integral domains* and *fields*, which come with desirable properties: We can cancel in integral domains and we can "divide" (by non-zero elements) in fields.

We also begin our study of *polynomial rings*, whose elements are familiar objects such as $f(x) = 2x^3 - 4x + 5$. An important underlying theme for the remainder of the book is the structural similarities between the ring of integers $\mathbb{Z}$ and the ring of polynomials whose coefficients are in a field. For example, some polynomials can be factored into a product of smaller polynomials; e.g., $x^2 - 6x + 8 = (x - 2)(x - 4)$. Likewise, some integers can be factored into a product of smaller integers; e.g., $15 = 3 \cdot 5$.

Here is a taste of what you'll be able to accomplish in this unit:

- Prove that $0 \cdot a = 0$ in a ring. By definition, 0 is the *additive identity*; i.e., $0 + a = a$ and $a + 0 = a$ for all elements a. There's nothing in that definition which says how 0 behaves under multiplication.
- Prove that every field is an integral domain, and show that not every integral domain is a field.
- Quickly determine whether or not $p(x) = x^3 + x + 1$ is factorable in $\mathbb{Z}_3[x]$. (**Answer:** It is!)

26

Introduction to Rings

For the remainder of this book, we will study an algebraic structure called a *ring*, although groups will continue to play a prominent role in our work. Unlike groups, where we consider one operation for each set (e.g., $\mathbb{Z}$ under addition, U_5 under multiplication), a ring has *two* operations. In fact, our most familiar example, namely, the set of integers $\mathbb{Z}$, is actually a ring.

This chapter introduces the notion of a ring (with lots of examples, of course). We'll see what makes the distributive law, i.e., $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$, so special. We'll *prove* familiar formulas such as $0 \cdot a = 0$ and $-1 \cdot a = -a$. Previous examples such as $\mathbb{Z}$, $\mathbb{Z}_{12}$, and $M(\mathbb{Z}_{10})$ will be revisited, but this time as rings.

26.1 Examples and definition

Example 26.1. In our study of groups, we viewed $\mathbb{Z}$ as an *additive* group. The set $\mathbb{Z}$ is not a *multiplicative* group, since most of its elements do not have multiplicative inverses. For instance, there is no integer n such that $5 \cdot n = 1$. Thus, 5 does not have a multiplicative inverse in $\mathbb{Z}$; i.e., 5^{-1} does not exist in $\mathbb{Z}$.

Although $\mathbb{Z}$ is not a multiplicative group, it is still *closed* under multiplication (i.e., $a \cdot b \in \mathbb{Z}$ for all $a, b \in \mathbb{Z}$), allowing us to add *and* multiply in $\mathbb{Z}$. The same can be said about $\mathbb{Z}_{12}$ and $\mathbb{R}$. Furthermore, there are properties of these operations that are common to $\mathbb{Z}$, $\mathbb{Z}_{12}$, and $\mathbb{R}$. For example, $a + b = b + a$ in all three sets. What other common properties can you find?

We acknowledge that $\mathbb{Z}$ has *two* operations: addition and multiplication, denoted $+$ and $\cdot$, respectively. Here are the essential properties of these operations. Observe that properties (1) through (5) involve addition, properties (6) through (8) involve multiplication, and property (9) involves both operations.

(1) $\mathbb{Z}$ is closed under addition: If $a, b \in \mathbb{Z}$, then $a + b \in \mathbb{Z}$.

(2) Associative law for addition: $(a + b) + c = a + (b + c)$ for all $a, b, c \in \mathbb{Z}$.

(3) There exists an additive identity $0 \in \mathbb{Z}$ such that $0 + a = a$ and $a + 0 = a$ for all $a \in \mathbb{Z}$.

(4) For $a \in \mathbb{Z}$, there exists an additive inverse $-a \in \mathbb{Z}$ such that $a + (-a) = 0$ and $(-a) + a = 0$.

(5) Commutative law for addition: $a + b = b + a$ for all $a, b \in \mathbb{Z}$.

(6) $\mathbb{Z}$ is closed under multiplication: If $a, b \in \mathbb{Z}$, then $a \cdot b \in \mathbb{Z}$.

(7) Associative law for multiplication: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in \mathbb{Z}$.

(8) There exists a multiplicative identity $1 \in \mathbb{Z}$ such that $1 \cdot a = a$ and $a \cdot 1 = a$ for all $a \in \mathbb{Z}$.

(9) Distributive law: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ for all $a, b, c \in \mathbb{Z}$.

Note the omission of the commutative law for multiplication (i.e., $a \cdot b = b \cdot a$ for all $a, b \in \mathbb{Z}$), even though this property is satisfied in $\mathbb{Z}$. This is intentional and we will explain why soon. Furthermore, property (9) (i.e., the distributive law) is the only property that connects addition and multiplication.

These properties define the algebraic structure of a *ring*.

Definition 26.2 (Ring)**.** A set R is called a *ring* if it has two operations (denoted $+$ and $\cdot$) satisfying properties (1) through (9) that are satisfied by $\mathbb{Z}$.

Remark. According to Definition 26.2 above, the set $2\mathbb{Z}$ of even integers is *not* a ring, because it does not contain the multiplicative identity element 1. (Note that $2\mathbb{Z}$ does satisfy all other ring properties.) However, many abstract algebra textbooks use a definition of a ring that does *not* require the multiplicative identity, which would make $2\mathbb{Z}$ a ring. And in this alternative definition, a ring that contains the multiplicative identity is called a *ring with unity* or a *ring with identity*.

In this textbook, a ring will contain the multiplicative identity by definition. We do so for two reasons. First, we wanted our definition of a ring to closely mimic what we observe in the ring $\mathbb{Z}$. Second, every relevant example of a ring that we examine contains the multiplicative identity.

We make some observations about the definition of a ring. Let R be a ring. Properties (1) through (5) say that R is a commutative group under addition. Thus, all the group properties that we've proved still apply to R, as long as we consider its addition operation. For instance, Theorem 8.9 states that a group has a unique identity element. Since R is an additive group, Theorem 8.9 implies that R has a unique *additive* identity element 0. It turns out that R also has a unique multiplicative identity element 1, which you'll prove in an exercise at the end of the chapter. But be careful: You can not claim that this follows from Theorem 8.9, since R is not a multiplicative group, as $0 \in R$ does not have a multiplicative inverse.

In a group, we used the symbol ε to denote the identity element. But a ring contains *two* identity elements. To differentiate the two, we use the symbols 0 and 1 to denote the additive and multiplicative identities, respectively. Given an element $a \in R$, we use $-a$ and a^{-1} to denote its additive and multiplicative inverses, respectively. Note that $-a$ always exists and a^{-1} sometimes exists, as shown in the example below.

Example 26.3. Consider the set $\mathbb{Z}_{10}$. We'll leave it to you as an exercise to verify that $\mathbb{Z}_{10}$ is indeed a ring under addition and multiplication modulo 10. We have $3 + 7 = 0$ in $\mathbb{Z}_{10}$ so that 3 and 7 are additive inverses of each other. We write $-3 = 7$ and $-7 = 3$. We also have $3 \cdot 7 = 1$ in $\mathbb{Z}_{10}$, so that 3 and 7 are multiplicative inverses of each other as well. We write $3^{-1} = 7$ and $7^{-1} = 3$. However, $5 \cdot x = 1$ is not possible in $\mathbb{Z}_{10}$, since $5 \cdot x = 0$ or 5 for all $x \in \mathbb{Z}_{10}$. Hence, 5^{-1} does not exist in $\mathbb{Z}_{10}$.

Example 26.4 (Examples of rings)**.** Here are some examples of rings, many of which we've seen before:

- $\mathbb{Z}$, the set of integers.
- $\mathbb{Q}$, the set of rational numbers.
- $\mathbb{R}$, the set of real numbers.
- $\mathbb{C}$, the set of complex numbers of the form $a + bi$ where $a, b \in \mathbb{R}$. (Here, $i = \sqrt{-1}$ so that $i^2 = -1$.)
- $\mathbb{Z}_n = \{0, 1, 2, 3, \ldots, n-1\}$ where addition and multiplication are done modulo n.

Example 26.5 (Polynomial rings)**.** Later, we will study these *polynomial rings* in depth.

- $\mathbb{R}[x]$ is the set of all polynomials with coefficients in $\mathbb{R}$.
- $\mathbb{Z}_p[x]$ is the set of all polynomials with coefficients in $\mathbb{Z}_p$. (Here, p is a prime number.)

For example, let $f(x) = 3x^4 + 2x^3 + 4$ and $g(x) = 4x^3 + 1$ be elements of $\mathbb{Z}_5[x]$. We then have

$$f(x) + g(x) = (3x^4 + 2x^3 + 4) + (4x^3 + 1) = 3x^4 + \mathbf{6}x^3 + \mathbf{5} = 3x^4 + \mathbf{1}x^3 + \mathbf{0},$$

so that $f(x) + g(x) = 3x^4 + x^3$. Observe how we reduce the coefficients in $\mathbb{Z}_5$; i.e., $6 = 1$ for the coefficient of x^3 and $5 = 0$ for the constant term. Now we compute their product:

$$\begin{aligned} f(x) \cdot g(x) &= (3x^4 + 2x^3 + 4) \cdot (4x^3 + 1) \\ &= 12x^7 + 3x^4 + 8x^6 + 2x^3 + 16x^3 + 4 \\ &= 2x^7 + 3x^6 + 3x^4 + 3x^3 + 4. \end{aligned}$$

Notice again that we reduce the coefficients in $\mathbb{Z}_5$.

Example 26.6. Consider the set $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$, where $i = \sqrt{-1}$ so that $i^2 = -1$. Here are some examples that illustrate how to add and multiply in $\mathbb{Z}_3[i]$:

- $(1 + 2i) + (2 + i) = 3 + 3i = 0 + 0i$ (or just 0), since $3 = 0$ in $\mathbb{Z}_3$.
- $\begin{aligned}(1 + 2i) \cdot (2 + i) &= 1 \cdot 2 + 1 \cdot i + 2i \cdot 2 + 2i \cdot i \\ &= 2 + 5i + 2 \cdot \boldsymbol{i^2} = 2 + 5i + 2 \cdot (-\mathbf{1}) = 0 + 5i = 2i, \text{ since } 5 = 2 \text{ in } \mathbb{Z}_3.\end{aligned}$

Note how we can replace i^2 with -1 to simplify the product. In an exercise, you'll verify that $\mathbb{Z}_3[i]$ is indeed a ring and perform some computations in it.

By definition, a ring need *not* be commutative under multiplication. In all of our examples thus far, multiplication has been commutative. In Example 26.8, we will see rings with non-commutative multiplication. This distinction is captured in the following definition.

Definition 26.7 (Commutative/non-commutative ring)**.** A ring R is called a *commutative ring* if its multiplication is commutative; i.e., $a \cdot b = b \cdot a$ for all $a, b \in R$. Otherwise, R is called a *non-commutative ring*.

Remark. By definition, a ring is always commutative under addition.

Example 26.8 (Matrix rings)**.** Consider the *matrix rings*:

- $M(\mathbb{R})$ is the set of all 2×2 matrices with entries in $\mathbb{R}$.
- $M(\mathbb{Z}_n)$ is the set of all 2×2 matrices with entries in $\mathbb{Z}_n$.

In an exercise, you'll exhibit matrices $\alpha, \beta \in M(\mathbb{Z}_{10})$ such that $\alpha \cdot \beta \neq \beta \cdot \alpha$, so that $M(\mathbb{Z}_{10})$ is a non-commutative ring. Similar examples show that $M(\mathbb{R})$ is non-commutative, too. In these rings, the additive and multiplicative identities (typically denoted 0 and 1) are the matrices $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, respectively.

26.2 Fundamental properties

In this section, we'll prove several familiar algebraic formulas, starting with perhaps the most interesting one. In the definition of a ring R, the element 0 is defined as the *additive* identity; i.e., $0 + a = a$ and $a + 0 = a$ for all $a \in R$. There is nothing in the definition that says how 0 behaves under multiplication. But we *know* from experience that $0 \cdot a = 0$ and $a \cdot 0 = 0$. Here's the theorem and its proof.

Theorem 26.9. *In a ring R, $0 \cdot a = 0$ and $a \cdot 0 = 0$ for all $a \in R$.*

PROOF. Let $a \in R$. Since 0 is an additive identity, $0+0 = 0$. Left-multiplying both sides by a, we obtain $a \cdot (0+0) = a \cdot 0$. Then use the distributive law to get $a \cdot 0 + a \cdot 0 = a \cdot 0$. Since $a, 0 \in R$ and R is closed under multiplication, we have $a \cdot 0 \in R$. Thus $a \cdot 0$ has an additive inverse, which we will call b, where $a \cdot 0 + b = 0$ and $b + a \cdot 0 = 0$. Add b to both sides of $a \cdot 0 + a \cdot 0 = a \cdot 0$ to obtain $(a \cdot 0 + a \cdot 0) + b = a \cdot 0 + b$, and then apply the associative law to find that $a \cdot 0 + (a \cdot 0 + b) = a \cdot 0 + b$. Since $a \cdot 0 + b = 0$, the above equation becomes $a \cdot 0 + 0 = 0$, so that $a \cdot 0 = 0$ as desired.

The argument for $0 \cdot a = 0$ follows similarly. We'll leave the details up to you as an exercise. ∎

Proof know-how. We began the above proof with a key step of writing $0 + 0 = 0$. This is the same "inserting the identity" technique that we saw in proofs about groups (e.g., the proof of Theorem 9.6). Since a ring is an additive group and also satisfies many (though not all) group properties under multiplication, the proof techniques from group theory often transfer well to proofs about rings.

Remark. Due to Theorem 26.9, there is no element $r \in R$ such that $0 \cdot r = 1$. Thus, we've proved that 0 does not have a multiplicative inverse in any ring. Hence, a ring is never a multiplicative group.

Below are several formulas involving additive inverses. Note that $-a$ and $-b$ refer to additive inverses of a and b, respectively, and -1 denotes the additive inverse of the multiplicative identity 1.

Theorem 26.10. *Let a and b be elements of a ring. Then:*

(a) $-(-a) = a$.

(b) $-1 \cdot a = -a$.

(c) $-1 \cdot -1 = 1$.

(d) $a \cdot (-b) = -(a \cdot b)$ *and* $(-a) \cdot b = -(a \cdot b)$.

(e) $(-a) \cdot (-b) = a \cdot b$.

We'll prove formulas (a) and (d). The rest are left for you as an exercise at the end of the chapter.

PROOF. First, we prove formula (a). The expression $-(-a)$ means "the additive inverse of $-a$." Observe that the additive inverse of $-a$ is the element a, since $(-a) + a = 0$ and $a + (-a) = 0$ by the definition of $-a$. Therefore, we have $-(-a) = a$ as desired.

For formula (d), we'll prove $a \cdot (-b) = -(a \cdot b)$ and leave the other one for you as an exercise. We will compute the sum $(a \cdot b) + (a \cdot (-b))$ and show that it equals 0. We have

$$(a \cdot b) + (a \cdot (-b)) = a \cdot (b + (-b)) = a \cdot 0 = 0.$$

Thus $(a \cdot b) + (a \cdot (-b)) = 0$. Since addition in a ring is commutative, we also have $(a \cdot (-b)) + (a \cdot b) = 0$. Hence, $a \cdot (-b)$ is the additive inverse of $a \cdot b$. Symbolically, this is written $a \cdot (-b) = -(a \cdot b)$, as desired. ■

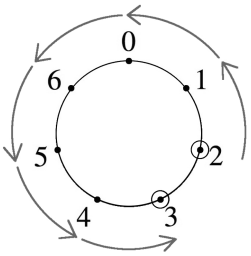
Proof know-how. In Example 8.12, we proved that $(g^{-1})^{-1} = g$ for a group element g. Formula (a) is simply the additive version of this. (Remember, every ring is an additive group.) In fact, the argument in Example 8.12 is almost identical to the one given in the proof above.

For formula (d), we use the additive version of the technique from the proof of Theorem 8.11. To show that $-\alpha = \beta$ (i.e., "the additive inverse of α is β"), we verify that $\alpha + \beta = 0$ and $\beta + \alpha = 0$. In the above proof, we apply this technique with $\alpha = a \cdot b$ and $\beta = a \cdot (-b)$.

When we worked with additive groups, we freely used subtraction, even though we never defined what we mean by $a - b$ where a and b are elements of an additive group. For instance, let H be a subgroup of an additive group G. Then for $a, b \in G$, we have $a + H = b + H$ if and only if $a - b \in H$ and $b - a \in H$. We'll continue to use subtraction in this manner. But we'll also define it formally once and for all.

Definition 26.11. Let a and b be elements of a ring. We define $a - b$ to mean $a + (-b)$.

Example 26.12 (Example 4.1 revisited)**.** When we first introduced $\mathbb{Z}_7$, here is how we described the computation of $2 - 6$: Start at 2 on the $\mathbb{Z}_7$ clock and move 6 units *counter-clockwise*. We land on 3, so that $2 - 6 = 3$ in $\mathbb{Z}_7$.



Using Definition 26.11, we have $2 - 6 = 2 + (-6)$. Moreover, $-6 = 1$ in $\mathbb{Z}_7$ (i.e., the additive inverse of 6 is 1), since $6 + 1 = 0$ and $1 + 6 = 0$. Therefore, $2 - 6 = 2 + (-6) = 2 + 1 = 3$, as before.

The proof of the following theorem is left for you as an exercise.

Theorem 26.13. *Let a, b, c be elements of a ring. Then* $a \cdot (b - c) = (a \cdot b) - (a \cdot c)$ *and* $(b - c) \cdot a = (b \cdot a) - (c \cdot a)$.

26.3 Units and zero divisors

Example 26.14. Recall that in $\mathbb{Z}_{10}$ we have $3 \cdot 7 = 1$ so that 3 and 7 are multiplicative inverses of each other. We write $3^{-1} = 7$ and $7^{-1} = 3$. Also in $\mathbb{Z}_{10}$, it is possible to have a pair of non-zero elements whose product is zero: $5 \cdot 2 = 0$ and $6 \cdot 5 = 0$, for instance.

The above example motivates the following definitions.

Definition 26.15 (Unit)**.** Let R be a ring. An element $u \in R$ is called a *unit* if it has a multiplicative inverse $u^{-1} \in R$ such that $u \cdot u^{-1} = 1$ and $u^{-1} \cdot u = 1$.

Definition 26.16 (Zero divisor)**.** Let R be a ring. A *non-zero* element $a \in R$ is called a *zero divisor* if there exists a non-zero $b \in R$ such that $a \cdot b = 0$.

Using these new terminologies, we can say that in $\mathbb{Z}_{10}$, 3 and 7 are units and 5, 2, and 6 are zero divisors. (What about the rest of the elements in $\mathbb{Z}_{10}$?) In any ring, the multiplicative identity 1 is a unit, because $1 \cdot 1 = 1$; i.e., it's a self-inverse under multiplication. The additive identity 0 is not a unit, because it does not have a multiplicative inverse; nor is 0 a zero divisor, since by definition a zero divisor must be non-zero.

Example 26.17. In each ring, we classify the *non-zero* elements as a unit, a zero divisor, or neither.

- In $\mathbb{Z}_{12}$: The units are 1, 5, 7, 11, as each is a self-inverse; i.e., $1 \cdot 1 = 5 \cdot 5 = 7 \cdot 7 = 11 \cdot 11 = 1$. The remaining non-zero elements of $\mathbb{Z}_{12}$, i.e., 2, 3, 4, 6, 8, 9, 10, are zero divisors, since $2 \cdot 6 = 0$, $3 \cdot 4 = 0$, $8 \cdot 9 = 0$, and $6 \cdot 10 = 0$.
- In $\mathbb{Z}_7$: The units are 1, 2, 3, 4, 5, 6. There are no zero divisors in $\mathbb{Z}_7$.
- In $\mathbb{Z}$: The units are 1 and -1, since $1 \cdot 1 = 1$ and $(-1) \cdot (-1) = 1$. There are no zero divisors in $\mathbb{Z}$. All other integers are neither a unit nor a zero divisor.
- In $\mathbb{R}$: The units are all non-zero elements of $\mathbb{R}$. There are no zero divisors in $\mathbb{R}$.

None of the rings in Example 26.17 has an element that is *both* a unit and a zero divisor. Here is why.

Theorem 26.18. *Let α be a non-zero ring element. Then α cannot be both a unit and a zero divisor.*

PROOF. Assume for contradiction that such an element exists. Let R be a ring, and suppose $\alpha \in R$, $\alpha \neq 0$, is both a unit and a zero divisor. Thus, there exists $\alpha^{-1} \in R$ such that $\alpha^{-1} \cdot \alpha = 1$; and there also exists $\beta \in R$, $\beta \neq 0$, such that $\alpha \cdot \beta = 0$. Multiplying both sides of $\alpha \cdot \beta = 0$ on the left by α^{-1}, we obtain

$$\alpha^{-1} \cdot (\alpha \cdot \beta) = \alpha^{-1} \cdot 0 \implies (\alpha^{-1} \cdot \alpha) \cdot \beta = \alpha^{-1} \cdot 0 \implies 1 \cdot \beta = 0.$$

(Note that $\alpha^{-1} \cdot 0 = 0$ follows from Theorem 26.9.) The last equation $1 \cdot \beta = 0$ implies that $\beta = 0$, which contradicts $\beta \neq 0$. Thus, such an element α does not exist. ■

26.4 Subrings

Analogous to a subgroup of a group is the notion of a *subring* of a ring. For example, $\mathbb{Z}$ is a subring of the ring $\mathbb{R}$, because (1) $\mathbb{Z}$ is a subset of $\mathbb{R}$ and (2) $\mathbb{Z}$ is a ring using the same operations as $\mathbb{R}$.

Definition 26.19 (Subring)**.** Let R be a ring. A subset $S \subseteq R$ is a *subring* of R if S is a ring using the operations of R.

Example 26.20. Consider the following subset of the ring $M(\mathbb{Z}_{10})$:

$$S = \left\{ \begin{bmatrix} a & b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{Z}_{10} \right\}.$$

To show that S is a subring of $M(\mathbb{Z}_{10})$, we must show that S satisfies the ring properties using the operations of $M(\mathbb{Z}_{10})$, i.e., matrix addition and multiplication. But S *inherits* these operations from $M(\mathbb{Z}_{10})$, and so we already know, for instance, that $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in S$ (because it's true for all $\alpha, \beta \in M(\mathbb{Z}_{10})$). So, we do not need to address this property when proving that S is a subring. This is analogous to how we don't address associativity when writing subgroup proofs. Thus, among the ring properties (1) through (9) from Section 26.1, we must show that S satisfies the following:

(1) S is closed under addition: If $a, b \in S$, then $a + b \in S$.

(3) S contains the additive identity of $M(\mathbb{Z}_{10})$, namely $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

(4) If $a \in S$, then $-a \in S$.

(6) S is closed under multiplication: If $a, b \in S$, then $a \cdot b \in S$.

(8) S contains the multiplicative identity of $M(\mathbb{Z}_{10})$, namely $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.

Note that since $M(\mathbb{Z}_{10})$ is a ring, we already know that $a+b$, $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, $-a$, $a \cdot b$, and $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ are all in $M(\mathbb{Z}_{10})$. Our goal is to show that these elements are contained in S.

Proof know-how. Let R be a ring. To show that a subset $S \subseteq R$ is a subring, we must show the following:

- S is closed under both operations.
- S contains both identity elements.
- S contains the *additive* inverses of its elements.

Theorem 26.21. *Let* $S = \left\{ \left[\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right] \mid a, b \in \mathbb{Z}_{10} \right\}$. *Then* S *is a subring of* $M(\mathbb{Z}_{10})$.

PROOF. We first show closure. Let $\alpha, \beta \in S$ so that $\alpha = \left[\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} c & d \\ d & c \end{smallmatrix}\right]$ for some $a, b, c, d \in \mathbb{Z}_{10}$. Then $\alpha + \beta = \left[\begin{smallmatrix} a+c & b+d \\ b+d & a+c \end{smallmatrix}\right]$ and $\alpha \cdot \beta = \left[\begin{smallmatrix} ac+bd & ad+bc \\ ad+bc & ac+bd \end{smallmatrix}\right]$. Both $\alpha + \beta$ and $\alpha \cdot \beta$ are in S, so that S is closed under addition and multiplication. The additive and multiplicative identities of $M(\mathbb{Z}_{10})$ are $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, respectively. Both have the form of the matrices in S and thus are contained in S. Lastly, we have $-\alpha = \left[\begin{smallmatrix} -a & -b \\ -b & -a \end{smallmatrix}\right] \in S$ so that S is a subring of $M(\mathbb{Z}_{10})$, as desired. ■

26.5 Group of units

In Chapter 10, we considered the following line of reasoning. We observed that $\mathbb{Z}_{10}$ admits both addition *and* multiplication. (We would now say that $\mathbb{Z}_{10}$ is a ring.) Although $\mathbb{Z}_{10}$ is an additive group, it is *not* a group under multiplication, because not every element of $\mathbb{Z}_{10}$ has a multiplicative inverse. For example, $5 \cdot x = 1$ isn't possible in $\mathbb{Z}_{10}$, and thus 5^{-1} doesn't exist. We salvaged the situation by defining the subset $U_{10} = \{a \in \mathbb{Z}_{10} \mid a \text{ is a unit}\}$. (**Note:** "$a$ is a unit" means the same as "a has a multiplicative inverse.") In essence, we removed from $\mathbb{Z}_{10}$ the elements without multiplicative inverses. And we showed that the resulting set U_{10} is, indeed, a multiplicative group.

Similarly, we started with the ring $M(\mathbb{Z}_{10})$ and obtained a multiplicative group $G(\mathbb{Z}_{10})$, defined by

$$G(\mathbb{Z}_{10}) = \{\alpha \in M(\mathbb{Z}_{10}) \mid \alpha \text{ is a unit}\}.$$

These two are examples of something more generally called the *group of units*.

Definition 26.22 (Group of units)**.** Let R be a ring. Then the subset $R^* = \{a \in R \mid a \text{ is a unit}\}$ is called the *group of units* of R.

Example 26.23. For each ring R, we find the group of units R^*.

- If $R = \mathbb{Z}_{10}$, then $R^* = U_{10}$.
- If $R = M(\mathbb{Z}_{10})$, then $R^* = G(\mathbb{Z}_{10})$.
- If $R = \mathbb{Z}$, then $R^* = \{1, -1\}$.
- If $R = \mathbb{R}$, then $R^* = \mathbb{R}^* = \{a \in \mathbb{R} \mid a \neq 0\}$, i.e., the set non-zero elements of $\mathbb{R}$.

The proof of the following theorem is similar to the proof showing that $G(\mathbb{Z}_{10})$ is a multiplicative group (in Section 10.2). We'll leave the details up to you.

Theorem 26.24. *Let R be a ring and define $R^* = \{a \in R \mid a \text{ is a unit}\}$. Then R^* is a multiplicative group.*

Exercises

1. Explain why each R is *not* a ring.

   (a) $R = \{0, 1, 2, 3, \ldots\}$ with integer addition and multiplication.

   (b) $R = 3\mathbb{Z}$ with integer addition and multiplication.

   (c) $R = U_7$ with addition and multiplication modulo 7.

2. For each $a \in \mathbb{Z}_{10}$, find $-a$ and a^{-1}, or explain why $-a$ or a^{-1} (or possibly both) doesn't exist.

   (a) $a = 3$. (b) $a = 4$. (c) $a = 1$. (d) $a = 0$. (e) $a = 9$.

3. For each $\alpha \in M(\mathbb{Z}_{10})$, find $-\alpha$ and α^{-1}, or explain why $-\alpha$ or α^{-1} (or possibly both) doesn't exist.

   (a) $\alpha = \left[\begin{smallmatrix} 4 & 7 \\ 1 & 7 \end{smallmatrix}\right]$. (b) $\alpha = \left[\begin{smallmatrix} 4 & 6 \\ 1 & 7 \end{smallmatrix}\right]$. (c) $\alpha = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. (d) $\alpha = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. (e) $\alpha = \left[\begin{smallmatrix} 9 & 0 \\ 0 & 9 \end{smallmatrix}\right]$.

4. Complete the proof of Theorem 26.9 by showing $0 \cdot a = 0$.

5. Complete the proof of Theorem 26.10 by proving parts (b), (c), (e) and the formula $(-a) \cdot b = -(a \cdot b)$ in part (d).

6. Prove Theorem 26.13.

7. Let a, b, c, d be elements of a ring. Prove that $(a+b)\cdot(c+d) = a\cdot c + a\cdot d + b\cdot c + b\cdot d$.

   Note: Since addition is a *binary* operation, we add only two elements at a time. But due to the associative law, there is no ambiguity when we write an expression such as $a \cdot c + a \cdot d + b \cdot c + b \cdot d$.

8. **Prove:** Let R be a ring. Then R has a unique (i.e., only one) multiplicative identity element.

   Note: Since R is an additive group, Theorem 8.9 implies R has a unique *additive* identity element.

9. **Prove:** Let R be a ring, and let $a \in R$. If a is a unit, then it has a unique multiplicative inverse in R.

   Note: Again, as R is an additive group, Theorem 8.10 ensures a has a unique *additive* inverse in R.

10. Find elements $a, b \in \mathbb{Z}_m$ such that a and b are additive inverses *and* multiplicative inverses of each other. Note that you must find the modulus m as well. (See Example 26.3.)

11. Verify that each set in Example 26.4 satisfies the ring properties.

12. Consider the set $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$ where $i = \sqrt{-1}$ so that $i^2 = -1$. See Example 26.6 for how to add and multiply in $\mathbb{Z}_3[i]$.
    (a) Verify that $\mathbb{Z}_3[i]$ satisfies the ring properties and that its multiplication is commutative.
    (b) Find all elements of $\mathbb{Z}_3[i]$. How many of them are there?
    (c) $1 + 2i \in \mathbb{Z}_3[i]$ has a multiplicative inverse. Find it.
    (d) Classify each *non-zero* element of $\mathbb{Z}_3[i]$ as a unit, a zero divisor, or neither.

    (This exercise is referenced in Example 27.14.)

13. (a) Find all zero divisors in $\mathbb{Z}_{20}$. Explain your reasoning.
    (b) Repeat part (a) with $\mathbb{Z}_{21}$.
    (c) Repeat part (a) with $\mathbb{Z}_{17}$.

14. Let $M(\mathbb{Z}_{10})$ be the ring of all 2×2 matrices with entries in $\mathbb{Z}_{10}$.
    (a) Come up with examples of a unit in $M(\mathbb{Z}_{10})$.
    (b) Come up with examples of a zero divisor in $M(\mathbb{Z}_{10})$.

15. Is there an element in $M(\mathbb{Z}_{10})$ that is *neither* a unit nor a zero divisor? Why or why not?

16. Let $S = \left\{ \begin{bmatrix} a & a-b \\ a-b & b \end{bmatrix} \mid a, b \in \mathbb{Z}_{10} \right\}$.
    (a) Find a few elements in S.
    (b) Prove that S is a subring of $M(\mathbb{Z}_{10})$.

17. Let $S = \left\{ \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \mid a, b \in \mathbb{Z}_{10} \right\}$.
    (a) Find a few elements in S.
    (b) Prove that S is a subring of $M(\mathbb{Z}_{10})$.

18. Let $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbb{Z}_{10} \text{ and } a + c = b + d \right\}$.
    (a) Find a few elements in S.
    (b) Prove that S is a subring of $M(\mathbb{Z}_{10})$.

19. Come up with another subset S like the ones in Exercises #16, #17, and #18. Then prove that your subset S is a subring of $M(\mathbb{Z}_{10})$.

20. Let R be a ring, and let S and T be its subrings. Define $M = S \cap T = \{r \in R \mid r \in S \text{ and } r \in T\}$; i.e., M is the *intersection* of S and T. Prove that M is a subring of R.

    Hint: How can you use the result of Chapter 11, Exercise #13 to reduce your work here?

21. Prove Theorem 26.24.

22. Let R be a ring. Using the ring properties (1) through (9), but *without* assuming property (5) (i.e., the commutative law for addition), prove that $a + b = b + a$ for all $a, b \in R$.

23. Given a ring R, define the *center* of R by
$$Z(R) = \{z \in R \mid z \cdot r = r \cdot z \text{ for all } r \in R\}.$$
    Prove that $Z(R)$ is a subring of R.

27

Integral Domains and Fields

The rings $\mathbb{Z}_{13}$ and $\mathbb{Z}_{15}$ are similar in many respects. They're both finite rings with 13 and 15 elements, respectively. Each is a *commutative ring*, which means its multiplication is commutative; i.e., $a \cdot b = b \cdot a$ for all of its elements. (Recall that by definition, a ring is always commutative under addition.) However, $\mathbb{Z}_{13}$ and $\mathbb{Z}_{15}$ are also very different, and how they're different is the focus of this chapter.

For instance, consider the equation $5x = 10$. In $\mathbb{Z}_{13}$, this equation has one solution; namely $x = 2$. However, the same equation has *five* solutions in $\mathbb{Z}_{15}$. Aside from $x = 2$, we have $5 \cdot 8 = 10 \pmod{15}$, so that $x = 8$ is another solution in $\mathbb{Z}_{15}$. We'll leave it up to you to verify that $x = 5, 11, 14$ are also solutions to this equation in $\mathbb{Z}_{15}$. Why are there so many solutions in $\mathbb{Z}_{15}$? Let's find out!

27.1 Integral domains

Example 27.1. In $\mathbb{Z}_{12}$, there is a pair of *non-zero* elements whose product is zero: $3 \cdot 4 = 0$, for instance. Recall that such elements are called *zero divisors* in $\mathbb{Z}_{12}$. Other zero divisors in $\mathbb{Z}_{12}$ include 2, 6, 8, 9, 10. Note that $2 \cdot 6 = 0$, $8 \cdot 9 = 0$, and $6 \cdot 10 = 0$ in $\mathbb{Z}_{12}$.

Some rings do not have any zero divisors, which motivates the following definition.

Definition 27.2 (Integral domain)**.** A commutative ring is called an *integral domain* if it does *not* contain any zero divisors.

Example 27.3. If a and b are non-zero integers, then their product ab is also non-zero. Hence, there are no zero divisors in $\mathbb{Z}$, which implies that $\mathbb{Z}$ is an integral domain.

Example 27.4. In the ring $\mathbb{Z}_{13} = \{0, 1, 2, \ldots, 12\}$, every non-zero element is a unit, i.e., an element with a multiplicative inverse. Then by Theorem 26.18, these units cannot be zero divisors. Thus, there are no zero divisors in $\mathbb{Z}_{13}$, so that $\mathbb{Z}_{13}$ is an integral domain. The same argument shows that the set of real numbers $\mathbb{R}$ is an integral domain as well.

Example 27.5 (Non-examples)**.** We saw in Example 27.1 that $\mathbb{Z}_{12}$ contains zero divisors. Thus, it's not an integral domain. Likewise, we have $2 \cdot 5 = 0$ in $\mathbb{Z}_{10}$ and $5 \cdot 6 = 0$ in $\mathbb{Z}_{15}$. Since they have zero divisors, neither $\mathbb{Z}_{10}$ nor $\mathbb{Z}_{15}$ is an integral domain.

A useful feature of integral domains is cancellation. If $5b = 5c$ in $\mathbb{Z}$, for example, we can conclude that $b = c$. Here is the generalization.

Theorem 27.6 (Cancellation in an integral domain)**.** *Let a, b, c be elements of an integral domain R and suppose $a \neq 0$. If $ab = ac$, then $b = c$.*

Remark. An important assumption in this theorem is that $a \neq 0$. Otherwise, we could have a scenario such as $0 \cdot 4 = 0 \cdot 5$ in $\mathbb{Z}$, which does *not* imply that $4 = 5$.

How can we prove this theorem? Observe that it looks just like left cancellation in groups (Theorem 8.18). Here's how we proved it in a group setting: Assuming $ab = ac$, we left-multiplied both sides of the equation by a^{-1}. Unfortunately, the multiplicative inverse a^{-1} need not exist in a ring. For instance, 5^{-1} does not exist in $\mathbb{Z}$. However, we can subtract ac from both sides of the equation. This is a valid step, since additive inverses always exist in a ring. With this key insight, let's proceed to the proof of Theorem 27.6.

PROOF. Assume $ab = ac$. Then $ab - ac = 0$ and thus $a \cdot (b - c) = 0$. Since R is an integral domain, $a \cdot (b - c) = 0$ implies that a or $b - c$ must be zero. But $a \neq 0$; hence $b - c$ must equal 0. Thus $b = c$. ■

Proof know-how. In the above proof, we already *know* that R is an integral domain. Thus if $\alpha \cdot \beta = 0$, then we can conclude that either α or β (or possibly both) must be zero. Otherwise, α and β would both be non-zero, making them zero divisors (since $\alpha \cdot \beta = 0$), which contradicts the fact that R is an integral domain. For the above proof, we apply this technique with $\alpha = a$ and $\beta = b - c$.

The Proof know-how above captures an important feature of a ring that we define here.

Definition 27.7. The following property of a ring R is called the *zero product property*: Let $\alpha, \beta \in R$. If $\alpha \cdot \beta = 0$, then $\alpha = 0$ or $\beta = 0$.

Theorem 27.8. *Let R be a commutative ring. Then R is an integral domain if and only if R satisfies the zero product property.*

PROOF. We must prove two implications:

- If R is an integral domain, then R satisfies the zero product property.
- If R satisfies the zero product property, then R is an integral domain.

The first implication was proved in the Proof know-how above. We'll leave the proof of the second implication to you as an exercise at the end of the chapter. ■

Example 27.9. Consider the equation $5x = 10$ at the beginning of this chapter. Let's view it in $\mathbb{Z}_{13}$, which is an integral domain. (See Example 27.4.) Our equation can be rewritten as $5 \cdot x = 5 \cdot 2$, and $5 \neq 0$ in $\mathbb{Z}_{13}$. Therefore, Theorem 27.6 applies and thus $5 \cdot x = 5 \cdot 2$ implies $x = 2$, as expected.

Example 27.10 (Non-example)**.** Consider again the equation $5x = 10$ or $5 \cdot x = 5 \cdot 2$. Now view it in $\mathbb{Z}_{15}$, which is not an integral domain. (See Example 27.5.) Thus, Theorem 27.6 does not apply. Let's see what goes wrong if we proceed as in its proof. Starting with $5 \cdot x = 5 \cdot 2$, we subtract $5 \cdot 2$ from both sides to obtain $5 \cdot x - 5 \cdot 2 = 0$. Thus, $5 \cdot (x-2) = 0$. However, since $\mathbb{Z}_{15}$ is not an integral domain, we cannot conclude that 5 or $x-2$ must be zero. In fact, since $5 \cdot 6 = 0$ in $\mathbb{Z}_{15}$, we could have $5 \cdot (8-2) = 5 \cdot 6 = 0$, so that $x = 8$ is a possible solution to $5x = 10$ in $\mathbb{Z}_{15}$. We'll leave it to you as an exercise to find other solutions to $5x = 10$ in $\mathbb{Z}_{15}$ using the zero divisors β of the form $5 \cdot \beta = 0$.

Strange things happen when we're *not* in an integral domain, like $5x = 10$ having five solutions in $\mathbb{Z}_{15}$. Here is another interesting occurrence in $\mathbb{Z}_{15}$.

Example 27.11 (Non-example)**.** Consider the equation $x^2 - 6x + 8 = 0$. Factoring the left-hand side, we obtain $(x-2) \cdot (x-4) = 0$. When viewed in $\mathbb{Z}$, an integral domain, we can conclude via the zero product property that $x-2 = 0$ or $x-4 = 0$, so that $x = 2$ or $x = 4$ are the only solutions to this equation.

Now, $x = 2, 4$ are solutions in $\mathbb{Z}_{15}$ as well. But $\mathbb{Z}_{15}$ is not an integral domain; i.e., it has zero divisors. For instance, we could have $(7-2) \cdot (7-4) = 5 \cdot 3 = 0$. Thus $x = 7$ is another solution to the equation. We also have $(14-2) \cdot (14-4) = 12 \cdot 10 = 0$ in $\mathbb{Z}_{15}$, so that $x = 14$ is a solution, too. Thus, the quadratic equation $x^2 - 6x + 8 = 0$ has *four* solutions in $\mathbb{Z}_{15}$: $x = 2, 4, 7, 14$. We'll leave it to you to verify that there is no other solution in $\mathbb{Z}_{15}$. (Here is one approach: Substitute the remaining 11 values of $\mathbb{Z}_{15}$ into $x^2 - 6x + 8$.)

27.2 Fields

Example 27.12. In $\mathbb{Z}_7 = \{0, 1, 2, 3, 4, 5, 6\}$, all non-zero elements are *units*; i.e., they have multiplicative inverses. Note that $1 \cdot 1 = 1$, $2 \cdot 4 = 1$, $3 \cdot 5 = 1$, and $6 \cdot 6 = 1$ in $\mathbb{Z}_7$. Similarly, all non-zero elements of $\mathbb{R}$ are units. While a ring is *never* a multiplicative group, since the additive identity element 0 does not have a multiplicative inverse, examples like $\mathbb{Z}_7$ and $\mathbb{R}$ come awfully close!

Rings like $\mathbb{Z}_7$ and $\mathbb{R}$, which are *almost* multiplicative groups, are examples of a *field*.

Definition 27.13 (Field)**.** A commutative ring is called a *field* if every non-zero element is a unit.

Examples of a field include $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_7$, and $\mathbb{Z}_{13}$. Below are a couple more examples.

Example 27.14 (Example 26.6 revisited)**.** Consider the set $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$, where $i^2 = -1$. In Chapter 26, Exercise #12, you verified that $\mathbb{Z}_3[i]$ is a commutative ring with 9 elements; namely

$$\mathbb{Z}_3[i] = \{0,\ 1,\ 2,\ i,\ 1+i,\ 2+i,\ 2i,\ 1+2i,\ 2+2i\}.$$

It turns out that every non-zero element of $\mathbb{Z}_3[i]$ is a unit. For instance, to find the multiplicative inverse of $2 + i \in \mathbb{Z}_3[i]$, we must find an element $a + bi \in \mathbb{Z}_3[i]$ such

that $(2+i)\cdot(a+bi)=1$. The left-hand side of this equation expands to

$$\begin{aligned}(2+i)\cdot(a+bi) &= 2\cdot a+2\cdot bi+i\cdot a+i\cdot bi\\ &= 2a+(2b+a)i+bi^2\\ &= 2a+(2b+a)i+b\cdot(-1)\\ &= (2a-b)+(2b+a)i,\end{aligned}$$

so that $(2+i)\cdot(a+bi)=(2a-b)+(2b+a)i$. Thus, the equation $(2+i)\cdot(a+bi)=1$ can be rewritten as $(2a-b)+(2b+a)i=1+0i$, from which we obtain $2a-b=1$ and $2b+a=0$. Solving this system of equations in $\mathbb{Z}_3$, we find $a=1$, $b=1$. Therefore, $1+i$ is the multiplicative inverse of $2+i$.

Proceeding similarly, we can find the multiplicative inverse of every non-zero element in $\mathbb{Z}_3[i]$. (We'll leave the details for you to complete in Chapter 26, Exercise #12.) Thus, $\mathbb{Z}_3[i]$ is a field.

Example 27.15. Consider the set $\mathbb{Q}(\sqrt{2})=\{a+b\sqrt{2}\mid a,b\in\mathbb{Q}\}$. The following examples illustrate how to add and multiply in $\mathbb{Q}(\sqrt{2})$. For multiplication, note that $\sqrt{2}\cdot\sqrt{2}=2$.

- $(2+15\sqrt{2})+(7-8\sqrt{2})=(2+7)+(15-8)\sqrt{2}=9+7\sqrt{2}$.
- $$\begin{aligned}(2+15\sqrt{2})\cdot(7-8\sqrt{2}) &= 2\cdot 7+2\cdot(-8\sqrt{2})+(15\sqrt{2})\cdot 7+(15\sqrt{2})\cdot(-8\sqrt{2})\\ &= 14+89\sqrt{2}+(-120)\cdot(\sqrt{2})^2\\ &= -226+89\sqrt{2}.\end{aligned}$$

In an exercise, you'll verify that $\mathbb{Q}(\sqrt{2})$ is a commutative ring. Furthermore, we claim that every non-zero element of $\mathbb{Q}(\sqrt{2})$ is a unit. To find the multiplicative inverse of $7+3\sqrt{2}\in\mathbb{Q}(\sqrt{2})$, for instance, consider the calculation below, where $(7+3\sqrt{2})\cdot(7-3\sqrt{2})=7^2-3^2\cdot 2=31$:

$$\frac{1}{7+3\sqrt{2}}=\frac{1}{7+3\sqrt{2}}\cdot\frac{7-3\sqrt{2}}{7-3\sqrt{2}}=\frac{7}{31}-\frac{3}{31}\sqrt{2}.$$

Therefore $(7+3\sqrt{2})\cdot\left(\frac{7}{31}-\frac{3}{31}\sqrt{2}\right)=1$, so that $7+3\sqrt{2}$ and $\frac{7}{31}-\frac{3}{31}\sqrt{2}$ are multiplicative inverses of each other in $\mathbb{Q}(\sqrt{2})$. We'll leave it for you in an exercise to generalize this calculation and show that every non-zero element $a+b\sqrt{2}\in\mathbb{Q}(\sqrt{2})$ has a multiplicative inverse. Thus, $\mathbb{Q}(\sqrt{2})$ is a field.

Example 27.16 (Non-example)**.** Although $\mathbb{Z}$ is an integral domain, it is *not* a field. Most non-zero integers (except for 1 and -1) are not units, because they don't have multiplicative inverses in $\mathbb{Z}$.

Cancellation also holds in a field. We'll leave the proof of the following theorem to you as an exercise. Here's a hint: A field is *almost* a multiplicative group.

Theorem 27.17 (Cancellation in field)**.** *Let a, b, c be elements of a field R and suppose $a\neq 0$. If $ab=ac$, then $b=c$.*

Now, $\mathbb{Z}$ is an integral domain that's not a field. We'll examine other such examples soon. (If you're curious, the polynomial ring $\mathbb{Z}_5[x]$, defined in Example 26.5, is another integral domain that is not a field. More on this later.) But does there exist a field that's not an integral domain? The answer is "No," as implied by the following theorem.

Theorem 27.18. *Every field is an integral domain.*

To prove this theorem, we will use the second implication of Theorem 27.8; namely: If R satisfies the zero product property, then R is an integral domain.

PROOF. Let R be a field. To show that R is an integral domain, we will show that R satisfies the zero product property. Let $\alpha, \beta \in R$ and assume $\alpha \cdot \beta = 0$. We must show that $\alpha = 0$ or $\beta = 0$. Assume that $\alpha \neq 0$. (We will show that $\beta = 0$.) The equation $\alpha \cdot \beta = 0$ can be rewritten as $\alpha \cdot \beta = \alpha \cdot 0$. Since R is a field and $\alpha \neq 0$, we apply Theorem 27.17 to cancel α from $\alpha \cdot \beta = \alpha \cdot 0$. Thus, $\beta = 0$ as desired. ■

Proof know-how. In the above proof, we prove an "or" statement; i.e., we must show that (1) $\alpha = 0$ or (2) $\beta = 0$. Observe that conclusion (1) is either true or false. If it's true, then we're done with the proof. Thus, we *assume* that (1) is false and *prove* that (2) is true. In the proof itself, though, we leave out the rationale behind this approach and start right away with the assumption that (1) is false. (Compare this with the proof of Theorem 20.15.)

To summarize, every field is an integral domain (Theorem 27.18), but not every integral domain is a field ($\mathbb{Z}$, for instance). But it turns out that every *finite* integral domain is a field. This is based on the theorem below, whose proof is left for you as an exercise at the end of the chapter.

Theorem 27.19. *Let R be a finite ring. If $\alpha \in R$ is a non-zero element, then α is a unit or a zero divisor.*

Example 27.20 (Example 26.17 revisited)**.** In a finite ring $\mathbb{Z}_{12}$, every non-zero element is a unit or a zero divisor. The units are 1, 5, 7, 11. The zero divisors are 2, 3, 4, 6, 8, 9, 10. Every non-zero element of $\mathbb{Z}_{12}$ has been accounted for. In other words, there's no non-zero element of $\mathbb{Z}_{12}$ that's *neither* a unit nor a zero divisor. (Such "neither" elements exist in $\mathbb{Z}$. For instance, $5 \in \mathbb{Z}$ is neither a unit nor a zero divisor.)

Using Theorem 27.19, we can prove the following:

Theorem 27.21. *Every finite integral domain is a field.*

PROOF. Let R be a finite integral domain. Let $\alpha \in R$ be a non-zero element. Since R is an integral domain, α is *not* a zero divisor. But by Theorem 27.19, a non-zero element in a finite ring must be a unit or a zero divisor. Since α is not a zero divisor, it must be a unit. Thus R is a field, as desired. ■

Proof know-how. To show that R is a field, we must show that every non-zero element of R is a unit. Typically, we accomplish this by considering a non-zero element $\alpha \in R$ and showing that α is a unit.

27.3 Idempotent elements

Example 27.22. Consider $5 \in \mathbb{Z}_{10}$. We have $5^2 = 5$, from which we derive the following:

$$\begin{aligned} 5^2 &= 5, \\ 5^3 &= \mathbf{5^2} \cdot 5 = \mathbf{5} \cdot 5 = 5, \\ 5^4 &= \mathbf{5^3} \cdot 5 = \mathbf{5} \cdot 5 = 5, \\ 5^5 &= \mathbf{5^4} \cdot 5 = \mathbf{5} \cdot 5 = 5, \\ &\vdots \end{aligned}$$

Thus we have $5^n = 5$ for all positive integer exponents n.

Example 27.22 motivates the following definition.

Definition 27.23. A ring element a is called an *idempotent element* (or simply an *idempotent*) if $a^2 = a$.

Remark. Let a be an idempotent in a ring. Then $a^n = a$ for all positive integers n. (See Example 27.22.)

Example 27.24. In Example 27.22, we saw that $5 \in \mathbb{Z}_{10}$ is an idempotent. In $\mathbb{Z}_{10}$, we also have $0^2 = 0$, $1^2 = 1$, and $6^2 = 6$. Thus, the idempotents in $\mathbb{Z}_{10}$ are 0, 1, 5, and 6. (We'll leave it to you to verify that no other element of $\mathbb{Z}_{10}$ satisfies $a^2 = a$.)

Example 27.25. To find the idempotents in $\mathbb{Z}_7$, let's see which elements satisfy the equation $a^2 = a$.

$$\mathbf{0^2 = 0},\ \mathbf{1^2 = 1},\ 2^2 = 4,\ 3^2 = 2,\ 4^2 = 2,\ 5^2 = 4,\ 6^2 = 1.$$

Thus, the only idempotents in $\mathbb{Z}_7$ are 0 and 1.

Example 27.26. To find the idempotents in $\mathbb{Z}_8$, let's see which elements satisfy the equation $a^2 = a$.

$$\mathbf{0^2 = 0},\ \mathbf{1^2 = 1},\ 2^2 = 4,\ 3^2 = 1,\ 4^2 = 0,\ 5^2 = 1,\ 6^2 = 4,\ 7^2 = 1.$$

Thus, the only idempotents in $\mathbb{Z}_8$ are 0 and 1.

Example 27.27. In any ring, $0^2 = 0$ and $1^2 = 1$. Thus, the additive identity 0 and the multiplicative identity 1 are idempotents. Since they are always idempotents, 0 and 1 are often called *trivial idempotents*.

The following theorem is left for you to prove in an exercise. Note that its converse is false, since $\mathbb{Z}_8$ can serve as a counterexample. (See Example 27.26.)

Theorem 27.28. *Let R be a commutative ring. If R is an integral domain, then the only idempotents of R are* 0 *and* 1.

Example 27.29. The following calculations are done in the ring $\mathbb{Z}_{30}$:

- $6^2 = 6$ so that 6 is an idempotent. And $1 - 6 = 25$ and $25^2 = 25$. Thus 25 is an idempotent, too.
- $10^2 = 10$ so that 10 is an idempotent. And $1 - 10 = 21$ and $21^2 = 21$. Thus 21 is an idempotent, too.
- $15^2 = 15$ so that 15 is an idempotent. And $1 - 15 = 16$ and $16^2 = 16$. Thus 16 is an idempotent, too.
- $1^2 = 1$ so that 1 is an idempotent. And $1 - 1 = 0$ and $0^2 = 0$. Thus 0 is an idempotent, too.

Here is the generalization.

Theorem 27.30. *In a ring, if a is an idempotent, then $1 - a$ is also an idempotent.*

PROOF. Assume a is an idempotent so that $a^2 = a$. Then

$$(1 - a)^2 = 1 - 2a + \boldsymbol{a}^2 = 1 - 2a + \boldsymbol{a} = 1 - a,$$

so that $(1 - a)^2 = 1 - a$. Thus, $1 - a$ is an idempotent. ■

Exercises

1. Consider the following commutative rings: $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}_{19}$, $\mathbb{Z}_{20}$.
   (a) Determine whether or not each of them is an integral domain.
   (b) Determine whether or not each of them is a field.
2. Prove each of these claims about the commutative ring $\mathbb{Z}_m$.
   (a) If p is prime, then $\mathbb{Z}_p$ is a field.
   (b) If m is composite, then $\mathbb{Z}_m$ is not an integral domain.
3. Consider the ring $\mathbb{Z}_5[i] = \{a + bi \mid a, b \in \mathbb{Z}_5\}$. (Here, $i = \sqrt{-1}$ so that $i^2 = -1$.)
   (a) How many elements are in $\mathbb{Z}_5[i]$? Explain your reasoning.
   (b) The element $2 + 3i \in \mathbb{Z}_5[i]$ is a unit. Find its multiplicative inverse.
   (c) The element $\alpha = 2 + i \in \mathbb{Z}_5[i]$ is a zero divisor. Find a non-zero $\beta \in \mathbb{Z}_5[i]$ such that $\alpha \cdot \beta = 0$.
   (d) Is $\mathbb{Z}_5[i]$ an integral domain, a field, or neither?
4. Complete the proof of Theorem 27.8 by proving its second implication.
5. Prove Theorem 27.17.
6. **Prove:** Let a and b be elements of an integral domain. If $a^2 = b^2$, then $a = b$ or $a = -b$.
7. Using a counterexample, explain how the statement in Exercise #6 is false in a commutative ring that is *not* an integral domain.

8. **Prove:** Let a and b be elements of a commutative ring. If ab is a zero divisor, then a or b is a zero divisor.

   Note: The Proof know-how after Theorem 27.18 describes how to prove an "or" statement.

9. Let S be a subring of a commutative ring R (i.e., $S \subseteq R$). Determine if each statement is true or false. If it's true, prove it. If it's false, give a counterexample.
   (a) If S is a field, then R is a field.
   (b) If R is a field, then S is a field.
   (c) If S is an integral domain, then R is an integral domain.
   (d) If R is an integral domain, then S is an integral domain.

10. Exercise #9(b) happens to be a false statement. (Have you found a counterexample?) However, Anita thought that it was true and even wrote its proof:

    > Assume that R is a field. Let $\alpha \in S$ be a non-zero element. Since $S \subseteq R$, we have $\alpha \in R$. Thus, α is a non-zero element in a field R, which implies that α is a unit. Hence S is a field.

    Explain the error in this proof.

11. Let R be a commutative ring, and let $a, b \in R$. Determine if each statement is true or false. If it's true, prove it. If it's false, give a counterexample.
    (a) If a and b are zero divisors, then $a + b$ is a zero divisor. (Assume $a + b \neq 0$.)
    (b) If a and b are zero divisors, then $a \cdot b$ is a zero divisor. (Assume $a \cdot b \neq 0$.)

12. **Note:** Read Example 27.10 before working on this exercise.
    (a) Find all $\beta \in \mathbb{Z}_{15}$ such that $5 \cdot \beta = 0$ in $\mathbb{Z}_{15}$.
    (b) Use your work in part (a) to find all solutions to $5x = 10$ in $\mathbb{Z}_{15}$.

13. Consider the equation $x^2 - 7x + 12 = 0$.
    (a) Find all solutions in $\mathbb{Z}_{13}$.
    (b) Find all solutions in $\mathbb{Z}_{12}$.
    (c) Find all solutions in $\mathbb{Z}_{15}$.
    (d) Explain why this equation has only two solutions in an integral domain.

14. Consider $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. See Example 27.15 for how to add and multiply in $\mathbb{Q}(\sqrt{2})$.
    (a) Verify that $\mathbb{Q}(\sqrt{2})$ is a commutative ring.
    (b) Find the multiplicative inverse of $8+5\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Do likewise for $6-3\sqrt{2} \in \mathbb{Q}(\sqrt{2})$.
    (c) Prove that $\mathbb{Q}(\sqrt{2})$ is a field by showing that every non-zero element of $\mathbb{Q}(\sqrt{2})$ has a multiplicative inverse. (This exercise is referenced in Example 33.10.)

15. (a) Find all elements of $\mathbb{Z}$ that are self-inverses under multiplication; i.e., $a^2 = 1$.
    (b) Repeat part (a) with $\mathbb{Z}_7$; with $\mathbb{Z}_{12}$; with $\mathbb{Z}_{16}$; with $\mathbb{Z}_{13}$.
    (c) What conjecture do you have?

16. Determine all elements of an integral domain that are self-inverses under multiplication.

17. Let R be a commutative ring that satisfies the cancellation property; i.e., if $ab = ac$ with $a \neq 0$, then $b = c$. Prove that R is an integral domain.

18. **Prove:** Every non-zero element of the ring $\mathbb{Z}_m$ is a unit or a zero divisor.

    Remark. This is a special case of Theorem 27.19.

19. Prove Theorem 27.19.

20. Find all idempotents in $\mathbb{Z}_{12}$; in $\mathbb{Z}_{13}$; in $\mathbb{Z}_{14}$; in $\mathbb{Z}_{15}$.

21. **Prove:** If a ring element α is both a unit and an idempotent, then $\alpha = 1$.

22. (a) Verify that $3 \in \mathbb{Z}_6$, $5 \in \mathbb{Z}_{10}$, $7 \in \mathbb{Z}_{14}$, $9 \in \mathbb{Z}_{18}$ are all idempotents.
    (b) What conjecture do you have?

23. **Prove:** If m is odd, then $m \in \mathbb{Z}_{2m}$ is an idempotent.

24. Consider the following statement: If a ring element a is an idempotent, then $1-2a$ is a self-inverse under multiplication.
    (a) Create a few examples to illustrate this statement.
    (b) Prove the statement.

25. **Prove:** Let $m = 2^k$ where k is a positive integer. Then $\mathbb{Z}_m$ contains only the trivial idempotents.

26. **Prove:** Let R be a ring in which every element is an idempotent. Then R is commutative.

28

Polynomial Rings, Part I

Prior to studying abstract algebra, many students are used to thinking about *sets* as containing numbers, or number-like objects. This is true in calculus, which deals with functions on the set $\mathbb{R}$ of real numbers. In linear algebra, we study the vector space $\mathbb{R}^n$, which is the set of *vectors* or lists of real numbers. In abstract algebra, however, we study sets whose elements can be objects that are quite different from numbers. (We also study familiar sets like $\mathbb{Z}$ and $\mathbb{R}$.) We have examined the group D_4 which contains *symmetries* of a square; the group S_3 which contains *permutations* of the set $\{1, 2, 3\}$; and the quotient group G/H whose elements are *cosets* of H. The power of *abstraction* in mathematics lies in its ability to generalize and prove statements that apply to all of these groups, despite the vastly different types of objects they contain.

We will continue in this theme in the next several chapters and study an important family of commutative rings called *polynomial rings*. In calculus, we work with polynomials such as $f(x) = 3x^4 - 7x^2 + 4$, and we treat them as *functions*. In abstract algebra, we treat polynomials as *elements* of a ring, although we will still substitute values into them on occasion (e.g., $f(2) = 3 \cdot 2^4 - 7 \cdot 2^2 + 4 = 24$).

28.1 Examples and definition

Example 28.1. Let $\mathbb{Z}[x]$ be the set of all polynomials with coefficients in $\mathbb{Z}$. Examples of such polynomials include $f(x) = 3x^4 - 7x^2 + 4$ and $g(x) = 4x^2 + 1$. To add polynomials, we add the like terms, i.e., terms with the same exponent for the variable x:

$$f(x) + g(x) = (3x^4 - 7x^2 + 4) + (4x^2 + 1) = 3x^4 + (-7x^2 + 4x^2) + (4 + 1) = 3x^4 - 3x^2 + 5.$$

To multiply polynomials, we use the distribute law repeatedly:

$$\begin{aligned} f(x) \cdot g(x) &= (3x^4 - 7x^2 + 4) \cdot (4x^2 + 1) \\ &= 3x^4 \cdot (4x^2 + 1) - 7x^2 \cdot (4x^2 + 1) + 4 \cdot (4x^2 + 1) \\ &= (3x^4 \cdot 4x^2 + 3x^4 \cdot 1) + ((-7x^2) \cdot 4x^2 + (-7x^2) \cdot 1) + (4 \cdot 4x^2 + 4 \cdot 1) \\ &= (12x^6 + 3x^4) + (-28x^4 - 7x^2) + (16x^2 + 4) \\ &= 12x^6 + (3x^4 - 28x^4) + (-7x^2 + 16x^2) + 4 \\ &= 12x^6 - 25x^4 + 9x^2 + 4. \end{aligned}$$

The product $f(x) \cdot g(x)$ can also be computed using an *area model* shown below. Note how the terms of $f(x)$ are placed to the left of the rectangle, while the terms of $g(x)$ are placed on top. We compute the "area" of each square by multiplying its "sides." For example, the top-left square has area $3x^4 \cdot 4x^2 = 12x^6$. By adding the areas of the squares, we obtain the total area of the rectangle, i.e., the product $f(x) \cdot g(x)$.

	$4x^2$	$+$	1
$3x^4$	$12x^6$		$3x^4$
$+$			
$-7x^2$	$-28x^4$		$-7x^2$
$+$			
4	$16x^2$		4

Note that both the sum $f(x) + g(x)$ and the product $f(x) \cdot g(x)$ are in $\mathbb{Z}[x]$. You should convince yourself that $\mathbb{Z}[x]$ is indeed closed under addition and multiplication.

The additive and multiplicative identities of $\mathbb{Z}[x]$ are the constant polynomials 0 and 1, respectively. Moreover, every polynomial in $\mathbb{Z}[x]$ has an additive inverse in $\mathbb{Z}[x]$. For $f(x) = 3x^4 - 7x^2 + 4 \in \mathbb{Z}[x]$ above, its additive inverse is $-f(x) = -3x^4 + 7x^2 - 4$, which is also in $\mathbb{Z}[x]$. Note that

$$\begin{aligned} f(x) + (-f(x)) &= (3x^4 - 7x^2 + 4) + (-3x^4 + 7x^2 - 4) \\ &= (3x^4 - 3x^4) + (-7x^2 + 7x^2) + (4 - 4) \\ &= 0 + 0 + 0, \end{aligned}$$

so that $f(x) + (-f(x)) = 0$. A similar calculation shows that $-f(x) + f(x) = 0$.

Although we won't provide a rigorous proof, it turns out that $\mathbb{Z}[x]$ is a ring. In fact, it's a *commutative* ring, since $\alpha(x) \cdot \beta(x) = \beta(x) \cdot \alpha(x)$ for all $\alpha(x), \beta(x) \in \mathbb{Z}[x]$. In an exercise, you'll compute $g(x) \cdot f(x)$ using the polynomials above and compare the product to $f(x) \cdot g(x)$.

Example 28.2. The ring of integers $\mathbb{Z}$ is a subring of $\mathbb{Z}[x]$. Here, we're viewing the elements of $\mathbb{Z}$ (such as 0, 1, -3, and 6) as constant polynomials.

Here is the generalization of the above examples.

Definition 28.3 (Polynomial ring)**.** Let R be a commutative ring. Then $R[x]$, called a *polynomial ring*, is the set of all polynomials with *coefficients* in R. In this context, we refer to R as the *coefficient ring*.

Remark. An element $f(x) \in R[x]$ has the form $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, where the *coefficients* a_i are in R. (**Note:** n is a non-negative integer.)

Generalizing from the example of $\mathbb{Z}[x]$, here are some key properties of $R[x]$. We emphasize that, by the definition of $R[x]$, the coefficient ring R must be commutative.

- $R[x]$ is also a commutative ring.
- R is a subring of $R[x]$. (**Note:** The elements of R are the *constant* polynomials.)

Example 28.4. $\mathbb{Z}_5[x]$ is another example of a polynomial ring, with the coefficient ring $\mathbb{Z}_5$. Here, we view $\mathbb{Z}_5$ as containing the constant polynomials 0, 1, 2, 3, and 4. Therefore, $\mathbb{Z}_5$ is a subring of $\mathbb{Z}_5[x]$.

To add and multiply in $\mathbb{Z}_5[x]$, we use the same methods as in $\mathbb{Z}[x]$, described in Example 28.1, except that we reduce the coefficients modulo 5. Be careful, though. In $\mathbb{Z}_5[x]$, we have $7x = 2x$, but $x^7 \neq x^2$. The key here is that we reduce only the *coefficients* modulo 5, whereas the exponents are viewed as regular (non-negative) integers.

28.2 Degree of a polynomial

In this section, we will introduce the *degree* of a polynomial, which is a measure of how "big" the polynomial is. It's a useful tool that will have numerous applications in our work with polynomials.

Definition 28.5 (Degree of a polynomial)**.** Let $f(x)$ be a non-zero polynomial in $R[x]$, where R is a commutative ring. The *degree of* $f(x)$, denoted $\deg f(x)$, is the highest exponent in $f(x)$.

Remark. The degree of the zero polynomial $0 \in R[x]$ is undefined. You'll see why in an exercise.

Example 28.6. Consider the polynomial $f(x) = 3x^{15} + 4x^3 + 2 \in \mathbb{Z}[x]$. The highest exponent in $f(x)$ is 15, and thus we have $\deg f(x) = 15$.

Example 28.7. Let $f(x) = 3x + 7 \in \mathbb{Z}_{10}[x]$. Rewriting it as $f(x) = 3x^1 + 7x^0$ (since x^0 is defined to be 1), we see that the highest exponent in $f(x)$ is 1. Thus, $\deg f(x) = 1$.

Here is a generalization of Example 28.7.

Example 28.8. Let $f(x) \in R[x]$ be a *linear* polynomial; i.e., $f(x) = a_1x + a_0$ where $a_1 \neq 0$. Then $f(x) = a_1x^1 + a_0x^0$, so that $\deg f(x) = 1$. In other words, all linear polynomials have degree 1.

Example 28.9. Let $f(x) = 7 \in \mathbb{R}[x]$, i.e., a *constant* polynomial. Since $f(x) = 7x^0$, we have $\deg f(x) = 0$. A similar calculation shows that all *non-zero* constant polynomials have degree 0.

Remark. Now that we've seen several examples, let's give a bit more precise formulation of the *degree*. Let $f(x) \in R[x]$ be a *non-zero* polynomial where $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ and $a_n \neq 0$. Then $\deg f(x) = n$ and the coefficient a_n is called the *leading coefficient* of $f(x)$. Since n is a non-negative integer, it follows that $\deg f(x) \geq 0$ as well.

Definition 28.10. A polynomial is said to be *monic* if its leading coefficient equals 1.

Example 28.11. Let $f(x), g(x) \in \mathbb{Z}[x]$, where $f(x) = x^{10} + 7x^3 - 2x + 5$ and $g(x) = 4x^9 - 18x^5 + 3x^4$. Then $f(x)$ is monic since its leading coefficient, i.e., the coefficient of x^{10}, is 1. However, $g(x)$ is not monic, since its leading coefficient, i.e., the coefficient of x^9, is 4.

We've seen that the polynomial ring $R[x]$ is commutative. (Recall that by the definition of $R[x]$, the coefficient ring R must be commutative.) We'll soon see that if R is an integral domain, then so is $R[x]$. However, $R[x]$ is *never* a field, even if the coefficient ring R is a field. In particular, we'll show that the polynomial $x \in R[x]$ does not have a multiplicative inverse in $R[x]$.

Theorem 28.12. *Let R be a commutative ring. Then the polynomial $x \in R[x]$ does not have a multiplicative inverse in $R[x]$. Consequently, $R[x]$ is not a field.*

PROOF. Assume for contradiction that $x \in R[x]$ does have a multiplicative inverse. Then there exists $f(x) \in R[x]$ such that $x \cdot f(x) = 1$ and $f(x) \cdot x = 1$. Note that $f(x)$ is a non-zero polynomial, since if $f(x)$ were 0, we'd have $x \cdot f(x) = x \cdot 0 = 0$ by Theorem 26.9, whereas we know that $x \cdot f(x) = 1$.

Since $f(x)$ is non-zero, $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $a_n \neq 0$ and $n \geq 0$. Thus

$$\begin{aligned} x \cdot f(x) &= x \cdot (a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) \\ &= a_n x^{n+1} + a_{n-1}x^n + \cdots + a_1 x^2 + a_0 x. \end{aligned}$$

Hence, $\deg(x \cdot f(x)) = n + 1 \geq 1$. Then $x \cdot f(x)$ does not equal 1, since all non-zero constant polynomials (such as 1) have degree 0. This contradicts $x \cdot f(x) = 1$, and thus x cannot have a multiplicative inverse. ■

Proof know-how. In the proof above, we used Theorem 26.9 (i.e., $0 \cdot a = 0$ and $a \cdot 0 = 0$ in any ring) to show that $x \cdot 0 = 0$ in $R[x]$. Here, we highlight the fact that *a polynomial ring is a ring*. Thus, any theorem that we've proven about a ring applies to the polynomial ring $R[x]$ as well. (Compare this with the recurring theme of Chapter 23; namely: *A quotient group is a group*.)

Proofs about a polynomial often involve writing out its terms. A *non-zero* polynomial $f(x) \in R[x]$ can be written as $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $a_n \neq 0$ and $n \geq 0$. Then $\deg f(x) = n$.

The next example shows how the degrees of $f(x)$, $g(x)$, and $f(x) \cdot g(x)$ are related.

Example 28.13. Consider the polynomials $f(x) = 3x^{15} + 4x^3 + 2$ and $g(x) = 6x^8 + 5x + 3$.

- In $\mathbb{Z}[x]$, we have $\deg f(x) = 15$ and $\deg g(x) = 8$. Moreover,

$$\begin{aligned} f(x) \cdot g(x) &= (3x^{15}) \cdot (6x^8) + (\text{lower-degree terms}) \\ &= (3 \cdot 6)x^{15+8} + (\text{lower-degree terms}) \\ &= 18x^{23} + (\text{lower-degree terms}) \end{aligned}$$

  so that the degree of $f(x) \cdot g(x)$ is 23, which is the sum of $\deg f(x) = 15$ and $\deg g(x) = 8$.

- In $\mathbb{Z}_7[x]$, we have the same result as in $\mathbb{Z}[x]$. The only difference is that the highest degree term of $f(x) \cdot g(x)$ is $4x^{23}$ instead of $18x^{23}$, because $18 = 4$ in $\mathbb{Z}_7$.

- Now let's view these polynomials in $\mathbb{Z}_9[x]$. We still have $\deg f(x) = 15$ and $\deg g(x) = 8$. However,

$$\begin{aligned} f(x) \cdot g(x) &= (3x^{15} + 4x^3 + 2) \cdot (6x^8 + 5x + 3) \\ &= (\mathbf{3} \cdot \mathbf{6})x^{15+8} + (3 \cdot 5)x^{15+1} + (\text{lower-degree terms}) \\ &= \mathbf{18}x^{23} + 15x^{16} + (\text{lower-degree terms}) \\ &= \mathbf{0}x^{23} + 6x^{16} + (\text{lower-degree terms}) \quad \leftarrow \text{Reduce coefficients modulo 9} \\ &= 6x^{16} + (\text{lower-degree terms}) \end{aligned}$$

  so that the degree of $f(x) \cdot g(x)$ is 16, *not* 23. What happened here? Observe that 3 and 6 are zero divisors in $\mathbb{Z}_9$. Since $3 \cdot 6 = 0$ in $\mathbb{Z}_9$, the leading term in $f(x) \cdot g(x)$, namely $18x^{23}$, vanishes when its coefficient is reduced modulo 9. Therefore, unlike in $\mathbb{Z}[x]$ and $\mathbb{Z}_7[x]$, the degree of $f(x) \cdot g(x)$ does *not* equal the sum of $\deg f(x) = 15$ and $\deg g(x) = 8$.

The above example leads to the following theorem.

Theorem 28.14 (Degree of a product)**.** *Suppose R is an integral domain. Let $f(x)$, $g(x) \in R[x]$ with $f(x)$, $g(x) \neq 0$. Then* $\deg f(x) \cdot g(x) = \deg f(x) + \deg g(x)$.

PROOF. Since $f(x)$ and $g(x)$ are non-zero polynomials, we write

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$

and

$$g(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0,$$

where $a_m, b_n \neq 0$. Hence $\deg f(x) = m$ and $\deg g(x) = n$. The product $f(x) \cdot g(x)$ is given by

$$\begin{aligned} f(x) \cdot g(x) &= (a_m x^m) \cdot (b_n x^n) + (\text{lower-degree terms}) \\ &= (a_m \cdot b_n)x^{m+n} + (\text{lower-degree terms}). \end{aligned}$$

Since R is an integral domain and $a_m, b_n \neq 0$, we have $a_m \cdot b_n \neq 0$ as well. Thus $\deg f(x) \cdot g(x) = m + n$, which equals $\deg f(x) + \deg g(x)$, as desired. ∎

Proof know-how. In the above proof, we used the following: If $a_m, b_n \neq 0$, then $a_m \cdot b_n \neq 0$. This is the contrapositive of the zero product property (Definition 27.7), which must hold as R is an integral domain.

The above proof shows that the product of two non-zero polynomials is also a non-zero polynomial. This confirms our earlier claim, which we state as a theorem here.

Theorem 28.15. *Let R be a commutative ring. If R is an integral domain, then $R[x]$ is an integral domain.*

28.3 Units and zero divisors

Example 28.16. Let's find all *units* in the polynomial ring $\mathbb{Z}_7[x]$, i.e., polynomials $f(x)$ with multiplicative inverse $g(x)$ such that $f(x) \cdot g(x) = 1$. The constant polynomials 1, 2, 3, 4, 5, 6 are units in $\mathbb{Z}_7[x]$, since $1 \cdot 1 = 1,\ 2 \cdot 4 = 1,\ 3 \cdot 5 = 1,$ and $6 \cdot 6 = 1$. Note that these are precisely the units of the coefficient ring $\mathbb{Z}_7$.

Could a *non-constant* polynomial be a unit in $\mathbb{Z}_7[x]$? The answer is "No," and here is the reason why. Let $f(x) \in \mathbb{Z}_7[x]$ be a non-constant polynomial, say with degree 4. Suppose $g(x) \in \mathbb{Z}_7[x]$ is any non-zero polynomial. As the coefficient ring $\mathbb{Z}_7$ is an integral domain (in fact, it's a field), Theorem 28.14 applies. Hence the degree of $f(x) \cdot g(x)$ equals $\deg f(x) + \deg g(x)$, which is at least 4. Then $f(x) \cdot g(x) \neq 1$, since 1 is a non-zero constant polynomial and thus has degree 0. Therefore, $f(x)$ cannot be a unit in $\mathbb{Z}_7[x]$.

Intuitively, this example shows that a non-constant polynomial is too "big" to be a unit in $\mathbb{Z}_7[x]$, or in any polynomial ring $R[x]$ where R is an integral domain.

Example 28.17. Since $\mathbb{Z}$ and $\mathbb{R}$ are both integral domains, Example 28.16 implies the following:

- The only units in $\mathbb{Z}[x]$ are 1 and -1. These are the units of the coefficient ring $\mathbb{Z}$.
- The only units in $\mathbb{R}[x]$ are the non-zero real numbers, i.e., the units of $\mathbb{R}$.

Here is the generalization of the above examples, whose proof is left for you as an exercise.

Theorem 28.18. *Let R be an integral domain. Then the only units in $R[x]$ are the units of R.*

Example 28.19 (Non-example)**.** Consider the polynomial ring $\mathbb{Z}_9[x]$. Here, the coefficient ring $\mathbb{Z}_9$ is *not* an integral domain, because $3 \cdot 6 = 0$ in $\mathbb{Z}_9$ (thus, 3 and 6 are zero divisors). We have

$$(3x + 1) \cdot (6x + 1) = \mathbf{18}x^2 + \mathbf{9}x + 1 = \mathbf{0}x^2 + \mathbf{0}x + 1 = 1,$$

where we reduced the coefficients 18 and 9 in $\mathbb{Z}_9$. Thus $(3x + 1) \cdot (6x + 1) = 1$, so that the non-constant polynomials $3x + 1$ and $6x + 1$ are units in $\mathbb{Z}_9[x]$.

The polynomial rings $\mathbb{Z}_7[x]$, $\mathbb{Z}[x]$, and $\mathbb{R}[x]$ have coefficient rings that are integral domains ($\mathbb{Z}_7$, $\mathbb{Z}$, and $\mathbb{R}$, respectively). Thus, Theorem 28.15 applies and we know that $\mathbb{Z}_7[x]$, $\mathbb{Z}[x]$, and $\mathbb{R}[x]$ are also integral domains. In other words, these polynomial rings do not have zero divisors.

Example 28.20 (Non-example)**.** Consider $\mathbb{Z}_6[x]$, where $\mathbb{Z}_6$ is *not* an integral domain, and non-zero polynomials $4x + 2$ and $3x$ in $\mathbb{Z}_6[x]$. We have

$$(4x + 2) \cdot 3x = \mathbf{12}x^2 + \mathbf{6}x = \mathbf{0}x^2 + \mathbf{0}x = 0,$$

where we reduced the coefficients 12 and 6 in $\mathbb{Z}_6$. Thus $(4x + 2) \cdot 3x = 0$, so that $4x + 2$ and $3x$ are zero divisors in $\mathbb{Z}_6[x]$. Therefore $\mathbb{Z}_6[x]$ is *not* an integral domain.

Polynomial rings that contain non-constant units like $\mathbb{Z}_9[x]$ or zero divisors like $\mathbb{Z}_6[x]$ can be suitable for fun and rich mathematical exploration. However, they're not conducive to producing the type of elegant results that we'll seek to develop in our

study of polynomials. Thus, we want to work with a polynomial ring $R[x]$ where the coefficient ring R is an integral domain. In fact, R being a *field* is even better (we'll see why soon), and thus we'll work with such polynomial rings (e.g., $\mathbb{Z}_7[x]$ and $\mathbb{R}[x]$) most of the time.

Exercises

1. Let $f(x) = 3x^4 - 7x^2 + 4$ and $g(x) = 4x^2 + 1$ be elements of $\mathbb{Z}[x]$.
   (a) Compute $g(x) \cdot f(x)$ using the distributive law repeatedly. (See Example 28.1.)
   (b) How does your product in part (a) compare with the product $f(x) \cdot g(x)$ from Example 28.1?

2. Again, let $f(x) = 3x^4 - 7x^2 + 4$ and $g(x) = 4x^2 + 1$ be elements of $\mathbb{Z}[x]$.
   (a) Compute $g(x) \cdot f(x)$ using the area model. (See Example 28.1.)
   (b) After working on part (a), Elizabeth says, "The rectangles for $f(x) \cdot g(x)$ and $g(x) \cdot f(x)$ are the same, with one just rotated on its side." What might she mean?

3. Let $f(x) = 2x^3 + 2x^2 + 1$ and $g(x) = x^2 + 2x + 2$ be elements of $\mathbb{Z}_3[x]$.
   (a) Compute $f(x) + g(x)$ and $g(x) + f(x)$. How do the sums compare?
   (b) Compute $f(x) \cdot g(x)$ and $g(x) \cdot f(x)$. How do the products compare?
   (c) Find $-f(x)$ and $-g(x)$, i.e., the additive inverses of $f(x)$ and $g(x)$, respectively.

4. Repeat Exercise #3, but with $f(x) = 4x^{10} + 3x^5 + 2$ and $g(x) = 5x^8 + 6x^4 + 3x^2$ in $\mathbb{Z}_7[x]$.

5. Let $f(x)$, $g(x)$, $h(x) \in R[x]$ where R is a commutative ring. In Example 28.1, we saw how to use the area model to compute the product $f(x) \cdot g(x)$.
   (a) Describe how to use the area model to compute $f(x) \cdot (g(x) + h(x))$.
   (b) Describe how to use the area model to compute $f(x) \cdot g(x) + f(x) \cdot h(x)$.
   (c) Use your answers in parts (a) and (b) to explain why the distributive law must hold in $R[x]$.

6. Proceed as in Exercise #5, but this time, use the area model to explain why the associative law for multiplication, i.e., $(f(x) \cdot g(x)) \cdot h(x) = f(x) \cdot (g(x) \cdot h(x))$, must hold in $R[x]$.

   Hint: Instead of a rectangle, what geometric shape must you create?

7. Let $f(x), g(x) \in \mathbb{Z}[x]$ where
   $$f(x) = a_5x^5 + a_4x^4 + \cdots + a_1x + a_0$$
   and
   $$g(x) = b_5x^5 + b_4x^4 + \cdots + b_1x + b_0.$$
   Write an expression for the coefficient of x^5 in the product $f(x) \cdot g(x)$.

8. (a) Find examples of a monic, non-constant polynomial.
   (b) What is a monic, constant polynomial?

9. Let $f(x) \in R[x]$ where R is a commutative ring (but not necessarily an integral domain). Prove each statement.
   (a) If $f(x)$ is monic and non-constant, then $f(x)$ is *not* a unit in $R[x]$.
   (b) If $f(x)$ is monic, then $f(x)$ is *not* a zero divisor in $R[x]$.

10. (a) Find the number of non-zero polynomials in $\mathbb{Z}_7[x]$ of degree 5.
    (b) Find the number of non-zero polynomials in $\mathbb{Z}_7[x]$ of degree less than or equal to 5.
    (c) Find the number of monic polynomials in $\mathbb{Z}_7[x]$ of degree less than or equal to 5.

11. Find all units in $\mathbb{Z}_{101}[x]$. Explain your reasoning.

12. Find all zero divisors in $\mathbb{Z}_{101}[x]$. Explain your reasoning.

13. (a) Find five units (other than 1 and 3) in $\mathbb{Z}_4[x]$.
    (b) Explain why $\mathbb{Z}_4[x]$ has infinitely many units.

14. (a) Find five zero divisors (other than 2) in $\mathbb{Z}_4[x]$.
    (b) Explain why $\mathbb{Z}_4[x]$ has infinitely many zero divisors.

15. Find non-zero polynomials $f(x), g(x) \in \mathbb{Z}_{10}[x]$ that satisfy each condition below, or explain why it's not possible to satisfy the condition.
    (a) $\deg f(x) \cdot g(x) = \deg f(x) + \deg g(x)$.
    (b) $\deg f(x) \cdot g(x) < \deg f(x) + \deg g(x)$.
    (c) $\deg f(x) \cdot g(x) > \deg f(x) + \deg g(x)$.

16. Our friends are discussing the degree of the constant polynomial 0.

    Anita: "Why can't we say $\deg(0) = 0$? The zero polynomial is a constant, right?"
    Elizabeth: "But Theorem 28.14 fails if $\deg(0) = 0$."

    What might Elizabeth mean?

17. (a) Find $p(x), q(x) \in \mathbb{Z}_{10}[x]$, both with degree 1, such that $p(x) \cdot q(x) = x + 7$.
    (b) What if $p(x)$ and $q(x)$ must each have degree greater than 1? Do such polynomials exist in $\mathbb{Z}_{10}[x]$? If so, find them. If not, explain why not.

18. The converse of Theorem 28.15 states: If $R[x]$ is an integral domain, then R is an integral domain. (Here, R is a commutative ring.) Determine if this converse is true or false. If it's true, prove it. If it's false, provide a counterexample.

19. Prove Theorem 28.18.

20. Let $f(x) = 5x + 1 \in \mathbb{Z}_{10}[x]$. Determine whether or not $f(x)$ is a unit in $\mathbb{Z}_{10}[x]$. Explain your reasoning.

21. The converse of Theorem 28.18 states: If the only units in $R[x]$ are the units of R, then R is an integral domain. (Here, R is a commutative ring.) Determine if this converse is true or false. If it's true, prove it. If it's false, provide a counterexample.

22. **(Challenge)**
    (a) Find *all* units in $\mathbb{Z}_9[x]$.
    (b) Find *all* zero divisors in $\mathbb{Z}_6[x]$.

29

Polynomial Rings, Part II

Chapter 28 introduced *polynomial rings*, in which polynomials such as $f(x) = 3x^4 - 5x^3 - 8$ are treated as *elements* of a ring. We added and multiplied polynomials, just as we do with integers. But we can also *evaluate* polynomials by substituting values into them; i.e., treat them like *functions* as we do in calculus. For instance, by setting $x = 2$, we obtain $f(2) = 3 \cdot 2^4 - 5 \cdot 2^3 - 8 = 0$. What's more, the result of this evaluation can reveal important information about the polynomial $f(x)$.

It turns out that the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$ where F is a field (e.g., $\mathbb{R}[x]$) have many *structural* similarities. For instance, both rings satisfy the *division algorithm*. When dividing 5,273 by 6, we get a remainder of 5, because $5{,}273 = 6 \cdot 878 + 5$. Note that the remainder is less than the divisor and also is non-negative (i.e., $0 \leq 5 < 6$). We'll see below that $F[x]$ satisfies an analogous division algorithm as well. Starting with this chapter and continuing through the rest of the book, each chapter ends with a short section called "Big picture stuff" where we discuss the similarities between these two rings.

29.1 Division algorithm in $F[x]$

In Chapter 12, we studied the notion of a *remainder* that we obtain when dividing an integer by another integer. Here is an example.

Example 29.1. When dividing 5,273 by 6, the remainder is 5 since $5{,}273 = 6 \cdot 878 + 5$. We also have $5{,}273 = 6 \cdot 877 + 11$ and $5{,}273 = 6 \cdot 879 + (-1)$. But we don't say that the remainder is 11 or -1. This is because the remainder r must be less than the divisor (i.e., $r < 6$) and also non-negative (i.e., $r \geq 0$). Thus, $r = 5$ is the only possible remainder.

We recall the theorem that generalizes the above example.

Theorem 12.16 (Division algorithm in $\mathbb{Z}$)**.** *Let a and b be integers, with $b > 0$. Then there exist $q, r \in \mathbb{Z}$ such that $a = b \cdot q + r$ with $0 \leq r < b$.*

The division algorithm in $\mathbb{Z}$ is a useful theorem, particularly for showing that an integer is a divisor of another integer. We will soon identify an analogous theorem for polynomial rings. But first, let's delve a bit deeper into finding remainders in $\mathbb{Z}$.

Example 29.2. To find the quotient and remainder when dividing 5,273 by 6, we use a technique called "long division," as shown below.

$$\begin{array}{rl}
 & \boxed{878} \longleftarrow \text{quotient} \\
6\,\overline{)\,5273} & \\
\underline{-\,48} & \\
47 & \\
\underline{-\,42} & \\
53 & \\
\underline{-\,48} & \\
\boxed{5} & \longleftarrow \text{remainder}
\end{array}$$

Therefore, $5{,}273 = 6 \cdot 878 + 5$. Observe that the remainder 5 is non-negative and less than the divisor 6, as required by the division algorithm in $\mathbb{Z}$.

Example 29.3. Consider $f(x) = 5x^4 + x^3 - 3x^2 + 4x - 3$ and $g(x) = x^2 + 1$ in $\mathbb{R}[x]$, where the coefficient ring $\mathbb{R}$ is a field. We use long division to find the quotient and remainder when dividing $f(x)$ by $g(x)$:

$$\begin{array}{rl}
 & \boxed{5x^2 + x - 8} \longleftarrow \text{quotient} \\
x^2 + 1\,\overline{)\,5x^4 + x^3 - 3x^2 + 4x - 3} & \\
\underline{-\,(5x^4 \qquad + 5x^2)} & \\
x^3 - 8x^2 + 4x & \\
\underline{-\,(x^3 \qquad + \ x)} & \\
-\,8x^2 + 3x - 3 & \\
\underline{-\,(-\,8x^2 \qquad - 8)} & \\
\boxed{3x + 5} & \longleftarrow \text{remainder}
\end{array}$$

Therefore, $f(x) = (x^2 + 1) \cdot q(x) + (3x + 5)$, where $q(x) = 5x^2 + x - 8$ is the quotient. Note how the degree of the remainder $r(x) = 3x+5$ is less than the degree of the divisor $g(x) = x^2 + 1$.

Example 29.4. Consider $f(x) = 4x^3 + 5x^2 + 2$ and $g(x) = 3x^2 + 5$ in $\mathbb{Z}_7[x]$, where $\mathbb{Z}_7$ is a field. We divide $f(x)$ by $g(x)$ using long division, emphasizing that the coefficients are computed in $\mathbb{Z}_7$:

$$
\begin{array}{rl}
 & \boxed{6x+4} \longleftarrow \text{quotient} \\
3x^2+5 \;\overline{)\; 4x^3+5x^2 \qquad +2} & \\
\quad -(4x^3 \qquad +2x) & \\
\hline
\qquad\qquad 5x^2+5x+2 & \\
\quad -(5x^2 \qquad +6) & \\
\hline
\qquad\qquad \boxed{5x+3} \longleftarrow \text{remainder} &
\end{array}
$$

Therefore, $f(x) = (3x^2 + 5) \cdot q(x) + (5x + 3)$, where $q(x) = 6x + 4$ is the quotient. Note how the degree of the remainder $r(x) = 5x + 3$ is less than the degree of the divisor $g(x) = 3x^2 + 5$.

Just as for its counterpart in $\mathbb{Z}$, we'll assume the following theorem without proof.

Theorem 29.5 (Division algorithm in $F[x]$)**.** *Let F be a field. Suppose $f(x), g(x) \in F[x]$ with $g(x) \neq 0$. Then there exist $q(x), r(x) \in F[x]$ such that $f(x) = g(x) \cdot q(x) + r(x)$ with $r(x) = 0$ or $\deg r(x) < \deg g(x)$.*

Remark. Since the degree of a non-zero polynomial is always non-negative, we do not need to specify that $\deg r(x) \geq 0$ in the above theorem. However, the degree of the zero polynomial $0 \in F[x]$ is undefined, so we must write "$r(x) = 0$ or $\deg r(x) < \deg g(x)$." In any case, the key takeaway from the theorem is that the degree of the remainder is smaller than the degree of the divisor.

To see why the coefficients must be in a field, let's see how Theorem 29.5 fails in $\mathbb{Z}[x]$. Recall that $\mathbb{Z}$ is an integral domain but not a field, since not every non-zero integer has a multiplicative inverse.

Example 29.6 (Non-example)**.** Let $f(x) = x^3$ and $g(x) = 2x$ in $\mathbb{Z}[x]$. The division algorithm would imply that there exist $q(x), r(x) \in \mathbb{Z}[x]$ such that $x^3 = 2x \cdot q(x) + r(x)$, with either $r(x) = 0$ or $\deg r(x) < \deg g(x)$. Since $\deg g(x) = 1$, we know that $r(x)$ is a constant polynomial, possibly 0. So, let $r(x) = m$ for some $m \in \mathbb{Z}$. For the equation $x^3 = 2x \cdot q(x) + r(x)$ to be true, $q(x)$ must be a quadratic; i.e., $q(x) = ax^2 + bx + c$ for some integers a, b, c. Thus, we have $x^3 = 2x \cdot (ax^2 + bx + c) + m$, so that $x^3 = 2ax^3 + 2bx^2 + 2cx + m$. Matching the coefficients of x^3, we get $1 = 2a$, which cannot occur in $\mathbb{Z}$.

29.2 Factor theorem

Example 29.7. Let $f(x) = 3x^4 - 5x^3 - 8 \in \mathbb{R}[x]$. By setting $x = 2$, we obtain $f(2) = 3 \cdot 2^4 - 5 \cdot 2^3 - 8 = 0$. As we'll soon see, the fact that $f(2) = 0$ implies that $x - 2$ is a factor of $f(x)$. In an exercise, you will find $q(x) \in \mathbb{R}[x]$ such that $f(x) = (x - 2) \cdot q(x)$.

Note that the following mean the same thing:

- $f(x)$ is a multiple of $x - 2$.
- $x - 2$ is a factor of $f(x)$.

We write $(x - 2) \mid f(x)$ as a shorthand for "$x - 2$ is a factor of $f(x)$." Here is the generalization.

Definition 29.8. Let F be a field. Suppose $f(x), g(x) \in F[x]$. We say that $g(x)$ is a *factor* of $f(x)$ when $f(x) = g(x) \cdot q(x)$ for some $q(x) \in F[x]$. We write $g(x) \mid f(x)$ to mean "$g(x)$ is a factor of $f(x)$."

Below is a theorem that generalizes Example 29.7.

Theorem 29.9 (Factor theorem)**.** *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $f(a) = 0$ if and only if $(x - a) \mid f(x)$.*

Before proving this "if and only if" theorem, we examine one implication: If $(x - a) \mid f(x)$, then $f(a) = 0$. For instance, suppose $f(x) = (x - 2) \cdot q(x)$, which is an equality of two polynomials $f(x)$ and $(x-2) \cdot q(x)$ in $\mathbb{R}[x]$. Setting $x = 2$ on both sides of the equation, we get $f(2) = (2-2) \cdot q(2)$, which equates two real numbers $f(2)$ and $(2 - 2) \cdot q(2)$ in $\mathbb{R}$. (Here, be careful *not* to write $f(2) = (2 - 2) \cdot q(x)$; i.e., don't forget to set $x = 2$ in the polynomial $q(x)$ as well.) The factor theorem is one example of the powerful interplay between $\mathbb{R}[x]$ and $\mathbb{R}$ and, more generally, between $F[x]$ and F.

PROOF. We must prove two implications:

- If $f(a) = 0$, then $(x - a) \mid f(x)$.
- If $(x - a) \mid f(x)$, then $f(a) = 0$.

We will prove the first implication. The proof of the second implication is left for you as an exercise.

Assume $f(a) = 0$. By the division algorithm, there exist $q(x), r(x) \in F[x]$ such that

$$f(x) = (x - a) \cdot q(x) + r(x)$$

with $r(x) = 0$ or $\deg r(x) < \deg(x - a)$. Since $\deg(x - a) = 1$, we conclude that $r(x)$ is a constant polynomial, possibly 0. Let $r(x) = \alpha$ for some $\alpha \in F$, so that $f(x) = (x-a) \cdot q(x) + \alpha$. Solving for α, we obtain $\alpha = f(x) - (x-a) \cdot q(x)$. Substituting $x = a$ and recalling that $f(a) = 0$, we obtain

$$\alpha = f(a) - (a - a) \cdot q(a) = 0 - 0 \cdot q(a) = 0.$$

Thus $\alpha = 0$, which implies $f(x) = (x - a) \cdot q(x)$ and hence $(x - a) \mid f(x)$. ∎

Remark. In the proof above, we have the equation $\alpha = f(x)-(x-a)\cdot q(x)$. But α is an element of F, while $f(x)-(x-a)\cdot q(x)$ is a polynomial in $F[x]$. How can they be equal? When expanding the polynomial $f(x)-(x-a)\cdot q(x)$, all of the x's cancel with each other and we're left with a constant that equals α. Thus, the equation $\alpha = f(x)-(x-a)\cdot q(x)$ is telling us that the right-hand side is also a constant polynomial (though perhaps not as easily detectable). Since all the x's cancel anyway, we can set it to any value that we like *without* affecting α. And setting $x = a$ allowed us to conclude that $\alpha = 0$.

Proof know-how. For the factor theorem, we prove $(x-a) \mid f(x)$ by writing $f(x) = (x-a)\cdot q(x) + r(x)$ and showing $r(x) = 0$. More generally, to prove that $g(x) \mid f(x)$, first write $f(x) = g(x)\cdot q(x) + r(x)$ with $r(x) = 0$ or $\deg r(x) < \deg g(x)$. Then show that $r(x) = 0$. (Compare this with the Proof know-how following the second proof of Theorem 12.17.)

Example 29.10. Consider $f(x) = 5x^{672} + 2x^{359} + 4x^{101} + x^{77} + 3x^{23} + 6$ in $\mathbb{Z}_7[x]$. We have

$$f(1) = 5\cdot 1^{672} + 2\cdot 1^{359} + 4\cdot 1^{101} + 1^{77} + 3\cdot 1^{23} + 6 = 5 + 2 + 4 + 1 + 3 + 6 = 21 = 0,$$

where the equality $21 = 0$ occurs in $\mathbb{Z}_7$. Thus, $f(1) = 0$ so that $x - 1$ is a factor of $f(x)$.

Let's also find $f(-1)$ by setting $x = -1$ in the polynomial $f(x)$. Since $(-1)^k = 1$ when k is even and $(-1)^k = -1$ when k is odd, we obtain

$$\begin{aligned} f(-1) &= 5\cdot(-1)^{672} + 2\cdot(-1)^{359} + 4\cdot(-1)^{101} + (-1)^{77} + 3\cdot(-1)^{23} + 6 \\ &= 5\cdot 1 + 2\cdot(-1) + 4\cdot(-1) + (-1) + 3\cdot(-1) + 6 \\ &= 5 - 2 - 4 - 1 - 3 + 6 = 1 \end{aligned}$$

so that $f(-1) = 1$. Since $f(-1) \neq 0$, the factor theorem implies that $x - (-1)$, or equivalently $x + 1$, is *not* a factor of the polynomial $f(x)$.

Example 29.11. Consider $f(x) = 5x^{451} + 11x^{274} + 1$ in $\mathbb{Z}_{13}[x]$. Let's find the remainder $r(x)$ when $f(x)$ is divided by $x - 1$. (Performing long division is *not* recommended!) By the division algorithm, there exist $q(x), r(x) \in \mathbb{Z}_{13}[x]$ such that $f(x) = (x-1)\cdot q(x) + r(x)$ with $r(x) = 0$ or $\deg r(x) < \deg(x-1)$. Since $\deg(x-1) = 1$, we note that $r(x)$ is a constant polynomial, possibly 0. Let $r(x) = \alpha$ for some $\alpha \in \mathbb{Z}_{13}$, so that $\alpha = f(x) - (x-1)\cdot q(x)$. Setting $x = 1$, we obtain $\alpha = f(1) - (1-1)\cdot q(1) = f(1) - 0\cdot q(1) = f(1)$. Thus, $\alpha = f(1) = 5\cdot 1^{451} + 11\cdot 1^{274} + 1 = 5 + 11 + 1 = 17 = 4$, where the equality $17 = 4$ occurs in $\mathbb{Z}_{13}$. Therefore, the remainder is the constant polynomial $r(x) = \alpha = 4$.

Example 29.11 motivates the following theorem, whose proof is left for you as an exercise.

Theorem 29.12 (Remainder theorem)**.** *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $f(a) \in F$ is the remainder when $f(x)$ is divided by $x - a$.*

29.3 Nilpotent elements

Example 29.13. Consider $3 \in \mathbb{Z}_{81}$. We have

$$3^1 = 3,\ \ 3^2 = 9,\ \ 3^3 = 27,\ \ \mathbf{3^4 = 0}.$$

After that, we have

$$\begin{aligned} 3^5 &= \mathbf{3^4}\cdot 3 = \mathbf{0}\cdot 3 = 0, \\ 3^6 &= \mathbf{3^5}\cdot 3 = \mathbf{0}\cdot 3 = 0, \\ 3^7 &= \mathbf{3^6}\cdot 3 = \mathbf{0}\cdot 3 = 0, \\ &\vdots \end{aligned}$$

Thus, we conclude that $3^n = 0$ for all positive integers $n \geq 4$.

Example 29.13 motivates the following definition.

Definition 29.14 (Nilpotent)**.** A ring element r is said to be *nilpotent* if $r^n = 0$ for some positive integer n.

Example 29.15. In Example 29.13, we saw that $3 \in \mathbb{Z}_{81}$ is nilpotent. In $\mathbb{Z}_{81}$, we also have $0^1 = 0$, $6^4 = 0$, and $9^2 = 0$, so that 0, 6, and 9 are nilpotent in $\mathbb{Z}_{81}$. In an exercise at the end of the chapter, you'll find all other nilpotent elements in $\mathbb{Z}_{81}$.

Example 29.16. In any ring, $0^1 = 0$. Thus, the additive identity 0 is nilpotent.

Example 29.17. In $\mathbb{Z}$, the only nilpotent element is 0. For $a \neq 0$, we have $a^n \neq 0$ for all positive integers n.

Example 29.17 illustrates the following theorem, whose proof is left for you as an exercise.

Theorem 29.18. *Let R be an integral domain. Then the only nilpotent element of R is* 0.

Is the converse of this theorem true? In other words, if the only nilpotent element of R is 0, must R be an integral domain? The answer is "No," as shown in the next example.

Example 29.19. Consider the following calculations with the non-zero elements of $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$.

- $1^n = 1$ for all positive integers n.
- $2^1 = 2,\ 2^2 = 4,\ 2^3 = 2,\ 2^4 = 4, \ldots$. Thus, $2^n = 2$ or 4 for all $n \geq 1$.
- $3^1 = 3,\ 3^2 = 3,\ 3^3 = 3,\ 3^4 = 3, \ldots$. Thus, $3^n = 3$ for all $n \geq 1$.
- $4^1 = 4,\ 4^2 = 4,\ 4^3 = 4,\ 4^4 = 4, \ldots$. Thus, $4^n = 4$ for all $n \geq 1$.
- $5^1 = 5,\ 5^2 = 1,\ 5^3 = 5,\ 5^4 = 1, \ldots$. Thus, $5^n = 5$ or 1 for all $n \geq 1$.

Hence, $\mathbb{Z}_6$ does not have any non-zero nilpotent element, although 3 and 4 are idempotents.

Example 29.20. Let $f(x) = 3x$ in $\mathbb{Z}_9[x]$. Then $f(x)$ is nilpotent, since $(3x)^2 = \mathbf{9}x^2 = \mathbf{0}x^2 = 0$ in $\mathbb{Z}_9[x]$. Moreover, $1 - 3x$ is a unit in $\mathbb{Z}_9[x]$, since

$$(1 - 3x) \cdot (1 + 3x) = 1 \cdot 1 + 1 \cdot 3x - 3x \cdot 1 - (3x)^2 = 1 - \mathbf{9}x^2 = 1 - \mathbf{0}x^2 = 1,$$

so that $(1 - 3x) \cdot (1 + 3x) = 1$.

Next, we view $f(x) = 3x$ as an element of $\mathbb{Z}_{27}[x]$. Then $f(x)$ is nilpotent, since $(3x)^3 = \mathbf{27}x^3 = \mathbf{0}x^3 = 0$ in $\mathbb{Z}_{27}[x]$. And as before, $1 - 3x$ is a unit in $\mathbb{Z}_{27}[x]$, since

$$(1-3x)\cdot(1+3x+9x^2) = \text{ (we'll leave the calculations to you) } = 1-\mathbf{27}x^3 = 1-\mathbf{0}x^3 = 1,$$

so that $(1-3x)\cdot(1+3x+9x^2) = 1$. In an exercise, you'll show that the same conclusions can be made (i.e., $3x$ is nilpotent and $1 - 3x$ is a unit) when working in $\mathbb{Z}_{81}[x]$ and in $\mathbb{Z}_m[x]$ where m is a power of 3.

Example 29.20 motivates the theorem below, whose proof is left for you. (Compare with Theorem 27.30.)

Theorem 29.21. *In a ring, if a is nilpotent, then $1 - a$ is a unit.*

Big picture stuff

An important underlying theme for the next several chapters is the myriad *structural* similarities between the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$, where F is a field. In this chapter, we saw the following:

- $F[x]$ has the *division algorithm* just as $\mathbb{Z}$ does.
- The notion of a *factor* in $F[x]$ is analogous to the notion of a *divisor* in $\mathbb{Z}$.

We will continue to identify these similarities as we encounter them. Stay tuned!

Exercises

1. (a) Use long division to find the quotient q and remainder r when dividing 5,696 by 7.
   (b) Verify that your result in part (a) satisfies the division algorithm in $\mathbb{Z}$.
2. Consider $f(x) = x^3 + 2x + 1$ and $g(x) = 2x + 3$ in $\mathbb{Z}_5[x]$.
   (a) Use long division to find the quotient $q(x)$ and remainder $r(x)$ when dividing $f(x)$ by $g(x)$. Keep in mind that the coefficients are in $\mathbb{Z}_5$.
   (b) Verify that your result in part (a) satisfies the division algorithm in $F[x]$.
3. Consider again $f(x) = x^3 + 2x + 1$ and $g(x) = 2x + 3$ in $\mathbb{Z}_5[x]$.
   (a) Find the quotient $q(x)$ and remainder $r(x)$ when dividing $g(x)$ by $f(x)$.
   Note: This is slightly different from Exercise #2. Long division is *not* needed!
   (b) Verify that your result in part (a) satisfies the division algorithm in $F[x]$.
4. Consider $f(x) = x$ and $g(x) = 2x + 1$ in $\mathbb{Z}_5[x]$.
   (a) Find the quotient $q(x)$ and remainder $r(x)$ when dividing $f(x)$ by $g(x)$.
   (b) Verify that your result in part (a) satisfies the division algorithm in $F[x]$.
5. Consider $f(x) = 3x^3 + 10x^2 + 5x - 4$ and $g(x) = x^2 + 2x - 1$ in $\mathbb{R}[x]$. Verify that $g(x)$ is a factor of $f(x)$.
6. Let $f(x) = 3x^4 - 5x^3 - 8 \in \mathbb{R}[x]$. Find $q(x) \in \mathbb{R}[x]$ such that $f(x) = (x-2) \cdot q(x)$. (See Example 29.7.)
7. Let $f(x) = x^4 - 3x^3 + x^2 + 4x - 1 \in \mathbb{R}[x]$. Determine which of $x - 1$, $x + 1$, and $x - 2$ is a factor of $f(x)$.
8. Let $f(x) = 5x^{672} + 2x^{359} + 4x^{101} + x^{77} + 3x^{23} + 6$ in $\mathbb{Z}_7[x]$. In Example 29.10, we saw that $x - 1$ is a factor of $f(x)$, but $x - (-1)$ is *not* a factor of $f(x)$. Determine whether or not each of $x - 2$, $x - 3$, $x - 4$, and $x - 5$ is a factor of $f(x)$. (Note that $x - (-1)$ is equivalent to $x - 6$ in $\mathbb{Z}_7[x]$.)
9. Consider $f(x) = 3x^{458} + 2x^{103} + 4$ in $\mathbb{Z}_5[x]$.
   (a) Find the remainder when $f(x)$ is divided by $x - 1$.
   (b) Find the remainder when $f(x)$ is divided by $x - 2$.
   (c) Find the remainder when $f(x)$ is divided by $x + 1$.

10. In $\mathbb{Z}_{13}[x]$, define the polynomials

$$f(x) = 3x^4 + 10x^3 + 6x^2 + 8x + 11$$

and

$$g(x) = 11x^4 + 8x^3 + 6x^2 + 10x + 3.$$

(a) Describe how the polynomials $f(x)$ and $g(x)$ are related.
(b) Verify that $f(4) = 0$ and $g(10) = 0$.
(c) Verify that $f(11) = 0$ and $g(6) = 0$.
(d) How are 4 and 10 related in $\mathbb{Z}_{13}$?
(e) How are 11 and 6 related in $\mathbb{Z}_{13}$?
(f) What conjecture do you have?

11. Consider $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $g(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$ in $F[x]$ where F is a field. Let α be a non-zero element of F. Show that if $f(\alpha) = 0$, then $g(\alpha^{-1}) = 0$.

12. Complete the proof of Theorem 29.9 by proving its second implication.

13. Prove Theorem 29.12.

14. **Prove:** Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $f(a) = 0$ and $f'(a) = 0$ if and only if $f(x) = (x - a)^2 \cdot q(x)$ for some $q(x) \in F[x]$.

Note: Here, $f'(x)$ refers to the derivative of $f(x)$.

Hint: The product rule from calculus says $(p(x)\cdot q(x))' = p'(x)\cdot q(x) + p(x)\cdot q'(x)$.

15. Recall that a ring element r is said to be *nilpotent* if $r^n = 0$ for some positive integer n.

(a) Find all nilpotent elements of $\mathbb{Z}_9$.
(b) Find all nilpotent elements of $\mathbb{Z}_{10}$.
(c) Find all nilpotent elements of $\mathbb{Z}_{15}$.
(d) Find all nilpotent elements of $\mathbb{Z}_{18}$.
(e) Find all nilpotent elements of $\mathbb{Z}_{81}$. (See Example 29.15.)
(f) Any conjectures about which $\mathbb{Z}_m$ has non-zero nilpotent elements?

16. **Prove:** $\mathbb{Z}_m$ has a non-zero nilpotent element if and only if m is divisible by a square of a prime.

Note: For example, $m = 18$ is divisible by $9 = 3^2$, which is a square of a prime $p = 3$.

17. **Prove:** If a is a nilpotent element in a ring, then $a^n = 0$ for all but finitely many positive integers n.

18. (a) Find five nilpotent elements (other than 0 and 2) in $\mathbb{Z}_4[x]$.
(b) Explain why $\mathbb{Z}_4[x]$ has infinitely many nilpotent elements.

19. (a) Categorize each non-zero element of $\mathbb{Z}_{12}$ as a unit, a zero divisor, or a nilpotent element.
    Note: Be careful. These categories are not mutually exclusive.
    (b) Repeat part (a) with $\mathbb{Z}_{18}$.
    (c) Repeat part (a) with $\mathbb{Z}_{20}$.
    (d) What conjecture do you have?

20. (a) **Prove:** Let $\alpha \in R$ where R is a ring. If α is a non-zero nilpotent element, then α is a zero divisor.
    (b) How about the converse of the statement in part (a)? If it's true, prove it. If it's false, provide a counterexample.

21. Prove Theorem 29.18.

22. In this exercise, you'll continue the work that we began in Example 29.20.
    (a) In $\mathbb{Z}_{27}$, verify that $(1 - 3x) \cdot (1 + 3x + 9x^2) = 1$.
    (b) In $\mathbb{Z}_{81}$, verify that $3x$ and $1-3x$ are a nilpotent element and a unit, respectively.
    (c) Repeat part (b), but in $\mathbb{Z}_m[x]$ where m is a power of 3.

23. Prove Theorem 29.21.

24. **Prove:** If a ring element α is both a nilpotent element and an idempotent, then $\alpha = 0$.

25. Let R be a ring with the following property: If $\alpha^2 = 0$, then $\alpha = 0$ (where $\alpha \in R$). Show that 0 is the only nilpotent element of R.
    Hint: Suppose $a \in R$ is nilpotent and show that $a = 0$.

26. Let α and β be nilpotent elements in a *commutative* ring.
    (a) Prove that $\alpha \cdot \beta$ is nilpotent.
    (b) **(Challenge)** Prove that $\alpha + \beta$ is nilpotent.

30

Factoring Polynomials

From here on, we will work (mostly) in polynomial rings $F[x]$ where F is a field. Examples include $\mathbb{R}[x]$ and $\mathbb{Z}_7[x]$, but *not* $\mathbb{Z}_9[x]$ or $\mathbb{Z}[x]$. As discussed in Chapter 29, there are many structural similarities between $F[x]$ and the ring of integers $\mathbb{Z}$. In this chapter, we will encounter another such similarity by studying the notion of *factoring*. For instance, consider the polynomial $f(x) = x^2 - 6x + 8$ in $\mathbb{R}[x]$. We can factor $f(x)$ as a product of two "smaller" polynomials, i.e., polynomials of lower degree, like this: $f(x) = (x-2)\cdot(x-4)$. Analogously, we can factor an integer, say 15, into a product of two smaller integers: $15 = 3 \cdot 5$. And just as 3 and 5 are prime numbers in $\mathbb{Z}$, the factors $x-2$ and $x-4$ are what we call *unfactorable* in $\mathbb{R}[x]$.

It turns out that factoring an integer or a polynomial, especially if it's large, is *very* difficult to do. This is a good thing, since the difficulty of factoring large integers is what keeps the internet secure through encryption methods such as the *RSA algorithm*. For small enough polynomials, there are techniques to determine if and how they can be factored, which will be the focus of this chapter.

30.1 Examples and definition

Example 30.1. We might say that the polynomial x^2+1 is *unfactorable*; i.e., it cannot be factored. Indeed, it's true that x^2+1 is unfactorable in $\mathbb{R}[x]$. But if we work in $\mathbb{Z}_5[x]$, then

$$(x+2)\cdot(x+3) = x^2 + 5x + 6 = x^2 + 0x + 1 = x^2 + 1,$$

as we reduce the coefficients $5 = 0$ and $6 = 1$ in $\mathbb{Z}_5$. Thus, $x^2+1 = (x+2)\cdot(x+3)$ is factorable in $\mathbb{Z}_5[x]$. The lesson here is that we need to specify the polynomial ring in which we're working (e.g., $\mathbb{R}[x]$ or $\mathbb{Z}_5[x]$).

Example 30.2. We might say that x^2+1 is actually factorable in $\mathbb{R}[x]$, because $x^2+1 = 3\cdot\left(\frac{1}{3}x^2 + \frac{1}{3}\right)$. But this is not a legitimate factorization. When we factor a polynomial, we must write it as a product of two "smaller" polynomials, i.e., polynomials of lower degree. However, $\frac{1}{3}x^2 + \frac{1}{3}$ is *not* smaller than x^2+1, because they both have degree 2.

These examples motivate the following definition. Note how factorable and unfactorable polynomials are analogous to composite and prime integers.

Definition 30.3 (Factorable/unfactorable polynomials)**.** Let F be a field. Suppose $f(x) \in F[x]$ with $\deg f(x) \geq 1$; i.e., $f(x)$ is *not* a constant polynomial.

- We say that $f(x)$ is *factorable* in $F[x]$ when $f(x) = p(x) \cdot q(x)$ for some $p(x), q(x) \in F[x]$ with $\deg p(x), \deg q(x) < \deg f(x)$.

- Otherwise, we say that $f(x)$ is *unfactorable* in $F[x]$.

We make a couple of observations about the above definition of *factorable*. First, the degree requirement $\deg p(x), \deg q(x) < \deg f(x)$ ensures a legitimate factorization of $f(x)$ where the factors $p(x)$ and $q(x)$ are "smaller" than $f(x)$. (See Example 30.2.) Second, you'll show in an exercise that $\deg p(x)$ and $\deg q(x)$ are greater than 0, and thus the factors $p(x)$ and $q(x)$ must be non-constant polynomials.

Remark. In mathematics, factorable and unfactorable polynomials are more commonly referred to as *reducible* and *irreducible*, respectively. In this textbook, however, we will use the terms "factorable" and "unfactorable," since these are likely more familiar to you from your prior experiences with polynomials.

Example 30.4. Let $f(x) = x^2 + 1$ in $\mathbb{Z}_5[x]$. In Example 30.1, we found $x^2 + 1 = (x+2) \cdot (x+3)$, and so $f(x)$ is factorable in $\mathbb{Z}_5[x]$. Note that $\deg f(x) = 2$, while its factors $x+2$ and $x+3$ have degree 1. Thus, the degree requirement $\deg(x+2), \deg(x+3) < \deg f(x)$ is satisfied.

Example 30.5. Let $f(x) = x^2 - 2$ in $\mathbb{R}[x]$. We have $x^2 - 2 = (x + \sqrt{2}) \cdot (x - \sqrt{2})$ where $x + \sqrt{2}, x - \sqrt{2} \in \mathbb{R}[x]$. Since $\deg f(x) = 2$ and its factors $x + \sqrt{2}$, $x - \sqrt{2}$ have degree 1, the degree requirement is satisfied. Thus we conclude that $f(x)$ is factorable in $\mathbb{R}[x]$.

Example 30.6. Let $f(x) = x^2 + 1$ in $\mathbb{R}[x]$. We will show that $f(x)$ is unfactorable in $\mathbb{R}[x]$. Assume for contradiction that $f(x)$ is factorable in $\mathbb{R}[x]$. Thus $f(x) = p(x) \cdot q(x)$ for some $p(x), q(x) \in \mathbb{R}[x]$ with $\deg p(x), \deg q(x) < \deg f(x)$. Since $\deg f(x) = \deg p(x) + \deg q(x)$ (Theorem 28.14) and $\deg f(x) = 2$, we must have $\deg p(x) = 1$ and $\deg q(x) = 1$; i.e., $p(x)$ and $q(x)$ are linear polynomials. In an exercise, you'll show that since $f(x)$ is monic, we may assume that its factors $p(x)$ and $q(x)$ are monic, too. Then, $p(x) = x + \alpha$ and $q(x) = x + \beta$ where $\alpha, \beta \in \mathbb{R}$. We thus have the factorization $x^2 + 1 = (x + \alpha) \cdot (x + \beta)$. Expanding the right-hand side, we obtain $(x + \alpha) \cdot (x + \beta) = x^2 + (\alpha + \beta) \cdot x + (\alpha \cdot \beta)$. As this equals $x^2 + 1$, we obtain $\alpha + \beta = 0$ and $\alpha \cdot \beta = 1$. We'll leave it up to you to verify that this system of equations does not have a solution in $\mathbb{R}$, which is a contradiction. Hence $f(x)$ cannot be factorable; i.e., it must be unfactorable. (**Note:** In the next section, we'll derive a much quicker way to show that $f(x)$ is unfactorable in $\mathbb{R}[x]$.)

Example 30.7. Let $f(x) = x^2 + x + 1$ in $\mathbb{Z}_2[x]$. We'll show that $f(x)$ is unfactorable in $\mathbb{Z}_2[x]$. Assume for contradiction that $f(x)$ is factorable in $\mathbb{Z}_2[x]$. Similar to Example 30.6, we have $x^2 + x + 1 = p(x) \cdot q(x)$, where $p(x)$ and $q(x)$ have degree 1. But the only degree 1 polynomials in $\mathbb{Z}_2[x]$ are x and $x + 1$. Therefore, the possibilities for $p(x)$ and

$q(x)$ are as follows:

- $p(x) = x$ and $q(x) = x$, which implies $p(x) \cdot q(x) = x^2$.
- $p(x) = x$ and $q(x) = x + 1$ (or vice versa), which implies $p(x) \cdot q(x) = x^2 + x$.
- $p(x) = x + 1$ and $q(x) = x + 1$, which implies $p(x) \cdot q(x) = x^2 + 2x + 1 = x^2 + 1$.

Thus $p(x) \cdot q(x)$ never equals $f(x)$, and hence we have a contradiction. We conclude that $f(x)$ cannot be factorable; i.e., it must be unfactorable. (**Note:** Again, we'll derive a much quicker way soon.)

In Example 30.6, we saw that $f(x) = x^2+1$ is unfactorable in $\mathbb{R}[x]$. Thus, if we have $f(x) = p(x) \cdot q(x)$ where $p(x), q(x) \in \mathbb{R}[x]$, this cannot be a legitimate factorization of $f(x)$. Instead, this faux factorization must resemble $x^2 + 1 = 3 \cdot \left(\frac{1}{3}x^2 + \frac{1}{3}\right)$, where the factor 3 is a constant polynomial. (See Example 30.2.) The following theorem captures this observation.

Theorem 30.8. *Let F be a field. Suppose $f(x) \in F[x]$ with $\deg f(x) \geq 1$. Then $f(x)$ is unfactorable if and only if $f(x)$ satisfies the following property: If $f(x) = p(x) \cdot q(x)$, then $\deg p(x) = 0$ or $\deg q(x) = 0$.*

Remark. Many abstract algebra textbooks use Theorem 30.8 as the *definition* of an unfactorable polynomial, and they define factorable polynomials as those that are not unfactorable. In this textbook, we chose the approach of defining factorable polynomials first, since it probably more closely resembles your prior experiences with polynomials.

PROOF. We must prove two implications:

- If $f(x)$ is unfactorable, then $f(x)$ satisfies the property.
- If $f(x)$ satisfies the property, then $f(x)$ is unfactorable.

We will prove the first implication. The proof of the second implication is left for you as an exercise.

Assume $f(x)$ is unfactorable. Further assume that $f(x) = p(x) \cdot q(x)$ for some $p(x), q(x) \in F[x]$. We must show that $\deg p(x) = 0$ or $\deg q(x) = 0$. By Theorem 28.14, $\deg f(x) = \deg p(x) + \deg q(x)$, so that $\deg p(x), \deg q(x) \leq \deg f(x)$. Note that $\deg p(x)$ and $\deg q(x)$ cannot both be strictly less than $\deg f(x)$, since that would imply that $f(x)$ is factorable. Thus at least one of $\deg p(x)$ or $\deg q(x)$ equals $\deg f(x)$.

Suppose $\deg p(x) = \deg f(x)$. (The argument for the case $\deg q(x) = \deg f(x)$ follows similarly.) Then $\deg f(x) = \deg p(x) + \deg q(x)$ implies that $\deg q(x) = 0$ as desired. ■

30.2 Factorable or unfactorable?

We will consider various ways to determine whether a polynomial is factorable or unfactorable.

Example 30.9. Consider $f(x) = 2x - 7 \in \mathbb{R}[x]$. Then $f(x)$ is unfactorable in $\mathbb{R}[x]$. Intuitively, it is not possible to factor $f(x)$ into "smaller" factors that are not constants. Similarly, $4x + 2 \in \mathbb{Z}_7[x]$ is unfactorable in $\mathbb{Z}_7[x]$, and $ax + b \in F[x]$ (with $a \neq 0$) is unfactorable in $F[x]$.

Theorem 30.10. *Let F be a field. Suppose $f(x) \in F[x]$ with $\deg f(x) = 1$. Then $f(x)$ is unfactorable in $F[x]$.*

Proof know-how. The proof below demonstrates a common technique for showing that a polynomial is unfactorable. We will assume $f(x) = p(x) \cdot q(x)$ where $p(x)$, $q(x) \in F[x]$. Then we will show that $\deg p(x) = 0$ or $\deg q(x) = 0$. Then Theorem 30.8 implies that $f(x)$ is unfactorable.

PROOF. Suppose $f(x) = p(x) \cdot q(x)$ for some $p(x)$, $q(x) \in F[x]$. Since $\deg f(x) = \deg p(x) + \deg q(x)$ and $\deg f(x) = 1$, it follows that either $\deg p(x)$ or $\deg q(x)$, which are non-negative integers, must be 0. Then Theorem 30.8 implies that $f(x)$ is unfactorable. ■

Before proceeding, we introduce a terminology. For example, consider again $f(x) = x^2 + 1 \in \mathbb{Z}_5[x]$. Then, $f(3) = 3^2 + 1 = 10 = 0$ in $\mathbb{Z}_5$ and we call $3 \in \mathbb{Z}_5$ a *root* of the polynomial $f(x)$. However, $f(4) = 4^2 + 1 = 17 \neq 0$ in $\mathbb{Z}_5$, and thus $4 \in \mathbb{Z}_5$ is *not* a root of $f(x)$. Here is the generalization.

Definition 30.11 (Root of a polynomial)**.** Let F be a field, and let $f(x) \in F[x]$. We say that an element $\alpha \in F$ is a *root* of the polynomial $f(x)$ if $f(\alpha) = 0$.

Example 30.12. Let $f(x) = 4x + 5 \in \mathbb{R}[x]$. This polynomial has a root, namely $\alpha = -\frac{5}{4}$. To verify, note that $f\left(-\frac{5}{4}\right) = 4 \cdot \left(-\frac{5}{4}\right) + 5 = -5 + 5 = 0$. We found this root by solving the equation $4 \cdot \alpha + 5 = 0$ for α.

Example 30.13. Consider again $f(x) = 4x + 5$, but this time in $\mathbb{Z}_{11}[x]$. To find its root, we solve the equation $4 \cdot \alpha + 5 = 0$ and obtain $\alpha = 4^{-1} \cdot (-5)$. In $\mathbb{Z}_{11}$, we have $4^{-1} = 3$ since $4 \cdot 3 = 1$, and $-5 = 6$ since $5 + 6 = 0$. Therefore, $\alpha = 3 \cdot 6 = 18 = 7 \pmod{11}$. We verify that $f(7) = 4 \cdot 7 + 5 = 33 = 0 \pmod{11}$, so that $\alpha = 7$ is indeed a root of $f(x)$.

Examples 30.12 and 30.13 suggest the following theorem, whose proof is left for you as an exercise.

Theorem 30.14. *Let F be a field, and let $f(x) \in F[x]$. If $\deg f(x) = 1$, then $f(x)$ has a root.*

Example 30.15. Let $f(x) = x^3 + x + 1 \in \mathbb{Z}_3[x]$. Then $f(1) = 3 = 0$ so that $1 \in \mathbb{Z}_3$ is a root of $f(x)$. Since $f(1) = 0$, the factor theorem implies that $f(x) = (x - 1) \cdot q(x)$ for some $q(x) \in \mathbb{Z}_3[x]$. To conclude that this is a legitimate factorization of $f(x)$, we must show that $\deg(x - 1)$ and $\deg q(x)$ are both less than $\deg f(x)$.

We have $\deg f(x) = \deg(x - 1) + \deg q(x)$. Moreover, we know that $\deg f(x) = 3$ and $\deg(x - 1) = 1$. Thus, we must have $\deg q(x) = 2$. Hence $\deg(x - 1)$, $\deg q(x) < \deg f(x)$, so that $f(x)$ is factorable in $\mathbb{Z}_3[x]$. In summary, $f(x)$ having a root allowed us to conclude that $f(x)$ is factorable.

Theorem 30.16. *Let F be a field, and let $f(x) \in F[x]$ with $\deg f(x) \geq 2$. If $f(x)$ has a root $\alpha \in F$, then $f(x)$ is factorable in $F[x]$.*

PROOF. Assume $f(x)$ has a root $\alpha \in F$, so that $f(\alpha) = 0$. By the factor theorem, $f(x) = (x - \alpha) \cdot q(x)$ for some $q(x) \in F[x]$. To conclude that this is a legitimate factorization of $f(x)$, we must show that $\deg(x - \alpha)$ and $\deg q(x)$ are both less than $\deg f(x)$.

First, since $\deg(x - \alpha) = 1$ and $\deg f(x) \geq 2$, we conclude that $\deg(x - \alpha) < \deg f(x)$. Next, note that $\deg f(x) = \deg(x - \alpha) + \deg q(x) = 1 + \deg q(x)$, since $\deg(x - \alpha) = 1$. Solving for $\deg q(x)$, we obtain $\deg q(x) = \deg f(x) - 1$, so that $\deg q(x) < \deg f(x)$. Hence we have $f(x) = (x - \alpha) \cdot q(x)$ with $\deg(x - \alpha)$, $\deg q(x) < \deg f(x)$. Thus $f(x)$ is factorable in $F[x]$. ■

Proof know-how. To prove that a polynomial $f(x)$ is factorable, it's not enough to show that $f(x)$ can be written as $f(x) = p(x) \cdot q(x)$. We must also show that $\deg p(x)$, $\deg q(x) < \deg f(x)$; i.e., the degree of each factor is less than the degree of $f(x)$. Only then can we conclude that $f(x)$ is factorable.

Remark. In Theorem 30.16, we need the condition $\deg f(x) \geq 2$, as the theorem is false when $\deg f(x) = 1$. In Example 30.12, we saw that $f(x) = 4x + 5 \in \mathbb{R}[x]$ has a root, namely $\alpha = -\frac{5}{4}$. But we also know from Theorem 30.10 that all degree 1 polynomials are unfactorable.

Example 30.17. Let $f(x) = 5x^{493} + 2x^{314} + 3x^{235} + x^{102} + 6 \in \mathbb{Z}_{17}[x]$. Then $\deg f(x) = 493$, so that $\deg f(x) \geq 2$. We have $f(1) = 5 + 2 + 3 + 1 + 6 = 0 \pmod{17}$, so that $\alpha = 1$ is a root of $f(x)$. Thus by Theorem 30.16, we conclude that $f(x)$ is factorable in $\mathbb{Z}_{17}[x]$.

Next, we seek a tool to determine that $f(x)$ is unfactorable. Based on Theorem 30.16, we might consider the following:

If $f(x)$ has no root in F, then $f(x)$ is unfactorable in $F[x]$.

Unfortunately, this statement isn't true, as shown by the counterexample below.

Example 30.18. Consider $f(x) = x^4 + 3x^2 + 2 \in \mathbb{R}[x]$. For all $\alpha \in \mathbb{R}$, we have $\alpha^4 \geq 0$ and $\alpha^2 \geq 0$ so that $f(\alpha) \geq 2$. Thus, $f(\alpha)$ never equals 0, which means that $f(x)$ has no root in $\mathbb{R}$. However, we have $f(x) = (x^2 + 1) \cdot (x^2 + 2)$ so that $f(x)$ is factorable in $\mathbb{R}[x]$. The fact that $f(x)$ has no root means $f(x)$ has no *linear* factor. However, $f(x)$ could still have a factor of degree 2 (or higher), as we found in this example.

Here is how we can salvage the situation we encountered in Example 30.18.

Theorem 30.19. *Let F be a field, and let $f(x) \in F[x]$ with $\deg f(x) = 2$ or 3. If $f(x)$ has no root in F, then $f(x)$ is unfactorable in $F[x]$.*

PROOF. We prove the contrapositive; namely: If $f(x)$ is factorable in $F[x]$, then $f(x)$ has a root in F. Assume that $f(x)$ is factorable in $F[x]$. Then, $f(x) = p(x) \cdot q(x)$ where $p(x)$, $q(x) \in F[x]$ with $\deg p(x)$, $\deg q(x) < \deg f(x)$. We have $\deg f(x) = \deg p(x) + \deg q(x)$ and $\deg f(x) = 2$ or 3. Thus $\deg p(x)$ or $\deg q(x)$ must be 1. Suppose $\deg p(x) = 1$. (The argument for the case $\deg q(x) = 1$ follows similarly.) Then by Theorem 30.14, $p(x)$ has a root $\alpha \in F$ such that $p(\alpha) = 0$. Then $f(\alpha) = p(\alpha) \cdot q(\alpha) = 0 \cdot q(\alpha) = 0$, so that α is a root of $f(x)$ as well. Thus, $f(x)$ has a root in F as desired. ■

Remark. In the above proof, we used the fact that $\deg f(x) = 2$ or 3 to deduce that $\deg p(x)$ or $\deg q(x)$ must be 1. If $\deg f(x) = 3$, for instance, then the only way to satisfy

$\deg f(x) = \deg p(x) + \deg q(x)$ with $\deg p(x), \deg q(x) < \deg f(x)$ is either $3 = 1 + 2$ (i.e., $\deg p(x) = 1$) or $3 = 2 + 1$ (i.e., $\deg q(x) = 1$). A similar argument can be made in the case of $\deg f(x) = 2$.

Example 30.20. Consider $f(x) = x^2 + 1 \in \mathbb{R}[x]$. In Example 30.6, we showed (somewhat painstakingly) that $f(x)$ is unfactorable in $\mathbb{R}[x]$. Here, we observe that $\deg f(x) = 2$ and $f(x)$ has no root in $\mathbb{R}$, since there is no real number α such that $\alpha^2 + 1 = 0$. By Theorem 30.19, we conclude that $f(x)$ is unfactorable in $\mathbb{R}[x]$.

Example 30.21. Let $f(x) = x^2 - 2$. In Example 30.5, we saw that $f(x)$ is factorable in $\mathbb{R}[x]$. Now let's view this polynomial in $\mathbb{Q}[x]$. We have $\deg f(x) = 2$. Moreover, $f(x)$ has no root in $\mathbb{Q}$, since there is no rational number α such that $\alpha^2 - 2 = 0$ or, equivalently, $\alpha^2 = 2$. (See Theorem 1.10.) By Theorem 30.19, we conclude that $f(x)$ is unfactorable in $\mathbb{Q}[x]$.

Example 30.22. Let $f(x) = x^2 + x + 1$ in $\mathbb{Z}_2[x]$. In Example 30.7, we showed (again, painstakingly) that $f(x)$ is unfactorable in $\mathbb{Z}_2[x]$. Since $\mathbb{Z}_2 = \{0, 1\}$, there are only two candidates for a root of $f(x)$. We have $f(0) = 1$ and $f(1) = 1$ in $\mathbb{Z}_2$, so neither 0 nor 1 is a root of $f(x)$. Thus $f(x)$ has degree 2 but no root in $\mathbb{Z}_2$. By Theorem 30.19, we conclude that $f(x)$ is unfactorable in $\mathbb{Z}_2[x]$.

Example 30.23 (Non-example)**.** Let $f(x) = x^4 + x^2 + 1 \in \mathbb{Z}_2[x]$. We have $f(0) = 1$ and $f(1) = 1$ in $\mathbb{Z}_2$, so that $f(x)$ has no root in $\mathbb{Z}_2$. We might be tempted to conclude that $f(x)$ is unfactorable in $\mathbb{Z}_2[x]$. However, Theorem 30.19 does not apply to $f(x)$, since $\deg f(x) = 4$. In fact, $x^4 + x^2 + 1 = (x^2 + x + 1) \cdot (x^2 + x + 1)$, and thus $f(x)$ is factorable in $\mathbb{Z}_2[x]$.

Big picture stuff

We again highlight the *structural* similarity between the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$, where F is a field. The focus of this chapter was on factoring of polynomials, which turns out to have a lot in common with factoring of integers.

Factorable and unfactorable polynomials are similar to composite and prime integers. A factorable polynomial can be written as a product of two smaller polynomials, such as $x^2 + 1 = (x + 2) \cdot (x + 3)$ in $\mathbb{Z}_5[x]$. Analogously, a composite integer can be written as a product of two smaller integers, such as $15 = 3 \cdot 5$.

It turns out that both $\mathbb{Z}$ and $F[x]$ satisfy a fundamental property called "unique factorization." In $\mathbb{Z}$, for example, there is only one way to factor $15 = 3 \cdot 5$ into a product of primes. Similarly, in $\mathbb{Z}_5[x]$, $x^2 + 1 = (x + 2) \cdot (x + 3)$ is the only way to factor $x^2 + 1$ into a product of unfactorable polynomials. If this does not seem like a big deal, consider that in $\mathbb{Z}_{15}[x]$, where the coefficient ring $\mathbb{Z}_{15}$ is *not* a field, we have $x^2 + 6x + 8 = (x + 2) \cdot (x + 4)$, but also $x^2 + 6x + 8 = (x + 7) \cdot (x + 14)$. Thus, there are two *different* ways to factor $x^2 + 6x + 8$ in $\mathbb{Z}_{15}[x]$.

Exercises

1. In the definition of *factorable* in Definition 30.3, explain why $\deg p(x)$ and $\deg q(x)$ must be greater than 0.
2. **Prove:** Let F be a field, and let $f(x) \in F[x]$. Assume $f(x)$ is monic. If $f(x)$ is factorable in $F[x]$, then $f(x) = p(x) \cdot q(x)$ for some *monic* $p(x), q(x) \in F[x]$ with $\deg p(x), \deg q(x) < \deg f(x)$.

3. Complete the proof of Theorem 30.8 by proving its second implication.

4. (a) Let $f(x) = 2x - 7 \in \mathbb{R}[x]$. Find $\alpha \in \mathbb{R}$ such that $f(\alpha) = 0$.
   (b) Same as part (a), but with $f(x) = 4x + 2 \in \mathbb{Z}_7[x]$.
   (c) Same as part (a), but with $f(x) = 3x + 8 \in \mathbb{Z}_{13}[x]$.

5. Prove Theorem 30.14.

6. Explain what goes wrong in the proof of Theorem 30.16 if $\deg f(x) = 1$.

7. (a) Find an example of fields R and S with $R \subseteq S$ and a polynomial $f(x) \in R[x]$ (so that $f(x) \in S[x]$ also) such that $f(x)$ is unfactorable in $R[x]$, but $f(x)$ is factorable in $S[x]$.
   (b) Explain whether or not it is possible that $f(x)$ is factorable in $R[x]$, but $f(x)$ is unfactorable in $S[x]$.

8. Let $f(x) = x^3 + x + 1 \in \mathbb{Z}_3[x]$.
   (a) Compute $f(\alpha)$ for each $\alpha \in \mathbb{Z}_3$.
   (b) Is $f(x)$ factorable or unfactorable in $\mathbb{Z}_3[x]$? Explain your reasoning.

9. Let $f(x) = x^{217} + 100 \in \mathbb{Z}_{101}[x]$.
   (a) Find $\alpha \in \mathbb{Z}_{101}$ such that $f(\alpha) = 0$.
   (b) Is $f(x)$ factorable or unfactorable in $\mathbb{Z}_{101}[x]$? Explain your reasoning.

10. (a) Let $f(x) = x^3 + 4 \in \mathbb{Z}_5[x]$. Write $f(x)$ as a product of unfactorable polynomials in $\mathbb{Z}_5[x]$.
    (b) Repeat part (a) with $f(x) = x^3 + 6 \in \mathbb{Z}_7[x]$.
    (c) Repeat part (a) with $f(x) = x^3 + 10 \in \mathbb{Z}_{11}[x]$.
    (d) Repeat part (a) with $f(x) = x^3 + 12 \in \mathbb{Z}_{13}[x]$.

11. Determine if each polynomial is factorable or unfactorable.
    (a) $x^3 + x^2 + x + 1$ in $\mathbb{R}[x]$.
    (b) $x^3 + 2x + 1$ in $\mathbb{Z}_3[x]$.
    (c) $x^2 + 1$ in $\mathbb{Z}_7[x]$.
    (d) $x^{273} + 3x^{152} + 5x^{17} + 10$ in $\mathbb{Z}_{19}[x]$.

12. Determine if each polynomial is factorable or unfactorable.
    (a) $x^{100} - 1$ in $\mathbb{R}[x]$.
    (b) $x^3 + x + 1$ in $\mathbb{Z}_5[x]$.
    (c) $x^2 + 1$ in $\mathbb{Z}_{13}[x]$.
    (d) $x^{1,071} + 10x^{282} + 4x^{123} + 2$ in $\mathbb{Z}_{17}[x]$.

13. Explain why each polynomial is factorable in $F[x]$ where F is any field.
    (a) $x^3 + x^2 + x + 1$.
    (b) $x^3 + x^2 + 4$.

14. Consider $f(x) = x^2 + 1$ in $\mathbb{Z}_3[x]$. Show that $f(x)$ is unfactorable using these two methods:

    (a) Find all polynomials in $\mathbb{Z}_3[x]$ of degree 1 and show that the product of any two of them (including possibly the same two) never equals $f(x)$. (See Example 30.7.)

    (b) Use Theorem 30.19.

    Which method do you prefer?

15. In Example 30.22, we saw that $f(x) = x^2 + x + 1$ is unfactorable in $\mathbb{Z}_2[x]$. Verify that $f(x)$ is the *only* polynomial in $\mathbb{Z}_2[x]$ of degree 2 that is unfactorable.

16. Find all polynomials in $\mathbb{Z}_2[x]$ of degree 3 that are unfactorable.

17. Consider $f(x) = x^5 + x^4 + 1$ in $\mathbb{Z}_2[x]$.

    (a) Verify that $f(x)$ has no root in $\mathbb{Z}_2$.

    (b) Can we conclude that $f(x)$ is unfactorable in $\mathbb{Z}_2[x]$? Why or why not?

    (c) It turns out that $f(x)$ is factorable in $\mathbb{Z}_2[x]$. Find a legitimate factorization of $f(x)$; i.e., find polynomials $p(x)$, $q(x) \in \mathbb{Z}_2[x]$ with $\deg p(x)$, $\deg q(x) < \deg f(x)$ such that $f(x) = p(x) \cdot q(x)$.
    Hint: Exercise #15 should help.

18. (a) In $\mathbb{Z}_5[x]$, find the number of polynomials of the form $x^2 + bx + c$, with b, $c \in \mathbb{Z}_5$.

    (b) Of the polynomials in part (a), determine how many of them are factorable.

    (c) Repeat parts (a) and (b), but in $\mathbb{Z}_7[x]$ and $\mathbb{Z}_7$.

    (d) Repeat parts (a) and (b), but in $\mathbb{Z}_{13}[x]$ and $\mathbb{Z}_{13}$.

    (e) Repeat parts (a) and (b), but in $\mathbb{Z}_p[x]$ and $\mathbb{Z}_p$, where p is prime.

19. Let $a \in \mathbb{Z}_p$, $a \neq 0$, where p is prime. Prove that $f(x) = x^p - a$ is factorable in $\mathbb{Z}_p[x]$.

    Hint: We may view a as an element of the multiplicative group U_p.

20. Let $f(x) \in \mathbb{Z}_2[x]$ with $\deg f(x) \geq 2$. Prove that if $f(x)$ is unfactorable in $\mathbb{Z}_2[x]$, then $f(x)$ has an odd number of non-zero terms.

    Note: Example of $f(x) \in \mathbb{Z}_2[x]$ with an odd number of non-zero terms: $f(x) = x^{18} + x^{11} + x^8 + x^3 + 1$. There are five terms, and the coefficients must all be 1.

21. Determine if the converse of the statement in Exercise #20 is true or false. If it's true, prove it. If it's false, provide a counterexample.

22. Consider $f(x) = x^4 - 1 \in \mathbb{Z}_5[x]$.

    (a) Elizabeth says, "I can tell right away that $f(x)$ is factorable." How might she know?

    (b) Compute the product $(x-1) \cdot (x-2) \cdot (x-3) \cdot (x-4)$ in $\mathbb{Z}_5[x]$ and verify that it equals $f(x)$.

23. (a) How do you think $x^6 - 1$ would factor in $\mathbb{Z}_7[x]$? Verify your conjecture.
    (b) How do you think $x^{10} - 1$ would factor in $\mathbb{Z}_{11}[x]$? Verify your conjecture.
    (c) How do you think $x^{12} - 1$ would factor in $\mathbb{Z}_{13}[x]$? Yes, please verify.
    (d) What's going on here? Can you generalize and justify?

24. **Prove:** Let F be a field, $a \in F$ with $a \neq 0$, and $f(x) \in F[x]$. If $f(x)$ is unfactorable in $F[x]$, then $a \cdot f(x)$ is unfactorable in $F[x]$.

25. For each odd prime $p < 100$, determine if $f(x) = x^2+1$ is factorable or unfactorable in $\mathbb{Z}_p[x]$.

26. Let p be an odd prime (i.e., $p > 2$), and let $f(x) = x^2 + 1 \in \mathbb{Z}_p[x]$. Prove each statement below:
    (a) If $p = 4k + 3$ for some $k \in \mathbb{Z}$, then $f(x)$ is unfactorable in $\mathbb{Z}_p[x]$.
    (b) **(Challenge)** If $p = 4k + 1$ for some $k \in \mathbb{Z}$, then $f(x)$ is factorable in $\mathbb{Z}_p[x]$.

Unit VII: Quotient Rings

In this final unit of the book, we will investigate *quotient rings*, which are analogous to quotient groups studied earlier. More specifically, we will focus on quotient rings formed from polynomial rings $F[x]$ where F is a field. Thus, our quotient rings will have the form $F[x]/\langle g(x)\rangle$. (**Note:** $\langle g(x)\rangle$ is the *principal ideal* generated by $g(x)$. That notion is explained in Chapter 31.) We will conjecture and prove a criterion about when $F[x]/\langle g(x)\rangle$ is a field, which, continuing our theme about the structural similarities between $\mathbb{Z}$ and $F[x]$, is closely related to how $\mathbb{Z}_n$ is a field precisely when n is prime.

Along the way, we will explore (spoiler alert!) an isomorphism $\mathbb{R}[x]/\langle x^2+1\rangle \cong \mathbb{C}$, which highlights a beautiful connection between polynomial rings and complex numbers.

Here is a taste of what you'll be able to accomplish in this unit:

- Very quickly "reduce" the element $(4x^5+2x^3+4x+1)+\langle x^2-1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$.
- Understand the role that ideals play in making coset multiplication valid in quotient rings.
- Use *maximal ideals* to prove the criterion about when $F[x]/\langle g(x)\rangle$ is a field.

31

Ring Homomorphisms

A *ring homomorphism*, the focus of this chapter, is similar to its counterpart in group theory; i.e., it's an operation-preserving function. However, the domain and codomain of a ring homomorphism are (not surprisingly) rings, and thus both addition and multiplication must be preserved. Because every ring is an additive group, a ring homomorphism may also be viewed as a homomorphism of additive groups. As such, familiar properties from group homomorphisms still hold in this new setting. We will also revisit the notions of *kernel* and *image* in the context of ring homomorphisms.

We've seen that the kernel of a group homomorphism is a normal subgroup (Example 24.13), which played an integral role in our work with quotient groups. Analogously, we'll see that the kernel of a ring homomorphism is an *ideal*. And as we'll see in the next chapter, ideals are a necessary ingredient in making *quotient rings* work.

31.1 Evaluation map

Example 31.1. Consider the function $\theta : \mathbb{R}[x] \to \mathbb{R}$ where $\theta(f(x)) = f(2)$ for all $f(x) \in \mathbb{R}[x]$. This is an example of an *evaluation map*, where the input is a polynomial $f(x)$ and the corresponding output is $f(2)$, i.e., the value of that polynomial evaluated at $x = 2$. There is nothing special about 2 here. We could have fixed any real number at which to evaluate the input polynomial.

We choose inputs $f(x) = x^2 + 1$ and $g(x) = 4x + 5$ in $\mathbb{R}[x]$, and perform these computations:

- $\theta(f(x)+g(x))$ means first add the polynomials and then evaluate the sum at $x = 2$. Thus,

$$\theta(f(x) + g(x)) = \theta((x^2 + 1) + (4x + 5)) = \theta(x^2 + 4x + 6) = 2^2 + 4 \cdot 2 + 6 = 18.$$

- $\theta(f(x)) + \theta(g(x))$ means first find $f(2)$ and $g(2)$ and then add those real numbers. Thus,

$$\theta(f(x)) + \theta(g(x)) = f(2) + g(2) = (2^2 + 1) + (4 \cdot 2 + 5) = 5 + 13 = 18.$$

- $\theta(f(x)\cdot g(x))$ means first multiply the polynomials and then evaluate the product at $x = 2$. Thus,

$$\theta(f(x)\cdot g(x)) = \theta((x^2+1)\cdot(4x+5)) = \theta(4x^3+5x^2+4x+5) = 4\cdot 2^3+5\cdot 2^2+4\cdot 2+5 = 65.$$

- $\theta(f(x))\cdot\theta(g(x))$ means first find $f(2)$ and $g(2)$ and then multiply those real numbers. Thus,

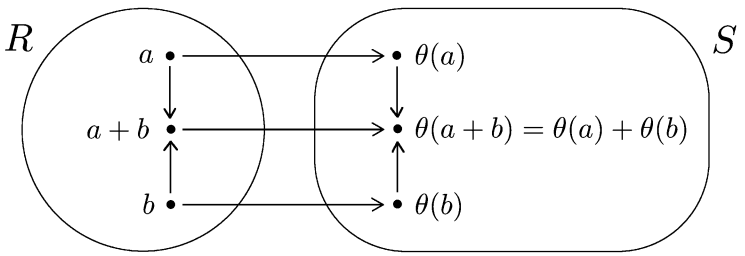
$$\theta(f(x))\cdot\theta(g(x)) = f(2)\cdot g(2) = (2^2+1)\cdot(4\cdot 2+5) = 5\cdot 13 = 65.$$

Thus, we have $\theta(f(x)+g(x)) = \theta(f(x))+\theta(g(x))$ and $\theta(f(x)\cdot g(x)) = \theta(f(x))\cdot\theta(g(x))$. In fact, these relationships hold for *all* $f(x), g(x) \in \mathbb{R}[x]$. We'll leave the justification to you as an exercise.

The relationships in Example 31.1 should feel familiar. We encountered such relationships when we studied group homomorphisms in Chapter 17. With *ring homomorphisms*, however, we require both operations addition and multiplication to be preserved.

Definition 31.2 (Ring homomorphism)**.** Let R and S be rings. A function $\theta : R \to S$ is a *ring homomorphism* if $\theta(a+b) = \theta(a)+\theta(b)$ and $\theta(a\cdot b) = \theta(a)\cdot\theta(b)$ for all $a, b \in R$.

The diagram below shows how addition is preserved by θ. The relationship $\theta(a+b) = \theta(a)+\theta(b)$ means it doesn't matter whether we first add in R and then apply θ (i.e., $\theta(a+b)$), or first apply θ to each and then add in S (i.e., $\theta(a)+\theta(b)$). A similar diagram and interpretation can be made for the multiplicative relationship $\theta(a\cdot b) = \theta(a)\cdot\theta(b)$, which we'll leave up to you.



Remark. In Definition 31.2, if we denote the addition operations in R and S by $+_R$ and $+_S$, respectively, then $\theta(a+b) = \theta(a)+\theta(b)$ would be written $\theta(a+_R b) = \theta(a)+_S\theta(b)$. Given our prior experience with homomorphisms (from group theory), we will simply write $\theta(a+b) = \theta(a)+\theta(b)$ and $\theta(a\cdot b) = \theta(a)\cdot\theta(b)$.

Example 31.3. The evaluation map $\theta : \mathbb{R}[x] \to \mathbb{R}$ where $\theta(f(x)) = f(2)$ for all $f(x) \in \mathbb{R}[x]$ is a ring homomorphism. (See Example 31.1.)

Example 31.4. Consider the function $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. We have

- $\varphi(26+17) = \varphi(43) = 43 \pmod 5 = 3 \pmod 5$ and
- $\varphi(26)+\varphi(17) = 26 \pmod 5 + 17 \pmod 5 = 1 \pmod 5 + 2 \pmod 5 = 3 \pmod 5$,

so that $\varphi(26 + 17) = \varphi(26) + \varphi(17)$. Likewise, we have

- $\varphi(26 \cdot 17) = \varphi(442) = 2 \pmod 5$ and
- $\varphi(26) \cdot \varphi(17) = 26 \pmod 5 \cdot 17 \pmod 5 = 1 \pmod 5 \cdot 2 \pmod 5 = 2 \pmod 5$,

so that $\varphi(26 \cdot 17) = \varphi(26) \cdot \varphi(17)$. You'll show in an exercise at the end of the chapter that the above relationships hold for all $a, b \in \mathbb{Z}$ (and not just for 26 and 17). Therefore, φ preserves both operations and hence is a ring homomorphism.

Example 31.5. Consider the function $\lambda : \mathbb{Z}_{10} \to \mathbb{Z}_{30}$ where $\lambda(a) = 6a$ for all $a \in \mathbb{Z}_{10}$. We have

- $\lambda(7 + 4) = \lambda(1) = 6 \cdot 1 = 6$ and
- $\lambda(7) + \lambda(4) = 6 \cdot 7 + 6 \cdot 4 = 66 = 6$,

so that $\lambda(7 + 4) = \lambda(7) + \lambda(4)$. Note that the reductions $7 + 4 = 1$ and $66 = 6$ are done in $\mathbb{Z}_{10}$ and $\mathbb{Z}_{30}$, respectively. We also have

- $\lambda(7 \cdot 4) = \lambda(8) = 6 \cdot 8 = 18$ and
- $\lambda(7) \cdot \lambda(4) = (6 \cdot 7) \cdot (6 \cdot 4) = 1{,}008 = 18$,

so that $\lambda(7 \cdot 4) = \lambda(7) \cdot \lambda(4)$. (We'll leave it up to you to identify where the reductions take place.)

Now let's generalize. For $a, b \in \mathbb{Z}_{10}$, we have $\lambda(a+b) = 6(a+b) = 6a+6b = \lambda(a)+\lambda(b)$. Therefore, $\lambda(a+b) = \lambda(a)+\lambda(b)$ so that λ preserves addition. For multiplication, observe first that $\lambda(a \cdot b) = 6(ab)$. We also have $\lambda(a) \cdot \lambda(b) = 6a \cdot 6b = 36(ab) = 6(ab)$, since $36 = 6$ in $\mathbb{Z}_{30}$. Therefore, $\lambda(a \cdot b) = \lambda(a) \cdot \lambda(b)$. Thus, λ is a ring homomorphism.

Example 31.6. In Example 26.6, we considered the ring $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$, where $i = \sqrt{-1}$ so that $i^2 = -1$. Define the function $\theta : \mathbb{Z}_3[i] \to \mathbb{Z}_3[i]$ where $\theta(a+bi) = a - bi$ for all $a + bi \in \mathbb{Z}_3[i]$. For instance, we have $\theta(1 + 2i) = 1 - 2i = 1 + i$, since $-2 = 1$ in $\mathbb{Z}_3$. For $a + bi, c + di \in \mathbb{Z}_3[i]$, we have

$$\begin{aligned}\theta((a + bi) + (c + di)) &= \theta((a + c) + (b + d)i) \\ &= (a + c) - (b + d)i \\ &= (a - bi) + (c - di) \\ &= \theta(a + bi) + \theta(c + di),\end{aligned}$$

so that $\theta((a + bi) + (c + di)) = \theta(a + bi) + \theta(c + di)$. Thus, θ preserves addition. We'll leave it up to you to verify that θ preserves multiplication as well. Therefore, θ is a ring homomorphism. In fact, θ is a ring *isomorphism*, as defined below. (We'll leave that verification to you, also.)

Definition 31.7 (Ring isomorphism)**.** Let R and S be rings. A function $\theta : R \to S$ is a *ring isomorphism* if θ is a bijection (i.e., one-to-one and onto) and θ preserves both addition and multiplication.

Example 31.8. Consider the function $\theta : \mathbb{Z}_2[x] \to \mathbb{Z}_2[x]$ where $\theta(f(x)) = f(x)^2$ for all $f(x) \in \mathbb{Z}_2[x]$. For instance, suppose $f(x) = x^2 + 1 \in \mathbb{Z}_2[x]$. Then

$$\theta(f(x)) = (x^2 + 1)^2 = (x^2 + 1) \cdot (x^2 + 1) = x^4 + \mathbf{2}x^2 + 1 = x^4 + \mathbf{0}x^2 + 1,$$

so that $\theta(f(x)) = x^4 + 1$. For $f(x), g(x) \in \mathbb{Z}_2[x]$, we have

$$\begin{aligned}\theta(f(x) + g(x)) &= (f(x) + g(x))^2 \\ &= (f(x) + g(x)) \cdot (f(x) + g(x)) \\ &= f(x) \cdot f(x) + f(x) \cdot g(x) + g(x) \cdot f(x) + g(x) \cdot g(x) \\ &= f(x)^2 + \mathbf{2} \cdot (f(x) \cdot g(x)) + g(x)^2 \\ &= f(x)^2 + \mathbf{0} \cdot (f(x) \cdot g(x)) + g(x)^2 \\ &= \theta(f(x)) + \theta(g(x))\end{aligned}$$

so that $\theta(f(x) + g(x)) = \theta(f(x)) + \theta(g(x))$. For multiplication, we have

$$\begin{aligned}\theta(f(x) \cdot g(x)) &= (f(x) \cdot g(x))^2 \\ &= (f(x) \cdot g(x)) \cdot (f(x) \cdot g(x)) \\ &= (f(x) \cdot f(x)) \cdot (g(x) \cdot g(x)) \\ &= f(x)^2 \cdot g(x)^2 \\ &= \theta(f(x)) \cdot \theta(g(x)),\end{aligned}$$

so that $\theta(f(x) \cdot g(x)) = \theta(f(x)) \cdot \theta(g(x))$. Thus, θ is a ring homomorphism.

Example 31.9. Let R and S be rings. Define the function $\theta : R \to S$ where $\theta(r) = 0_S$ for all $r \in R$. (Here, 0_S denotes the additive identity element of S.) For $a, b \in R$, we have $\theta(a + b) = 0_S = 0_S + 0_S = \theta(a) + \theta(b)$, so that $\theta(a + b) = \theta(a) + \theta(b)$. For multiplication, we have $\theta(a \cdot b) = 0_S = 0_S \cdot 0_S = \theta(a) \cdot \theta(b)$, so that $\theta(a \cdot b) = \theta(a) \cdot \theta(b)$. Thus, θ is a ring homomorphism, often called the *trivial homomorphism*.

Example 31.10 (Non-example)**.** Consider the determinant function $\delta : M(\mathbb{Z}_{10}) \to \mathbb{Z}_{10}$ where $\delta(\alpha) = \det \alpha$ for all $\alpha \in M(\mathbb{Z}_{10})$. For $\alpha, \beta \in M(\mathbb{Z}_{10})$, Theorem 7.24 implies

$$\delta(\alpha \cdot \beta) = \det(\alpha \cdot \beta) = \det \alpha \cdot \det \beta = \delta(\alpha) \cdot \delta(\beta),$$

so that $\delta(\alpha \cdot \beta) = \delta(\alpha) \cdot \delta(\beta)$ and δ preserves multiplication. (Recall from Example 17.5 that the above calculation shows that the determinant function is a *group* homomorphism from $G(\mathbb{Z}_{10})$ to U_{10}.)

However δ fails to preserve addition. As a counterexample, suppose $\alpha, \beta \in M(\mathbb{Z}_{10})$, where $\alpha = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ and $\beta = \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]$. Then $\delta(\alpha + \beta) = \delta\left(\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right] + \left[\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right]\right) = \delta\left(\left[\begin{smallmatrix} 6 & 8 \\ 0 & 2 \end{smallmatrix}\right]\right) = \det\left(\left[\begin{smallmatrix} 6 & 8 \\ 0 & 2 \end{smallmatrix}\right]\right) = 2$, so that $\delta(\alpha + \beta) = 2$. However, we have $\delta(\alpha) + \delta(\beta) = \det \alpha + \det \beta = 8 + 8 = 6$. Thus, $\delta(\alpha + \beta) \neq \delta(\alpha) + \delta(\beta)$, so that δ does *not* preserve addition and thus it's *not* a ring homomorphism.

31.2 Properties of ring homomorphisms

Recall that all rings are *additive* groups. Thus, a ring homomorphism $\theta : R \to S$ may be viewed as a homomorphism of additive groups. As such, familiar properties from group homomorphisms still hold.

First, we have $\theta(0_R) = 0_S$; i.e., θ maps the additive identity of R to the additive identity of S. This is the additive version of Theorem 17.9.

Example 31.11. Consider the ring homomorphism $\lambda : \mathbb{Z}_{10} \to \mathbb{Z}_{30}$ where $\lambda(a) = 6a$ for all $a \in \mathbb{Z}_{10}$. (See Example 31.5.) We then have $\lambda(0) = 6 \cdot 0 = 0$, so that the additive identity $0 \in \mathbb{Z}_{10}$ of the domain maps to the additive identity $0 \in \mathbb{Z}_{30}$ of the codomain.

Remark. We must be careful *not* to assume $\theta(1_R) = 1_S$, i.e., that θ maps the *multiplicative* identity of R to the multiplicative identity of S. This isn't necessarily true. In the ring homomorphism λ above, for instance, we have $\lambda(1) = 6 \cdot 1 = 6$, so that $\lambda(1) \neq 1$.

Second, we have $\theta(-r) = -\theta(r)$ for all $r \in R$, so that θ maps additive inverses in R to additive inverses in S. This is the additive version of Theorem 17.12.

Example 31.12. Consider the ring homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. (See Example 31.4.) Then $\varphi(-23) = -23 \pmod 5 = 2 \pmod 5$ and $-\varphi(23) = -(3 \pmod 5) = 2 \pmod 5$, so that $\varphi(-23) = -\varphi(23)$. In other words, it doesn't matter whether you first negate and then reduce mod 5, or first reduce mod 5 and then negate.

Third, we have $\theta(n \cdot r) = n \cdot \theta(r)$ for all $r \in R$ and $n \in \mathbb{Z}$. This is the additive version of Theorem 17.13. Here, the expression $n \cdot r$ means the element r is added to itself n times, and likewise for $n \cdot \theta(r)$. If $n = 3$, for instance, we have $\theta(3 \cdot r) = \theta(r + r + r) = \theta(r) + \theta(r) + \theta(r) = 3 \cdot \theta(r)$, so that $\theta(3 \cdot r) = 3 \cdot \theta(r)$. Similarly, $n = -3$ is interpreted as follows:

$$\begin{aligned} \theta(-3 \cdot r) &= \theta((-r) + (-r) + (-r)) \\ &= \theta(-r) + \theta(-r) + \theta(-r) \\ &= (-\theta(r)) + (-\theta(r)) + (-\theta(r)) \\ &= -3 \cdot \theta(r), \end{aligned}$$

so that $\theta(-3 \cdot r) = -3 \cdot \theta(r)$.

Fourth, repeated application of $\theta(a \cdot b) = \theta(a) \cdot \theta(b)$ implies $\theta(r^n) = \theta(r)^n$ for $r \in R$, as long as n is a positive integer. With $n = 3$, for instance, we have $\theta(r^3) = \theta(r \cdot r \cdot r) = \theta(r) \cdot \theta(r) \cdot \theta(r) = \theta(r)^3$. Do you see why n must be positive in this situation?

31.3 Kernel and image

Just as with group homomorphisms, we define the *kernel* and *image* of a ring homomorphism. Let's start with the kernel. Given a *group* homomorphism $\varphi : G \to H$, we had defined its kernel as

$$\ker \varphi = \{a \in G \mid \varphi(a) = \varepsilon_H\},$$

i.e., the set of elements in the domain G that map to the identity element ε_H in the codomain H. Now suppose $\theta : R \to S$ is a *ring* homormorphism. Since the codomain S has two identities 0_S and 1_S, there are two natural candidates for $\ker \theta$, namely

$$\{r \in R \mid \theta(r) = 0_S\} \quad \text{and} \quad \{r \in R \mid \theta(r) = 1_S\}.$$

As we did in Section 31.2, we want to take advantage of the fact that all rings are *additive* groups. Thus, we make the following choice for the definition of the *kernel* in the ring setting.

Definition 31.13 (Kernel of a ring homomorphism)**.** Let $\theta : R \to S$ be a ring homomorphism. The *kernel* of θ is the set $\ker \theta = \{r \in R \mid \theta(r) = 0_S\}$, where 0_S is the additive identity of S.

The image of a ring homomorphism is defined just like its counterpart in group theory.

Definition 31.14 (Image of a ring homomorphism)**.** Let $\theta : R \to S$ be a ring homomorphism. The *image* of θ is the set $\operatorname{im}\theta = \{\theta(r) \mid r \in R\}$.

Again, since we may view θ as a homomorphism of additive groups, we conclude that $\ker\theta$ is an additive subgroup of R (Theorem 18.6) and $\operatorname{im}\theta$ is an additive subgroup of S (Theorem 18.11). We record this as a theorem about ring homomorphisms.

Theorem 31.15. *Let $\theta : R \to S$ be a ring homomorphism. Then $\ker\theta$ is an additive subgroup of the domain R and $\operatorname{im}\theta$ is an additive subgroup of the codomain S.*

Example 31.16. Consider the ring homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ from Example 31.4 where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. Then $\ker\varphi$ contains integers a such that $\varphi(a) = 0$; i.e., $a = 0 \pmod 5$. These are precisely the multiples of 5, and thus $\ker\varphi = 5\mathbb{Z}$. We also have $\operatorname{im}\varphi = \mathbb{Z}_5$, which implies that φ is onto. Note that $\ker\varphi = 5\mathbb{Z}$ and $\operatorname{im}\varphi = \mathbb{Z}_5$ are additive subgroups of the domain $\mathbb{Z}$ and codomain $\mathbb{Z}_5$, respectively.

Example 31.17. Consider the ring homomorphism $\lambda : \mathbb{Z}_{10} \to \mathbb{Z}_{30}$ from Example 31.5 where $\lambda(a) = 6a$ for all $a \in \mathbb{Z}_{10}$. We have $\lambda(0) = 6 \cdot 0 = 0$, $\lambda(5) = 6 \cdot 5 = 0$, and $\lambda(a) \neq 0$ for all other elements $a \in \mathbb{Z}_{10}$. Thus $\ker\lambda = \{0, 5\}$, which is a subgroup of the domain $\mathbb{Z}_{10}$. Computing $\lambda(a)$ for each $a \in \mathbb{Z}_{10}$, we see that the distinct outputs of the function λ are 0, 6, 12, 18, and 24. Therefore $\operatorname{im}\lambda = \{0, 6, 12, 18, 24\}$, which is a subgroup of the codomain $\mathbb{Z}_{30}$.

Example 31.18. Consider the evaluation map $\theta : \mathbb{R}[x] \to \mathbb{R}$ from Example 31.1, where $\theta(f(x)) = f(2)$ for all $f(x) \in \mathbb{R}[x]$. Then $\ker\theta = \{f(x) \in \mathbb{R}[x] \mid \theta(f(x)) = 0\}$. For instance, let $p(x) = x^2 + x - 6$ and $q(x) = x^3 - 7$ be elements of $\mathbb{R}[x]$. Then $\theta(p(x)) = p(2) = 2^2 + 2 - 6 = 0$, so that $p(x) \in \ker\theta$. However, $\theta(q(x)) = q(2) = 2^3 - 7 = 1 \neq 0$, so that $q(x) \notin \ker\theta$.

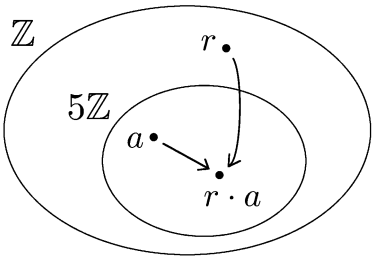
In general, a polynomial $f(x)$ is in $\ker\theta$ when $\theta(f(x)) = 0$, i.e., when $f(2) = 0$. By the factor theorem, these are precisely the polynomials that have $x-2$ as a factor, such as $p(x) = x^2+x-6 = (x-2)\cdot(x+3)$. We conclude that $\ker\theta = \{(x-2)\cdot g(x) \mid g(x) \in \mathbb{R}[x]\}$, i.e., the set of all polynomial multiples of $x - 2$.

Moreover, $\operatorname{im}\theta = \mathbb{R}$, or equivalently, θ is onto. Given $a \in \mathbb{R}$, let $f(x) = (x-2)+a \in \mathbb{R}[x]$. Then, $\theta(f(x)) = f(2) = (2-2)+a = a$. Thus, every element in $\mathbb{R}$ (the codomain) gets "hit" by the function θ.

31.4 Examples and definition of an ideal

Example 31.19. Consider the ring homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$, where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. We saw in Example 31.16 that $\ker\varphi = 5\mathbb{Z}$. While $5\mathbb{Z}$ is an additive subgroup of the domain $\mathbb{Z}$, it is *not* a subring of $\mathbb{Z}$, since $5\mathbb{Z}$ does not contain the multiplicative identity 1. Instead, $5\mathbb{Z}$ satisfies what we'll call the *product absorption* property: If $r \in \mathbb{Z}$ (the domain) and $a \in 5\mathbb{Z}$, then $r \cdot a \in 5\mathbb{Z}$. For instance, let $r = 7$ and $a = 10$. Then $r \cdot a = 70$, which is in $5\mathbb{Z}$. More generally, suppose $r \in \mathbb{Z}$ and $a \in 5\mathbb{Z}$ so that $a = 5n$ for some $n \in \mathbb{Z}$. Then $r \cdot a = r \cdot 5n = 5 \cdot (rn) \in 5\mathbb{Z}$, so that $r \cdot a \in 5\mathbb{Z}$.

Below is a visual depiction of product absorption. It's as if the element $a \in 5\mathbb{Z}$ *absorbs* the element $r \in \mathbb{Z}$ into the set $5\mathbb{Z}$ when they are multiplied together.



Example 31.20. Consider the ring homomorphism $\lambda : \mathbb{Z}_{10} \to \mathbb{Z}_{30}$, where $\lambda(a) = 6a$ for all $a \in \mathbb{Z}_{10}$. We saw in Example 31.17 that $\ker \lambda = \{0, 5\}$. Observe that $\{0, 5\}$ is an additive subgroup of the domain $\mathbb{Z}_{10}$, but not its subring. We'll show that $\{0, 5\}$ also satisfies the product absorption property: If $r \in \mathbb{Z}_{10}$ (the domain) and $a \in \{0, 5\}$, then $r \cdot a \in \{0, 5\}$. There are two cases to consider: $a = 0$ or $a = 5$. If $a = 0$, then $r \cdot a = 0 \in \{0, 5\}$ for all $r \in \mathbb{Z}_{10}$. If $a = 5$, then $r \cdot a = 0$ for $r = 0, 2, 4, 6$, and 8; and $r \cdot a = 5$ for $r = 1, 3, 5, 7$, and 9. Thus, in either case, $r \cdot a \in \{0, 5\}$.

Example 31.21. Consider the ring homomorphism $\theta : \mathbb{R}[x] \to \mathbb{R}$ from Example 31.1, where $\theta(f(x)) = f(2)$ for all $f(x) \in \mathbb{R}[x]$. We saw in Example 31.18 that $\ker \theta = \{(x-2) \cdot g(x) \mid g(x) \in \mathbb{R}[x]\}$. That $\ker \theta$ is an additive subgroup of $\mathbb{R}[x]$ follows from Theorem 31.15. To verify that $\ker \theta$ satisfies the product absorption property, let $f(x) \in \mathbb{R}[x]$ and $p(x) \in \ker \theta$, so that $p(x) = (x-2) \cdot g(x)$ for some $g(x) \in \mathbb{R}[x]$. We have $\theta(f(x) \cdot p(x)) = \theta(f(x)) \cdot \theta(p(x)) = f(2) \cdot p(2)$. But $p(2) = (2-2) \cdot g(2) = 0 \cdot g(2) = 0$. Thus, $\theta(f(x) \cdot p(x)) = f(2) \cdot 0 = 0$, and so $f(x) \cdot p(x)$ is contained in $\ker \theta$ as well.

In the above examples, $\ker \varphi = 5\mathbb{Z}$, $\ker \lambda = \{0, 5\}$, and $\ker \theta = \{(x-2) \cdot g(x) \mid g(x) \in \mathbb{R}[x]\}$ are additive subgroups of the domain and satisfy the product absorption property. They are examples of an *ideal*, defined below. As mentioned in the introduction to this chapter, an ideal of a ring is analogous to a normal subgroup of a group. Thus, ideals will play a critical role in our work with *quotient rings* in the next chapter.

Definition 31.22 (Ideal of a ring)**.** A subset A of a ring R is called an *ideal* of R if the following are true:

- A is an additive subgroup of R.
- A satisfies the *product absorption property*: If $r \in R$ and $a \in A$, then $r \cdot a \in A$.

Remark. For product absorption, we *should* say $r \cdot a,\ a \cdot r \in A$, since R isn't necessarily commutative. That being said, we'll mostly work with commutative rings for the remainder of this textbook. And when R is commutative, it's fine to consider $r \cdot a \in A$ only.

Example 31.23. Let R be a ring. Then R itself is an ideal of R. The subset $\{0\}$ containing only the additive identity of R is also an ideal of R, and it is called the *trivial* ideal.

Examples 31.19, 31.20, and 31.21 seem to suggest that the kernel of a ring homomorphism is an ideal of the domain. Here is the generalization. Compare this with Theorem 24.17, which states that the kernel of a group homomorphism is a normal subgroup of the domain.

Theorem 31.24 (Kernel is an ideal)**.** *Let $\theta : R \to S$ be a ring homomorphism with $K = \ker\theta$. Then K is an ideal of the domain R.*

PROOF. By Theorem 31.15, the kernel K is an additive subgroup of the domain R. Therefore, it suffices to show that K satisfies the product absorption property. Assume $r \in R$ and $k \in K$. Then $\theta(k) = 0_S$, and we have $\theta(r \cdot k) = \theta(r) \cdot \theta(k) = \theta(r) \cdot 0_S = 0_S$. Hence, $\theta(r \cdot k) = 0_S$ so that $r \cdot k \in K$. A similar argument shows that $k \cdot r \in K$. Thus, K is an ideal of R. ∎

Proof know-how. The above proof demonstrates a common technique for showing that a set K in a ring R satisfies the product absorption property. (**Note:** To prove that K is an ideal of R, you must separately show that K is an additive subgroup of R.) Begin by choosing two elements: $r \in R$ and $k \in K$. Then argue that the product $r \cdot k$ is contained in K. This argument typically uses the fact that k is an element of K. In the above proof, for instance, we used the fact that $\theta(k) = 0_S$ to show that $\theta(r \cdot k) = 0_S$.

Example 31.25. Thus far, we've seen the following examples of an ideal:

- $5\mathbb{Z}$ is an ideal of $\mathbb{Z}$, where $5\mathbb{Z} = \{5 \cdot n \mid n \in \mathbb{Z}\}$.
- $\{0,\ 5\}$ is an ideal of $\mathbb{Z}_{10}$, where $\{0,\ 5\} = \{5 \cdot n \mid n \in \mathbb{Z}_{10}\}$.
- $\{(x-2) \cdot g(x) \mid g(x) \in \mathbb{R}[x]\}$ is an ideal of $\mathbb{R}[x]$.

Each is a set of multiples of a particular element in the ring, which motivates the following definition.

Definition 31.26 (Principal ideal)**.** Let R be a commutative ring and fix an element $a \in R$. The set $\langle a\rangle = \{a \cdot r \mid r \in R\}$ is called the *principal ideal* generated by a; i.e., $\langle a\rangle$ is the set of all multiples of a.

Example 31.27. Recall that in the group setting, $\langle a\rangle$ means the *cyclic subgroup* generated by a, i.e., the set of all integer powers of a. The meaning of $\langle a\rangle$ should be clear from the context.

Example 31.28. Each ideal in Example 31.25 is a principal ideal:

- With $5 \in \mathbb{Z}$, we have $\langle 5\rangle = \{5 \cdot r \mid r \in \mathbb{Z}\} = 5\mathbb{Z}$, an ideal of $\mathbb{Z}$.
- With $5 \in \mathbb{Z}_{10}$, we have $\langle 5\rangle = \{5 \cdot r \mid r \in \mathbb{Z}_{10}\} = \{0,\ 5\}$, an ideal of $\mathbb{Z}_{10}$.
- With $x-2 \in \mathbb{R}[x]$, we have $\langle x-2\rangle = \{(x-2) \cdot g(x) \mid g(x) \in \mathbb{R}[x]\}$, an ideal of $\mathbb{R}[x]$.

Example 31.29. Let R be a ring. In Example 31.23, we noted that R itself and $\{0\}$ are ideals of R. These are principal ideals as well, generated by the multiplicative and additive identities of R, respectively:

- With $1 \in R$, we have $\langle 1\rangle = \{1 \cdot r \mid r \in R\} = R$.
- With $0 \in R$, we have $\langle 0\rangle = \{0 \cdot r \mid r \in R\} = \{0\}$.

As its name suggests, a principal *ideal* $\langle a\rangle = \{a \cdot r \mid r \in R\}$ is indeed an ideal of R. The proof of the following theorem is left for you as an exercise.

Theorem 31.30. *Let R be a commutative ring and fix an element $a \in R$. Then the set $\langle a\rangle = \{a \cdot r \mid r \in R\}$ is an ideal of R.*

31.5 Ideals in $\mathbb{Z}$ and in $F[x]$

All of the ideals we've examined so far have been *principal ideals* of the form $\langle a\rangle = \{a \cdot r \mid r \in R\}$, i.e., the set of all multiples of a fixed element $a \in R$. The next example shows that not every ideal is principal.

Example 31.31. Let A be the set of all polynomials in $\mathbb{Z}[x]$ with an even constant term. For instance, consider $f(x) = 3x^{101} - 171x^{52} + x + 12$ and $g(x) = 5x - 21$, which are elements of $\mathbb{Z}[x]$. The constant terms of $f(x)$ and $g(x)$ are 12 (which is even) and -21 (which isn't even), respectively. Therefore, $f(x) \in A$ and $g(x) \notin A$. In the exercises, you'll show that (1) A is an ideal of $\mathbb{Z}[x]$, but (2) A is *not* a principal ideal; i.e., there does *not* exist an element $\alpha(x) \in \mathbb{Z}[x]$ such that $A = \langle\alpha(x)\rangle$.

Example 31.32. We've seen that $\langle 5\rangle = 5\mathbb{Z}$ is an ideal of $\mathbb{Z}$. Other ideals of $\mathbb{Z}$ include $\langle 2\rangle = 2\mathbb{Z}$, $\langle 12\rangle = 12\mathbb{Z}$, and more generally, $\langle n\rangle = n\mathbb{Z}$, where n is a fixed integer. These include the two extremes, i.e., the ring $\mathbb{Z}$ itself, $\langle 1\rangle = \mathbb{Z}$, and the trivial ideal $\langle 0\rangle = \{0\}$.

It turns out that every ideal of $\mathbb{Z}$ is principal, as the following theorem states. We'll start the proof but will leave it to you as an exercise to finish writing it.

Theorem 31.33. *Every ideal of $\mathbb{Z}$ is a principal ideal.*

PROOF. Let A be an ideal of $\mathbb{Z}$. There are two cases to consider; namely: $A = \{0\}$ and $A \neq \{0\}$.

If $A = \{0\}$, then we have $A = \langle 0\rangle$ so that A is principal.

Thus, assume $A \neq \{0\}$ so that A contains a non-zero element. Further, A is an additive subgroup of $\mathbb{Z}$, which means if $a \in A$, then $-a \in A$. Therefore, A must contain a positive integer. Let d be the *smallest* positive integer in A. We will show that $A = \langle d\rangle$ by showing $A \subseteq \langle d\rangle$ and $\langle d\rangle \subseteq A$.

Below, we will show that $\langle d\rangle \subseteq A$. The other set inclusion is left to you as an exercise. Let $n \in \langle d\rangle$ so that $n = d \cdot r$ for some $r \in \mathbb{Z}$. We have $d \in A$, $r \in \mathbb{Z}$, and A satisfies the product absorption property, since it is an ideal. Therefore, $n = d \cdot r \in A$, so that $\langle d\rangle \subseteq A$ as desired. ■

In recent chapters, we've described the many structural similarities between $\mathbb{Z}$ and the polynomial ring $F[x]$ where F is a field. The following theorem, whose proof is left for you, captures another such similarity.

Theorem 31.34. *Let F be a field. Then every ideal of $F[x]$ is a principal ideal.*

Big picture stuff

We continue to highlight the connections between the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$, where F is a field. As Theorems 31.33 and 31.34 indicate, both rings satisfy the condition that every ideal is principal. More generally, an integral domain whose ideals are all principal is called a *principal ideal domain* (or PID). And $\mathbb{Z}$ and $F[x]$ are classic examples of a PID.

For another connection, we restate an early observation in the language of principal ideals. In Chapter 3, Exercises #12 and #13, we proved the following:

Let $m, n \in \mathbb{Z}$. Then $n \mid m$ if and only if $\langle m\rangle \subseteq \langle n\rangle$.

We wrote $m\mathbb{Z}$ and $n\mathbb{Z}$ in Chapter 3, but we saw in this chapter that those are equivalent to the principal ideals $\langle m \rangle$ and $\langle n \rangle$, respectively. Now, one of the exercises in this chapter states the following:

Let $f(x), g(x) \in F[x]$. Then $g(x) \mid f(x)$ if and only if $\langle f(x) \rangle \subseteq \langle g(x) \rangle$.

The actual proof is in $\mathbb{R}[x]$, but the argument remains the same in a more general setting of $F[x]$. Note how these statements are saying the same thing in two different rings $\mathbb{Z}$ and $F[x]$.

Exercises

1. Consider the function $\theta : \mathbb{Z}_7[x] \to \mathbb{Z}_7$ where $\theta(f(x)) = f(3)$ for all $f(x) \in \mathbb{Z}_7[x]$. Let $f(x) = 2x^2 + 5x + 4$ and $g(x) = 6x + 1$ in $\mathbb{Z}_7[x]$.
   (a) Verify that $\theta(f(x) + g(x)) = \theta(f(x)) + \theta(g(x))$.
   (b) Verify that $\theta(f(x) \cdot g(x)) = \theta(f(x)) \cdot \theta(g(x))$.

2. Let F be a field and fix $a \in F$. Define the *evaluation map* $\theta : F[x] \to F$ where $\theta(f(x)) = f(a)$ for all $f(x) \in F[x]$. Prove that θ is a ring homomorphism.

3. Consider the ring $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ (with $i = \sqrt{-1}$ so that $i^2 = -1$) and the ring $M(\mathbb{Z})$ of 2×2 matrices with entries in $\mathbb{Z}$. Define $\theta : \mathbb{Z}[i] \to M(\mathbb{Z})$ where $\theta(a + bi) = \left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right]$ for all $a + bi \in \mathbb{Z}[i]$.
   (a) Let $\alpha = 2 + 5i$ and $\beta = 7 - i$ be a pair of elements in $\mathbb{Z}[i]$. Verify that
   $$\theta(\alpha + \beta) = \theta(\alpha) + \theta(\beta) \quad \text{and} \quad \theta(\alpha \cdot \beta) = \theta(\alpha) \cdot \theta(\beta).$$
   (b) Show that θ is a ring homomorphism.
   (c) Find all elements in the kernel $\ker \theta$.
   (d) Describe all elements in the image $\operatorname{im} \theta$.

4. Consider the following subset of $M(\mathbb{Z}_{10})$: $S = \left\{ \left[\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right] \mid a, b \in \mathbb{Z}_{10} \right\}$. In Theorem 26.21, we saw that S is a subring of $M(\mathbb{Z}_{10})$. Define a function $\varphi : S \to \mathbb{Z}_{10}$ such that $\varphi\left(\left[\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right]\right) = a - b$ for all $\left[\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right] \in S$.
   (a) Show that φ is a ring homomorphism.
   (b) Find 5 elements in the kernel $\ker \varphi$.
   (c) Describe all elements in $\ker \varphi$.
   (d) Verify that $K = \ker \varphi$ satisfies the product absorption property:
   If $\rho \in S$ (the domain) and $\alpha \in K$, then $\rho \cdot \alpha \in K$.
   (e) Find all elements in the image $\operatorname{im} \varphi$. Is φ onto? Why or why not?

5. Define a function $\theta : M(\mathbb{Z}_{10}) \to \mathbb{Z}_{10}$ where $\theta\left(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\right) = a + d$ for all $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in M(\mathbb{Z}_{10})$. Determine whether or not θ is a ring homomorphism. If it is, prove it. If not, provide a counterexample.

6. Consider the ring $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$. Define $\theta : \mathbb{Z}_3[i] \to \mathbb{Z}_3[i]$ where $\theta(a+bi) = a-bi$ for all $a+bi \in \mathbb{Z}_3[i]$. In Example 31.6, we showed that θ preserves addition.
   (a) Show that θ preserves multiplication, so that it's a ring homomorphism.
   (b) Show that θ is a bijection, so that it's a ring isomorphism.

7. Consider the function $\theta : \mathbb{Z}_2[x] \to \mathbb{Z}_2[x]$ where $\theta(f(x)) = f(x)^2$ for all $f(x) \in \mathbb{Z}_2[x]$. In Example 31.8, we showed that θ is a ring homomorphism.
   (a) Find all elements in the kernel $\ker\theta$.
   (b) Describe all elements in the image $\operatorname{im}\theta$.

8. Consider the function $\theta : \mathbb{Z}_p[x] \to \mathbb{Z}_p[x]$ where $\theta(f(x)) = f(x)^p$ for all $f(x) \in \mathbb{Z}_p[x]$.
   (a) Prove that θ is a ring homomorphism when $p = 3$.
   (b) Repeat part (a) with $p = 5$.
   (c) Repeat part (a) with $p = 7$.
   (d) Repeat part (a) with any prime p.

9. Consider the field $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. (See Example 27.15.) Define $\theta : \mathbb{Q}(\sqrt{2}) \to \mathbb{Q}(\sqrt{2})$ where $\theta(a + b\sqrt{2}) = a - b\sqrt{2}$ for all $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Prove that θ is a ring isomorphism.

10. In each case, A is an additive subgroup of ring R. Determine whether or not A satisfies the product absorption property (so that A is an ideal of R). If it does, prove it. If it doesn't, give a counterexample showing how A does *not* satisfy product absorption.
    (a) $R = \mathbb{R}$; $A = \mathbb{Z}$.
    (b) $R = \mathbb{Z}[i]$; $A = \{n + ni \mid n \in \mathbb{Z}\}$. (For the definition of $\mathbb{Z}[i]$, see Exercise #3.)
    (c) $R = \mathbb{Z}[i]$; $A = \{a + bi \mid a, b \in \mathbb{Z}, b \text{ is even}\}$.
    (d) $R = \mathbb{Z}[x]$; A is the set of all polynomials with every coefficient divisible by 5.
    (e) $R = \mathbb{Z}_7[x]$; A is the set of all polynomials with constant term equaling 0.

11. Let A be the set of all polynomials in $\mathbb{Z}[x]$ with an even constant term. (See Example 31.31.)
    (a) Prove that A is an ideal of $\mathbb{Z}[x]$.
    (b) Anita claims $A = \langle 2\rangle$, where $\langle 2\rangle = \{2 \cdot f(x) \mid f(x) \in \mathbb{Z}[x]\}$. Do you agree or disagree? Explain.
    (c) Elizabeth claims $A = \langle x\rangle$, where $\langle x\rangle = \{x \cdot f(x) \mid f(x) \in \mathbb{Z}[x]\}$. Agree or disagree? Explain.

    (This exercise and Exercises #12 and #13 below are referenced in Example 35.7. This exercise is also referenced in Chapter 37, Exercise #13.)

12. Let A be the set of all polynomials in $\mathbb{Z}[x]$ with an even constant term. Prove that A is *not* a principal ideal; i.e., there does *not* exist any $\alpha(x) \in \mathbb{Z}[x]$ such that $A = \langle\alpha(x)\rangle$.

    Hint: Suppose for contradiction that $A = \langle\alpha(x)\rangle$ for some $\alpha(x) \in \mathbb{Z}[x]$. That means every element of A is a (polynomial) multiple of $\alpha(x)$. Then what conclusions can you make about $\alpha(x)$?

13. Let A be the set of all polynomials in $\mathbb{Z}[x]$ with an even constant term. Prove that $A = \langle x, 2\rangle$, where $\langle x, 2\rangle = \{x \cdot f(x) + 2 \cdot g(x) \mid f(x), g(x) \in \mathbb{Z}[x]\}$. (This exercise is also referenced in Chapter 36, Exercise #21.)

14. Prove Theorem 31.30.

15. **Prove:** Let $\theta : R \to S$ be a ring homomorphism. Suppose B is an ideal of S, and define the set $A = \{r \in R \mid \theta(r) \in B\}$. Then A is an ideal of R.

16. **Prove:** Let $\theta : R \to S$ be an *onto* ring homomorphism. Suppose A is an ideal of R, and define the set $B = \{\theta(a) \mid a \in A\}$. Then B is an ideal of S.

17. Let $\theta : R \to S$ be an *onto* ring homomorphism.

    (a) **Prove:** $\theta(1_R) = 1_S$; i.e., θ maps the multiplicative identity of R to that of S.

    (b) Using an example, show how the statement in part (a) is false if θ is not onto.

18. **Prove:** Let $\theta : R \to S$ be a ring homomorphism. Suppose that the only nilpotent elements of $\ker\theta$ and S are 0_R and 0_S, respectively. Then 0_R is the only nilpotent element of R.

    Hint: Suppose $a \in R$ is nilpotent and show that $a = 0_R$.

19. **Prove:** Let $\theta : R \to S$ be a ring homomorphism, where R is a field (and S is a ring). If θ is *not* one-to-one, then θ is trivial; i.e., $\theta(r) = 0_S$ for all $r \in R$.

20. **Prove:** Let $f(x), g(x) \in \mathbb{R}[x]$. Then $g(x) \mid f(x)$ if and only if $\langle f(x)\rangle \subseteq \langle g(x)\rangle$. (This exercise is referenced in Chapter 36, Exercise #3.)

21. Let I and J be ideals of a ring R. Prove each of the following:

    (a) $I + J = \{i + j \mid i \in I,\ j \in J\}$ is an ideal of R. (This is the statement of Theorem 37.3.)

    (b) $I \cap J = \{r \in R \mid r \in I \text{ and } r \in J\}$ is an ideal of R. (This exercise is referenced in Chapter 35, Exercise #15.)

22. **Prove:** Let R be a commutative ring. Then the only ideals of R are $\{0\}$ and R itself if and only if R is a field.

23. Let S be the subset of $\mathbb{R}[x]$ containing polynomials whose sum of coefficients equals zero. For example, $f(x) = 5x^4 - 3x^2 - 6x + 4 \in S$ because $5 + (-3) + (-6) + 4 = 0$. Prove that $S = \langle x - 1\rangle$, i.e., the principal ideal generated by $x - 1$.

    Note: You must show both $S \subseteq \langle x - 1\rangle$ and $\langle x - 1\rangle \subseteq S$.

24. Complete the proof of Theorem 31.33 by showing the set inclusion $A \subseteq \langle d\rangle$.

25. Prove Theorem 31.34.

26. Let $X \subseteq R$ be a non-empty subset of a ring R, and define $A = \{a \in R \mid a \cdot x = 0 \text{ for all } x \in X\}$.

    (a) Let $R = \mathbb{Z}_{16}$ and $X = \{2, 4, 8\}$. Find the set A.

    (b) Let $R = \mathbb{Z}_{16}$ and $X = \{4, 8\}$. Find the set A.

    (c) Let $R = \mathbb{Z}_{16}$ and $X = \{8\}$. Find the set A.

    (d) Set A is called the *annihilator* of set X. Why do you think so?

    (e) **Prove:** A is an ideal of R.

32

Introduction to Quotient Rings

In this chapter, we begin our study of *quotient rings*, which will be the focus of the rest of the textbook. As you can imagine, quotient rings are like quotient groups, but with two operations instead of just one. Here is a quick recap of quotient groups. Given an additive group G and a *normal subgroup* H, we formed the quotient group G/H which contains the cosets $g + H$ where $g \in G$. (**Note:** An additive group is always commutative; thus its subgroups are always normal.) We add a pair of cosets $a+H$ and $b + H$ using coset addition, i.e., by adding every element of $a + H$ to every element of $b+H$. But we also found a convenient shortcut; namely: $(a+H)+(b+H) = (a+b)+H$.

Our development of quotient rings will take a similar path. Given a commutative ring R and an *ideal* A, we'll form the quotient ring R/A consisting of the cosets $r + A$ where $r \in R$. (**Note:** We use additive cosets to take advantage of the fact that every ring is an additive group.) We'll learn how to add and multiply cosets in this new setting and the role that the ideals play in making everything fit together. We'll also revisit the First Isomorphism Theorem (see Chapter 25) and develop an analogous theorem for rings.

32.1 From a quotient group to a quotient ring

We begin by reviewing the notion of *cosets* from group theory. (**Suggestion:** Re-read Section 21.3 for a quick refresher.) Consider the *additive* group $\mathbb{Z}$ and its subgroup $5\mathbb{Z}$. Then the cosets of $5\mathbb{Z}$ have the form $a + 5\mathbb{Z}$ where $a \in \mathbb{Z}$. Examples of such cosets include $4{,}378 + 5\mathbb{Z}$, $85 + 5\mathbb{Z}$, and $-23 + 5\mathbb{Z}$.

Recall that different coset representatives can generate the same coset. As we'll see below, we have the equality of cosets $4{,}378 + 5\mathbb{Z} = 3 + 5\mathbb{Z}$ in $\mathbb{Z}/5\mathbb{Z}$, even though their coset representatives are different; i.e., $4{,}378 \neq 3$ in $\mathbb{Z}$. Below are two important properties to help us determine when cosets are equal:

- $a + 5\mathbb{Z} = b + 5\mathbb{Z} \iff a - b \in 5\mathbb{Z}$.
- $a + 5\mathbb{Z} = 0 + 5\mathbb{Z} \iff a \in 5\mathbb{Z}$, which is a special case of the above property with $b = 0$.

We can "reduce" cosets using these properties. For instance, the division algorithm says $4{,}378 = 5 \cdot 873 + 3$, so that $4{,}378 - 3 = 5 \cdot 873 \in 5\mathbb{Z}$. Thus, we conclude $4{,}378 + 5\mathbb{Z} = 3 + 5\mathbb{Z}$. Likewise, $85 + 5\mathbb{Z} = 0 + 5\mathbb{Z}$, because $85 = 5 \cdot 17 \in 5\mathbb{Z}$. (We'll leave it up to you as an exercise to reduce $-23 + 5\mathbb{Z}$.) Therefore, there are only five distinct cosets of $5\mathbb{Z}$, and these cosets form the *quotient group*

$$\mathbb{Z}/5\mathbb{Z} = \{0 + 5\mathbb{Z},\ 1 + 5\mathbb{Z},\ 2 + 5\mathbb{Z},\ 3 + 5\mathbb{Z},\ 4 + 5\mathbb{Z}\}$$

under the operation of coset addition. Coset addition is a rather tedious process (see Section 32.2 for details), but we found a convenient shortcut; namely: $(a + 5\mathbb{Z}) + (b + 5\mathbb{Z}) = (a + b) + 5\mathbb{Z}$. For example,

$$(3 + 5\mathbb{Z}) + (4 + 5\mathbb{Z}) = (3 + 4) + 5\mathbb{Z} = 7 + 5\mathbb{Z} = 2 + 5\mathbb{Z}.$$

We've seen that $\mathbb{Z}/5\mathbb{Z}$ and $\mathbb{Z}_5$ are isomorphic as additive groups via the correspondence $a + 5\mathbb{Z} \leftrightarrow a \pmod 5$. But $\mathbb{Z}_5$ is also a ring, equipped with the multiplication operation: $3 \cdot 4 = 2$ in $\mathbb{Z}_5$, for instance. We wish to define multiplication of cosets in $\mathbb{Z}/5\mathbb{Z}$ and turn it into a ring as well. The key is to adapt the shortcut for coset multiplication. For example, $(3 + 5\mathbb{Z}) \cdot (4 + 5\mathbb{Z}) = 3 \cdot 4 + 5\mathbb{Z} = 12 + 5\mathbb{Z} = 2 + 5\mathbb{Z}$. More generally, we define coset multiplication in $\mathbb{Z}/5\mathbb{Z}$ by $(a + 5\mathbb{Z}) \cdot (b + 5\mathbb{Z}) = a \cdot b + 5\mathbb{Z}$.

Below are some calculations in $\mathbb{Z}/5\mathbb{Z}$:

- $(1 + 5\mathbb{Z}) \cdot (4 + 5\mathbb{Z}) = 4 + 5\mathbb{Z}$ and $(4 + 5\mathbb{Z}) \cdot (1 + 5\mathbb{Z}) = 4 + 5\mathbb{Z}$. We'll leave it to you to verify that $1 + 5\mathbb{Z}$ is the multiplicative identity of $\mathbb{Z}/5\mathbb{Z}$.
- $(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z}) = 6 + 5\mathbb{Z} = 1 + 5\mathbb{Z}$ and $(3 + 5\mathbb{Z}) \cdot (2 + 5\mathbb{Z}) = 6 + 5\mathbb{Z} = 1 + 5\mathbb{Z}$. Thus, $2 + 5\mathbb{Z}$ and $3 + 5\mathbb{Z}$ are multiplicative inverses of each other with

$$(2 + 5\mathbb{Z})^{-1} = 3 + 5\mathbb{Z} \quad \text{and} \quad (3 + 5\mathbb{Z})^{-1} = 2 + 5\mathbb{Z}.$$

In an exercise, you'll verify that coset multiplication in $\mathbb{Z}/5\mathbb{Z}$ satisfies the ring properties that are outlined in Definition 26.2. (**Note:** We already know that $\mathbb{Z}/5\mathbb{Z}$ is an additive group, so that its coset addition satisfies the ring properties.) Thus we conclude that $\mathbb{Z}/5\mathbb{Z}$ is a *quotient ring* under coset addition and multiplication. We also have a ring isomorphism $\mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}_5$ via the correspondence $a + 5\mathbb{Z} \leftrightarrow a \pmod 5$. In fact, we can generalize to obtain a ring isomorphism $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$. We'll leave the details up to you as an exercise.

Definition 32.1 (Quotient ring)**.** Let R be a commutative ring, and let A be an ideal of R. The set of cosets $R/A = \{r + A \mid r \in R\}$ is a *quotient ring* under the operations

$$(r + A) + (s + A) = (r + s) + A \quad \text{and} \quad (r + A) \cdot (s + A) = r \cdot s + A.$$

Remark. In the next section, we'll see the role played by an *ideal* in a quotient ring.

32.2 Role of an ideal in a quotient ring

Consider again the additive group $\mathbb{Z}$ and its subgroup $5\mathbb{Z}$. Here are a couple of cosets of $5\mathbb{Z}$:

$$2 + 5\mathbb{Z} = \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\},$$
$$3 + 5\mathbb{Z} = \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\}.$$

Recall that the coset $2 + 5\mathbb{Z}$ is obtained by adding 2 to each element of $5\mathbb{Z}$, and the coset $3 + 5\mathbb{Z}$ is found similarly. The coset sum is defined as follows:

$$(2 + 5\mathbb{Z}) + (3 + 5\mathbb{Z}) = \{\alpha + \beta \mid \alpha \in 2 + 5\mathbb{Z},\ \beta \in\ 3 + 5\mathbb{Z}\};$$

i.e., add every element of $2 + 5\mathbb{Z}$ to those of $3 + 5\mathbb{Z}$. We learned that, since $5\mathbb{Z}$ is a *normal subgroup* of $\mathbb{Z}$, coset addition satisfies a shortcut; namely, $(2+5\mathbb{Z})+(3+5\mathbb{Z}) = (2+3)+5\mathbb{Z}$.

The shortcut is a convenient property, not the definition of coset addition. As an analogy, we *define* U_{13} to be the set of elements of $\mathbb{Z}_{13}$ with multiplicative inverses. According to this definition, we know that $6 \in U_{13}$, because $6 \cdot 11 = 1 \pmod{13}$. But we also found a convenient *property*; namely: $6 \in U_{13}$ because 6 and 13 are relatively prime (i.e., they don't share a common factor other than 1).

Similarly to coset addition, we might define coset multiplication as follows:

$$(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z}) = \{\alpha \cdot \beta \mid \alpha \in 2 + 5\mathbb{Z},\ \beta \in\ 3 + 5\mathbb{Z}\};$$

i.e., multiply every element of $2 + 5\mathbb{Z}$ with those of $3 + 5\mathbb{Z}$. Let's see what happens:

$$\begin{aligned}
&(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z}) \\
&\quad = \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\} \cdot \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\} \\
&\quad = \{\ldots, -39, -34, -24, -14, -9, -4, 6, 16, 21, 26, 36, 46, \ldots\} \ \leftarrow \text{call this set } S.
\end{aligned}$$

We would hope that $(2+5\mathbb{Z})\cdot(3+5\mathbb{Z}) = 2\cdot 3+5\mathbb{Z} = 1+5\mathbb{Z}$, but the resulting set S seems to be missing some elements that are in $1 + 5\mathbb{Z}$. In particular, 1 is not in S. This isn't surprising, since the only ways to obtain 1 as a product of two integers are $1 = 1 \cdot 1$ and $1 = -1 \cdot -1$; but $2+5\mathbb{Z}$ and $3+5\mathbb{Z}$ do not contain 1 or -1. A similar reasoning explains why, for instance, 11 and 31 are not in S, even though they are in $1 + 5\mathbb{Z}$. Therefore, if we define coset multiplication like we defined coset addition, then $(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z})$ does *not* equal $1 + 5\mathbb{Z}$ (i.e., the shortcut fails). Even worse, the resulting set S isn't even a coset of $5\mathbb{Z}$.

How can we salvage the situation? **We *define* coset multiplication using the shortcut.** Thus, we define $(2+5\mathbb{Z})\cdot(3+5\mathbb{Z})$ to be the coset $2\cdot 3+5\mathbb{Z}$, which simplifies to $1 + 5\mathbb{Z}$. Now the shortcut isn't some convenient *property*. Instead, it's built into the *definition* of coset multiplication.

This seems simple enough on the surface, but we must attend to an important technicality. The product $(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z}) = 2 \cdot 3 + 5\mathbb{Z}$ depends only on the coset representatives 2 and 3, rather than the elements in those cosets. But we also have $2 + 5\mathbb{Z} = 22 + 5\mathbb{Z}$ and $3 + 5\mathbb{Z} = 58 + 5\mathbb{Z}$. So, does $(22 + 5\mathbb{Z}) \cdot (58 + 5\mathbb{Z})$ yield the same product as $(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z})$? Let's find out:

$$\begin{aligned}
(22 + 5\mathbb{Z}) \cdot (58 + 5\mathbb{Z}) &= 22 \cdot 58 + 5\mathbb{Z} && \leftarrow \text{definition of coset multiplication} \\
&= 1{,}276 + 5\mathbb{Z} && \leftarrow \text{because } 22 \cdot 58 = 1{,}276 \\
&= 1 + 5\mathbb{Z} && \leftarrow \text{because } 1{,}276 - 1 \in 5\mathbb{Z}.
\end{aligned}$$

As we had hoped, $(22+5\mathbb{Z})\cdot(58+5\mathbb{Z})$ does equal $(2+5\mathbb{Z})\cdot(3+5\mathbb{Z})$, since both products equal $1 + 5\mathbb{Z}$.

Let's dig deeper into why $(22 + 5\mathbb{Z}) \cdot (58 + 5\mathbb{Z})$ *should* equal $(2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z})$. We have $22 + 5\mathbb{Z} = 2 + 5\mathbb{Z}$, because $22 - 2 = 5 \cdot 4 \in 5\mathbb{Z}$. Likewise, $58 + 5\mathbb{Z} = 3 + 5\mathbb{Z}$, since

$58 - 3 = 5 \cdot 11 \in 5\mathbb{Z}$. Writing $22 = 2 + 5 \cdot 4$ and $58 = 3 + 5 \cdot 11$, we have

$$\begin{aligned}
&(22 + 5\mathbb{Z}) \cdot (58 + 5\mathbb{Z}) \\
&\quad = ((2 + 5 \cdot 4) + 5\mathbb{Z}) \cdot ((3 + 5 \cdot 11) + 5\mathbb{Z}) \\
&\quad = (2 + 5 \cdot 4) \cdot (3 + 5 \cdot 11) + 5\mathbb{Z} && \leftarrow \text{definition of coset multiplication} \\
&\quad = (\mathbf{2 \cdot 3} + 2 \cdot (5 \cdot 11) + (5 \cdot 4) \cdot 3 + (5 \cdot 4) \cdot (5 \cdot 11)) + 5\mathbb{Z} \\
&\quad = \mathbf{2 \cdot 3} + 5\mathbb{Z} && \leftarrow \text{see below for reason behind this step} \\
&\quad = (2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z}) && \leftarrow \text{definition of coset multiplication again.}
\end{aligned}$$

We have the coset equality $(\mathbf{2 \cdot 3} + 2 \cdot (5 \cdot 11) + (5 \cdot 4) \cdot 3 + (5 \cdot 4) \cdot (5 \cdot 11)) + 5\mathbb{Z} = \mathbf{2 \cdot 3} + 5\mathbb{Z}$, because the difference of the coset representatives is $2 \cdot (5 \cdot 11) + (5 \cdot 4) \cdot 3 + (5 \cdot 4) \cdot (5 \cdot 11)$, which is in $5\mathbb{Z}$, since each term in the sum contains a 5. Therefore, $(22+5\mathbb{Z}) \cdot (58+5\mathbb{Z}) = (2 + 5\mathbb{Z}) \cdot (3 + 5\mathbb{Z})$ as desired.

The following theorem shows that this technicality is always satisfied.

Theorem 32.2 (Coset multiplication in $\mathbb{Z}/5\mathbb{Z}$ is well-defined)**.** *Suppose* $a+5\mathbb{Z} = c+5\mathbb{Z}$ *and* $b+5\mathbb{Z} = d+5\mathbb{Z}$ *for some* $a, b, c, d \in \mathbb{Z}$*. Then* $(a+5\mathbb{Z}) \cdot (b+5\mathbb{Z}) = (c+5\mathbb{Z}) \cdot (d+5\mathbb{Z})$.

PROOF. Since $a + 5\mathbb{Z} = c + 5\mathbb{Z}$, we have $a - c \in 5\mathbb{Z}$, which implies $a = c + x$ for some $x \in 5\mathbb{Z}$. Similarly, $b + 5\mathbb{Z} = d + 5\mathbb{Z}$ implies $b = d + y$ for some $y \in 5\mathbb{Z}$. Thus, we have

$$\begin{aligned}
(a + 5\mathbb{Z}) \cdot (b + 5\mathbb{Z}) &= ((c + x) + 5\mathbb{Z}) \cdot ((d + y) + 5\mathbb{Z}) \\
&= (c + x) \cdot (d + y) + 5\mathbb{Z} \\
&= c \cdot d + (c \cdot y + x \cdot d + x \cdot y) + 5\mathbb{Z}.
\end{aligned}$$

Consider the element $c \cdot y + x \cdot d + x \cdot y$. **Since $5\mathbb{Z}$ is an ideal of $\mathbb{Z}$, it satisfies product absorption and is closed under addition.** Since $x, y \in 5\mathbb{Z}$, we have $c \cdot y,\ x \cdot d,\ x \cdot y \in 5\mathbb{Z}$ by product absorption. Then $c \cdot y + x \cdot d + x \cdot y \in 5\mathbb{Z}$, because $5\mathbb{Z}$ is closed. Therefore, we have

$$\begin{aligned}
(a + 5\mathbb{Z}) \cdot (b + 5\mathbb{Z}) &= c \cdot d + (c \cdot y + x \cdot d + x \cdot y) + 5\mathbb{Z} \\
&= c \cdot d + 5\mathbb{Z} && \leftarrow \text{since } c \cdot y + x \cdot d + x \cdot y \in 5\mathbb{Z} \\
&= (c + 5\mathbb{Z}) \cdot (d + 5\mathbb{Z})
\end{aligned}$$

so that $(a + 5\mathbb{Z}) \cdot (b + 5\mathbb{Z}) = (c + 5\mathbb{Z}) \cdot (d + 5\mathbb{Z})$ as desired. ■

Proof know-how. In the above proof, we could have written $5i$ and $5j$ for the elements of $5\mathbb{Z}$ (with $i, j \in \mathbb{Z}$), rather than x and y. However, by writing them as x and y, we can better highlight the fact that they are elements of an ideal. For instance, we concluded that $c \cdot y,\ x \cdot d,\ x \cdot y \in 5\mathbb{Z}$, not because they're all multiples of 5 (which they are), but because of the product absorption property of $5\mathbb{Z}$.

Remark. Theorem 32.2 says coset multiplication in $\mathbb{Z}/5\mathbb{Z}$ is *well-defined*; i.e., the product does not depend on the choice of coset representatives. The key, as shown in the proof, is that $5\mathbb{Z}$ is an *ideal* of $\mathbb{Z}$.

Example 32.3 (Non-example)**.** Here is a fictitious rule for multiplying cosets that is *not* well-defined. In $\mathbb{Z}/5\mathbb{Z}$, suppose we had defined $(a + 5\mathbb{Z}) \cdot (b + 5\mathbb{Z}) = a^{|b|} + 5\mathbb{Z}$, where $|b|$ denotes the absolute value of b. (**Note:** We use $a^{|b|}$ rather than a^b, which

may not be an integer if b is negative.) For instance, we have $(2+5\mathbb{Z})\cdot(3+5\mathbb{Z}) = 2^{|3|}+5\mathbb{Z} = 8+5\mathbb{Z} = 3+5\mathbb{Z}$. As before, $(22+5\mathbb{Z})\cdot(58+5\mathbb{Z})$ should yield the same product as $(2+5\mathbb{Z})\cdot(3+5\mathbb{Z})$, since $22+5\mathbb{Z} = 2+5\mathbb{Z}$ and $58+5\mathbb{Z} = 3+5\mathbb{Z}$. But we have

$$(22+5\mathbb{Z})\cdot(58+5\mathbb{Z}) = 22^{|58|}+5\mathbb{Z} = (\text{a huge integer ending in } 4)+5\mathbb{Z} = 4+5\mathbb{Z},$$

so that $(22+5\mathbb{Z})\cdot(58+5\mathbb{Z}) \neq (2+5\mathbb{Z})\cdot(3+5\mathbb{Z})$. Since this product depends on the choice of coset representatives, we say that it's *not* well-defined and thus is not a valid multiplication rule for cosets.

Below is a generalization of Theorem 32.2, whose proof is left for you as an exercise. Observe that from its conclusion $a\cdot b+A = c\cdot d+A$, we can deduce that $(a+A)\cdot(b+A) = (c+A)\cdot(d+A)$, so that coset multiplication in R/A is well-defined.

Theorem 32.4. *Let R be a commutative ring, and let A be an ideal of R. Suppose $a+A = c+A$ and $b+A = d+A$ for some $a, b, c, d \in R$. Then $a\cdot b+A = c\cdot d+A$.*

32.3 Quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$

Consider the polynomial ring $\mathbb{Z}_3[x]$ and a subset $\langle x^2\rangle = \{x^2\cdot q(x) \mid q(x)\in\mathbb{Z}_3[x]\}$, i.e., the *principal ideal* generated by x^2. Note that $\langle x^2\rangle$ is the set of all multiples of x^2. Then the quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$ contains cosets of the form $a(x)+\langle x^2\rangle$ where $a(x)\in\mathbb{Z}_3[x]$.

Example 32.5. Consider the polynomials $a(x) = x^9+2x^5+x^3$ and $b(x) = 2x^7+x^4+2x$ in $\mathbb{Z}_3[x]$. We have $a(x) = x^2\cdot(x^7+2x^3+x)$, so that $a(x)$ is an element of $\langle x^2\rangle$. However, $b(x)$ is *not* contained in $\langle x^2\rangle$, since $b(x)$ is not a polynomial multiple of x^2.

Example 32.6. Let $\alpha(x), \beta(x)\in\mathbb{Z}_3[x]$ where

$$\alpha(x) = x^9+2x^5+x^3+\mathbf{2x+1} \quad \text{and} \quad \beta(x) = \mathbf{2x+1}.$$

Then $\alpha(x)\neq\beta(x)$ in $\mathbb{Z}_3[x]$; i.e., they're *different* polynomials. However, in $\mathbb{Z}_3[x]/\langle x^2\rangle$, their cosets are the *same*; i.e., $\alpha(x)+\langle x^2\rangle = \beta(x)+\langle x^2\rangle$, because $\alpha(x)-\beta(x) = x^9+2x^5+x^3 = x^2\cdot(x^7+2x^3+x)\in\langle x^2\rangle$.

Example 32.7. Let $\alpha(x) = x^{13}+2x^{10}+2x^5+\boldsymbol{x}+\mathbf{2}\in\mathbb{Z}_3[x]$. We will "reduce" the coset $\alpha(x)+\langle x^2\rangle$; i.e., we'll find $\beta(x)\in\mathbb{Z}_3[x]$ of the smallest degree such that $\alpha(x)+\langle x^2\rangle = \beta(x)+\langle x^2\rangle$. Let $\beta(x) = \boldsymbol{x}+\mathbf{2}$. Then

$$\alpha(x)-\beta(x) = x^{13}+2x^{10}+2x^5 = x^2\cdot(x^{11}+2x^8+2x^3)\in\langle x^2\rangle.$$

Therefore,

$$(x^{13}+2x^{10}+2x^5+x+2)+\langle x^2\rangle = (x+2)+\langle x^2\rangle.$$

Note how this is analogous to reducing $4{,}378+5\mathbb{Z} = 3+5\mathbb{Z}$ in the quotient ring $\mathbb{Z}/5\mathbb{Z}$.

Generalizing from the above examples, suppose $f(x)\in\mathbb{Z}_3[x]$ with

$$f(x) = a_nx^n+a_{n-1}x^{n-1}+\cdots+a_2x^2+\boldsymbol{a_1x+a_0},$$

where $a_i\in\mathbb{Z}_3$. Then $f(x)+\langle x^2\rangle = (\boldsymbol{a_1x+a_0})+\langle x^2\rangle$, because

$$f(x)-(a_1x+a_0) = a_nx^n+a_{n-1}x^{n-1}+\cdots+a_2x^2\in\langle x^2\rangle.$$

Hence, every element of $\mathbb{Z}_3[x]/\langle x^2\rangle$ can be reduced to $(ax+b)+\langle x^2\rangle$, where $a, b\in\mathbb{Z}_3$.

Thus, we have

$$\begin{aligned}\mathbb{Z}_3[x]/\langle x^2\rangle &= \{(ax+b)+\langle x^2\rangle \mid a,\, b \in \mathbb{Z}_3\}\\ &= \{0+\langle x^2\rangle,\ 1+\langle x^2\rangle,\ 2+\langle x^2\rangle,\\ &\quad x+\langle x^2\rangle,\ (x+1)+\langle x^2\rangle,\ (x+2)+\langle x^2\rangle,\\ &\quad 2x+\langle x^2\rangle,\ (2x+1)+\langle x^2\rangle,\ (2x+2)+\langle x^2\rangle\},\end{aligned}$$

where the additive and multiplicative identities are $0+\langle x^2\rangle$ and $1+\langle x^2\rangle$, respectively.

We verify that these 9 cosets are distinct. Suppose for contradiction that $(2x+1)+\langle x^2\rangle = (x+2)+\langle x^2\rangle$. Then $(2x+1)-(x+2) \in \langle x^2\rangle$; i.e., $x-1 \in \langle x^2\rangle$. Thus, $x-1$ is a multiple of x^2, which is a contradiction. In an exercise, you'll show that any pair of the 9 cosets above are, in fact, distinct.

Example 32.8. The calculation below shows $(x+1)+\langle x^2\rangle$ and $(2x+1)+\langle x^2\rangle$ are multiplicative inverses of each other, and hence units in the quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$:

$$\begin{aligned}((x+1)+\langle x^2\rangle)\cdot((2x+1)+\langle x^2\rangle) &= (x+1)\cdot(2x+1)+\langle x^2\rangle\\ &= (2x^2+\mathbf{3}\boldsymbol{x}+1)+\langle x^2\rangle\\ &= (2x^2+1)+\langle x^2\rangle \qquad \longleftarrow \text{since } 3x=0 \text{ in } \mathbb{Z}_3[x]\\ &= 1+\langle x^2\rangle.\end{aligned}$$

Example 32.9. The element $x+\langle x^2\rangle \in \mathbb{Z}_3[x]/\langle x^2\rangle$ is non-zero; i.e., it does not equal $0+\langle x^2\rangle$, because x is not a multiple of x^2. Moreover, we have $(x+\langle x^2\rangle)\cdot(x+\langle x^2\rangle) = x^2+\langle x^2\rangle = 0+\langle x^2\rangle$. Thus, $x+\langle x^2\rangle$ is a zero divisor in $\mathbb{Z}_3[x]/\langle x^2\rangle$, which means it is not a unit. And since not every non-zero element of $\mathbb{Z}_3[x]/\langle x^2\rangle$ is a unit, we conclude that $\mathbb{Z}_3[x]/\langle x^2\rangle$ is *not* a field.

32.4 First Isomorphism Theorem for rings

Example 32.10. Consider the ring homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}_5$ where $\varphi(a) = a \pmod 5$ for all $a \in \mathbb{Z}$. The kernel and image of φ are $K = 5\mathbb{Z}$ and $\operatorname{im}\varphi = \mathbb{Z}_5$, respectively. (See Example 31.16.) In Section 32.1, we observed a ring isomorphism $\mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}_5$ via the correspondence $a+5\mathbb{Z} \leftrightarrow a \pmod 5$. Putting these pieces together, we conclude that $\mathbb{Z}/K \cong \operatorname{im}\varphi$, where $a+K \in \mathbb{Z}/K$ corresponds to $\varphi(a) \in \operatorname{im}\varphi$.

Here is the generalization of Example 32.10, namely the First Isomorphism Theorem for rings.

Theorem 32.11 (First Isomorphism Theorem for rings)**.** *Let $\theta : R \to S$ be a ring homomorphism with $K = \ker\theta$. Then there is a ring isomorphism $R/K \cong \operatorname{im}\theta$, where $r+K \in R/K$ corresponds to $\theta(r) \in \operatorname{im}\theta$.*

Example 32.12. Consider the ring homomorphism $\theta : \mathbb{R}[x] \to \mathbb{R}$ where $\theta(f(x)) = f(2)$ for all $f(x) \in \mathbb{R}[x]$. The kernel and image of θ are $K = \langle x-2\rangle$ and $\operatorname{im}\theta = \mathbb{R}$, respectively. (See Example 31.18.) Therefore, the First Isomorphism Theorem for rings implies that there is a ring isomorphism $\mathbb{R}[x]/\langle x-2\rangle \cong \mathbb{R}$, where $f(x)+\langle x-2\rangle \in \mathbb{R}[x]/\langle x-2\rangle$ corresponds to $f(2) \in \mathbb{R}$.

Example 32.13. Consider the ring $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. In fact, $\mathbb{Q}(\sqrt{2})$ is a field, as we saw in Example 27.15. Define the function $\theta : \mathbb{Q}[x] \to \mathbb{Q}(\sqrt{2})$ where $\theta(f(x)) = f(\sqrt{2})$ for all $f(x) \in \mathbb{Q}[x]$. For instance, suppose $f(x) = 4x^3 - 5x^2 + 6x + 2 \in \mathbb{Q}[x]$. Then

$$\theta(f(x)) = f(\sqrt{2}) = 4(\sqrt{2})^3 - 5(\sqrt{2})^2 + 6\sqrt{2} + 2 = 8\sqrt{2} - 10 + 6\sqrt{2} + 2 = -8 + 14\sqrt{2},$$

so that $\theta(f(x)) = -8 + 14\sqrt{2}$, which is an element of $\mathbb{Q}(\sqrt{2})$. In an exercise, you'll show that θ is a ring homomorphism with kernel $\langle x^2 - 2\rangle$ and image $\mathbb{Q}(\sqrt{2})$ (i.e., θ is onto). Thus the First Isomorphism Theorem implies that $\mathbb{Q}[x]/\langle x^2 - 2\rangle \cong \mathbb{Q}(\sqrt{2})$, where $f(x) + \langle x^2 - 2\rangle \in \mathbb{Q}[x]/\langle x^2 - 2\rangle$ corresponds to $f(\sqrt{2}) \in \mathbb{Q}(\sqrt{2})$.

Big picture stuff

We continue to find structural similarities between the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$, where F is a field. In this chapter, we studied the quotient rings $\mathbb{Z}/5\mathbb{Z}$ and $\mathbb{Z}_3[x]/\langle x^2\rangle$. In fact, we can write $5\mathbb{Z}$ as the principal ideal $\langle 5\rangle$, i.e., the set of all multiples of 5. In Example 32.6, we saw how to reduce cosets:

> Consider $\alpha(x) = x^9 + 2x^5 + x^3 + \mathbf{2x + 1}$ and $\beta(x) = \mathbf{2x + 1}$ in $\mathbb{Z}_3[x]$. Then $\alpha(x) \neq \beta(x)$ in $\mathbb{Z}_3[x]$; i.e., they're *different* polynomials. However, in $\mathbb{Z}_3[x]/\langle x^2\rangle$, their cosets are the *same*; i.e., $\alpha(x) + \langle x^2\rangle = \beta(x) + \langle x^2\rangle$, because $\alpha(x) - \beta(x) = x^9 + 2x^5 + x^3 = x^2 \cdot (x^7 + 2x^3 + x) \in \langle x^2\rangle$.

This is analogous to how $4{,}378 \neq 3$ in $\mathbb{Z}$; i.e., they're *different* integers. However, in $\mathbb{Z}/\langle 5\rangle$, their cosets are the *same*; i.e., $4{,}378 + \langle 5\rangle = 3 + \langle 5\rangle$, because $4{,}378 - 3 = 5 \cdot 873 \in \langle 5\rangle$.

We also found that $\mathbb{Z}/\langle 5\rangle$ is a field, but $\mathbb{Z}_3[x]/\langle x^2\rangle$ is *not* a field. In the upcoming chapters, we'll study the factors (pun intended!) that lead to this distinction.

Exercises

1. "Reduce" each element of $\mathbb{Z}/5\mathbb{Z}$ by writing it in the form $a + 5\mathbb{Z}$ where $a = 0$, 1, 2, 3, or 4.
   (a) $-23 + 5\mathbb{Z}$.
   (b) $172 + 5\mathbb{Z}$.
   (c) $1{,}437 + 5\mathbb{Z}$.
   (d) $-2{,}908 + 5\mathbb{Z}$.

2. For a fixed positive integer n, explain why the distinct elements of $\mathbb{Z}/n\mathbb{Z}$ are
   $$\mathbb{Z}/n\mathbb{Z} = \{0 + n\mathbb{Z},\ 1 + n\mathbb{Z},\ 2 + n\mathbb{Z},\ \ldots,\ (n-1) + n\mathbb{Z}\}.$$
   Be sure to explain why these cosets are *distinct* from each other.

3. Coset multiplication in $\mathbb{Z}/5\mathbb{Z}$ is defined by $(a + 5\mathbb{Z}) \cdot (b + 5\mathbb{Z}) = a \cdot b + 5\mathbb{Z}$. Verify that this operation satisfies the ring properties (6) through (9) that are outlined in Definition 26.2.
   Note: For property (9) (the distributive law), you'll also need coset addition. Feel free to use the coset addition shortcut; namely: $(a + 5\mathbb{Z}) + (b + 5\mathbb{Z}) = (a + b) + 5\mathbb{Z}$.

4. In Section 32.2, we attempted to define coset multiplication by saying that $(2+5\mathbb{Z})\cdot(3+5\mathbb{Z})$ should equal the set $S = \{\alpha \cdot \beta \mid \alpha \in 2+5\mathbb{Z},\ \beta \in 3+5\mathbb{Z}\}$. Unfortunately, we found that $S \neq 1+5\mathbb{Z}$. Explain why S isn't even a coset of $5\mathbb{Z}$.

5. Here is a fictitious rule for multiplying cosets in $\mathbb{Z}/5\mathbb{Z}$: $(a+5\mathbb{Z})\cdot(b+5\mathbb{Z}) = |a\cdot b|+5\mathbb{Z}$, where $|a \cdot b|$ denotes the absolute value of $a \cdot b$. Explain why this rule is not well-defined.

6. Fix $x^2 \in \mathbb{Z}_3[x]$ and define $\langle x^2\rangle = \{x^2 \cdot q(x) \mid q(x) \in \mathbb{Z}_3[x]\}$.
   (a) Explain why $2x^7 + x^5 + 2x^4$ is contained in $\langle x^2\rangle$.
   (b) Explain why $x^6 + 2x$ is *not* contained in $\langle x^2\rangle$.
   (c) List three more elements of $\mathbb{Z}_3[x]$ that are contained in $\langle x^2\rangle$.
   (d) List three more elements of $\mathbb{Z}_3[x]$ that are *not* contained in $\langle x^2\rangle$.

7. Let $\alpha(x), \beta(x) \in \mathbb{Z}_3[x]$, where $\alpha(x) = 2x^7 + x^5 + 2x^4 + x + 2$ and $\beta(x) = x + 2$. Explain why the cosets $\alpha(x) + \langle x^2\rangle$ and $\beta(x) + \langle x^2\rangle$ are equal in the quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$.

8. (a) Let $f(x) = x^5 + 2x^2 + x + 2 \in \mathbb{Z}_3[x]$. Find $g(x) \in \mathbb{Z}_3[x]$ of smallest degree such that
   $$f(x) + \langle x^2\rangle = g(x) + \langle x^2\rangle.$$
   (b) Repeat part (a), this time with $f(x) = 2x^9 + x^6 + 2x^3 + 1 \in \mathbb{Z}_3[x]$.

9. In the quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$, do the following:
   (a) Explain why $(x+1) + \langle x^2\rangle \neq 2x + \langle x^2\rangle$.
   (b) Explain why $(x+2) + \langle x^2\rangle \neq (2x+2) + \langle x^2\rangle$.
   (c) **Prove:** If $a_1x+a_0 \neq b_1x+b_0$ in $\mathbb{Z}_3[x]$, then $(a_1x+a_0)+\langle x^2\rangle \neq (b_1x+b_0)+\langle x^2\rangle$ in $\mathbb{Z}_3[x]/\langle x^2\rangle$.
   Note: This confirms that $\mathbb{Z}_3[x]/\langle x^2\rangle$ indeed has 9 distinct elements. (See Section 32.3.)

10. Let F be a field and fix $g(x) \in F[x]$. Suppose $p(x), q(x) \in F[x]$ with $\deg p(x)$, $\deg q(x) < \deg g(x)$. Prove that if $p(x) \neq q(x)$ in $F[x]$, then $p(x) + \langle g(x)\rangle \neq q(x) + \langle g(x)\rangle$ in $F[x]/\langle g(x)\rangle$.
    Note: This is a generalization of Exercise #9(c) above.

11. The quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$ is finite, so every non-zero element must be either a unit or a zero divisor (i.e., it cannot be neither). For each non-zero element $a(x) + \langle x^2\rangle$, determine if it's a unit or a zero divisor. Moreover, find a non-zero $b(x) + \langle x^2\rangle$ such that the following hold:
    - For a unit: $(a(x) + \langle x^2\rangle) \cdot (b(x) + \langle x^2\rangle) = 1 + \langle x^2\rangle$.
    - For a zero divisor: $(a(x) + \langle x^2\rangle) \cdot (b(x) + \langle x^2\rangle) = 0 + \langle x^2\rangle$.

12. Consider the quotient ring $\mathbb{Z}_2[x]/\langle x^2\rangle$. (**Be careful:** The coefficient ring is $\mathbb{Z}_2$, not $\mathbb{Z}_3$.)
    (a) Find all distinct elements of $\mathbb{Z}_2[x]/\langle x^2\rangle$, i.e., distinct cosets $f(x) + \langle x^2\rangle$ where $f(x) \in \mathbb{Z}_2[x]$.
    (b) Construct the addition and multiplication tables for $\mathbb{Z}_2[x]/\langle x^2\rangle$.
    (c) Is $\mathbb{Z}_2[x]/\langle x^2\rangle$ a field? Why or why not?

13. (a) How many distinct elements does $\mathbb{Z}_5[x]/\langle x^2\rangle$ contain?
    (b) How about $\mathbb{Z}_7[x]/\langle x^2\rangle$?
    (c) How about $\mathbb{Z}_{11}[x]/\langle x^2\rangle$?
    (d) How about $\mathbb{Z}_{101}[x]/\langle x^2\rangle$?
    (e) How about $\mathbb{Z}_p[x]/\langle x^2\rangle$, where p is prime?

14. Repeat Exercise #13, but replace $\langle x^2\rangle$ with $\langle x^3\rangle$.

15. Repeat Exercise #13, but replace $\langle x^2\rangle$ with $\langle x^n\rangle$, where n is a positive integer.

16. Consider the element $(x^2+1)+\langle x^4+x^2\rangle$ in the quotient ring $\mathbb{R}[x]/\langle x^4+x^2\rangle$.
    (a) Explain why $(x^2+1)+\langle x^4+x^2\rangle \neq 0+\langle x^4+x^2\rangle$ in $\mathbb{R}[x]/\langle x^4+x^2\rangle$.
    (b) Verify that $(x^2+1)+\langle x^4+x^2\rangle$ is an idempotent in $\mathbb{R}[x]/\langle x^4+x^2\rangle$.
    Recall: A ring element a is called an *idempotent* if $a^2 = a$.

17. Find a field F and $f(x), g(x) \in F[x]$ where $f(x)+\langle g(x)\rangle$ is a nontrivial idempotent in $F[x]/\langle g(x)\rangle$.
    Note: In other words, come up with your own Exercise #16.

18. Consider the element $(x+1)+\langle x^4+1\rangle$ in the quotient ring $\mathbb{Z}_2[x]/\langle x^4+1\rangle$.
    (a) Explain why $(x+1)+\langle x^4+1\rangle \neq 0+\langle x^4+1\rangle$ in $\mathbb{Z}_2[x]/\langle x^4+1\rangle$.
    (b) Verify that $(x+1)+\langle x^4+1\rangle$ is nilpotent in $\mathbb{Z}_2[x]/\langle x^4+1\rangle$.
    Recall: A ring element r is called a *nilpotent* element if $r^n = 0$ for some positive integer n.

19. Find a field F and $f(x), g(x) \in F[x]$ where $f(x)+\langle g(x)\rangle$ is a non-zero nilpotent element in $F[x]/\langle g(x)\rangle$.
    Note: In other words, come up with your own Exercise #18.

20. Prove Theorem 32.4.

21. Use the First Isomorphism Theorem for rings to prove that there is a ring isomorphism $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$. (Compare with Chapter 25, Exercise #5. Also, this exercise is referenced in Section 35.1.)

22. Consider $\theta : \mathbb{Q}[x] \to \mathbb{Q}(\sqrt{2})$ where $\theta(f(x)) = f(\sqrt{2})$ for all $f(x) \in \mathbb{Q}[x]$. (See Example 32.13.) Prove each of the following claims about θ:
    (a) θ is a ring homomorphism.
    (b) $\ker\theta = \langle x^2-2\rangle$.
    (c) $\operatorname{im}\theta = \mathbb{Q}(\sqrt{2})$.

23. Consider the ring homomorphism $\lambda : \mathbb{Z}_{10} \to \mathbb{Z}_{30}$ where $\lambda(a) = 6a$ for all $a \in \mathbb{Z}_{10}$.
    (a) Find the kernel K and image of λ. (See Example 31.17 to verify your answer.)
    (b) The First Isomorphism Theorem for rings says that $\mathbb{Z}_{10}/K \cong \operatorname{im}\lambda$, where $a+K \in \mathbb{Z}_{10}/K$ corresponds to $\lambda(a) \in \operatorname{im}\lambda$. Create addition and multiplication tables for $\mathbb{Z}_{10}/K$ and for $\operatorname{im}\lambda$ (thus, 4 tables total) to verify this ring isomorphism.

33

Quotient Ring $\mathbb{Z}_7[x]/\langle x^2-1\rangle$

We will continue our investigation of quotient rings involving polynomials. In particular, we will consider the polynomial ring $\mathbb{Z}_7[x]$ and a subset $\langle x^2-1\rangle = \{(x^2-1)\cdot q(x) \mid q(x) \in \mathbb{Z}_7[x]\}$, i.e., the principal ideal generated by x^2-1, or, equivalently, the set of all multiples of x^2-1. The quotient ring $\mathbb{Z}_7[x]/\langle x^2-1\rangle$ contains cosets of the form $a(x)+\langle x^2-1\rangle$ where $a(x) \in \mathbb{Z}_7[x]$.

On the surface, it may seem that there are infinitely many distinct cosets $a(x)+\langle x^2-1\rangle$, since there are infinitely many distinct polynomials $a(x) \in \mathbb{Z}_7[x]$. However, different coset representatives can generate the same coset. Thus, we can "reduce" a coset such as $(4x^5+2x^3+4x+1)+\langle x^2-1\rangle$ to something simpler, namely $(3x+1)+\langle x^2-1\rangle$. (We'll see this in Example 33.3.) The focus of this chapter is on this reduction process. We will formalize the method that we implicitly used in Chapter 32 when we reduced the cosets in $\mathbb{Z}_3[x]/\langle x^2\rangle$. We will also learn a new reduction method that is much more efficient.

Moreover, we will begin to address the following: Let F be a field and fix $g(x) \in F[x]$. Then how can we determine whether or not the quotient ring $F[x]/\langle g(x)\rangle$ is a field? This is an overarching question that will guide our study for the remainder of this textbook.

33.1 Division algorithm revisited

Example 33.1 (Example 32.7 revisited)**.** Consider the quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$ whose elements are cosets of the form $a(x)+\langle x^2\rangle$ where $a(x) \in \mathbb{Z}_3[x]$. Let $\alpha(x) = x^{13} + 2x^{10} + 2x^5 + \boldsymbol{x} + \boldsymbol{2} \in \mathbb{Z}_3[x]$. To reduce the coset $\alpha(x)+\langle x^2\rangle$, let $\beta(x) = \boldsymbol{x} + \boldsymbol{2}$. Then $\alpha(x)-\beta(x) = x^{13}+2x^{10}+2x^5 = x^2\cdot(x^{11}+2x^8+2x^3)$, so that $\alpha(x)-\beta(x)$ is a multiple of x^2. Thus, $\alpha(x)-\beta(x) \in \langle x^2\rangle$, and so $\alpha(x)+\langle x^2\rangle = \beta(x)+\langle x^2\rangle$.

In the above example, we can quickly find $\beta(x)$ such that $\alpha(x)-\beta(x)$ is a multiple of x^2. This is why reducing a coset in $\mathbb{Z}_3[x]/\langle x^2\rangle$ is a relatively simple task. In this chapter, we'll work with the quotient ring $\mathbb{Z}_7[x]/\langle x^2-1\rangle$. Reducing cosets in this setting is more complicated, because determining whether or not a polynomial is a multiple of x^2-1 is more difficult.

Example 33.2. Let $a(x) = 3x^3 + 5x^2 + 4x + 2$ and $b(x) = 5x^3 + 5x^2 + 4$ be elements of $\mathbb{Z}_7[x]$. Long division calculations below show that $a(x) = (x^2-1)\cdot(3x+5)$ and $b(x) = (x^2-1)\cdot(5x+5)+(5x+2)$. Therefore, $a(x) \in \langle x^2-1\rangle$; i.e., $a(x)$ is a multiple of x^2-1, since there is no remainder when dividing $a(x)$ by x^2-1. But $b(x) \notin \langle x^2-1\rangle$; i.e., $b(x)$ is *not* a multiple of x^2-1, since there is a non-zero remainder $5x+2$.

$$\begin{array}{rl} & \boxed{3x+5} \longleftarrow \text{quotient} \\ x^2-1\,\overline{)\,3x^3+5x^2+4x+2} & \\ -(3x^3 \quad -3x) & \\ \hline 5x^2+2 & \\ -(5x^2 \quad -5) & \\ \hline \boxed{0} \longleftarrow \text{remainder} & \end{array} \qquad \begin{array}{rl} & \boxed{5x+5} \longleftarrow \text{quotient} \\ x^2-1\,\overline{)\,5x^3+5x^2 \quad +4} & \\ -(5x^3 \quad -5x) & \\ \hline 5x^2+5x+4 & \\ -(5x^2 \quad -5) & \\ \hline \boxed{5x+2} \longleftarrow \text{remainder} & \end{array}$$

Thus, we can use long division to determine whether a polynomial $f(x) \in \mathbb{Z}_7[x]$ is a multiple of x^2-1. The next example shows how long division can be used to reduce a coset in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$.

Example 33.3. Let $\alpha(x) = 4x^5+2x^3+4x+1 \in \mathbb{Z}_7[x]$. We will reduce $\alpha(x)+\langle x^2-1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$, i.e., find $\beta(x) \in \mathbb{Z}_7[x]$ of the smallest degree such that $\alpha(x)+\langle x^2-1\rangle = \beta(x)+\langle x^2-1\rangle$. First, we use long division to divide $\alpha(x)$ by x^2-1:

$$\begin{array}{rl} & \boxed{4x^3+6x} \longleftarrow \text{quotient} \\ x^2-1\,\overline{)\,4x^5+2x^3+4x+1} & \\ -(4x^5-4x^3) & \\ \hline 6x^3+4x & \\ -(6x^3-6x) & \\ \hline \boxed{3x+1} \longleftarrow \text{remainder} & \end{array}$$

We can also use a software such as *Mathematica* to perform long division, as shown below:

```
In[1]:= a = 4 x^5 + 2 x^3 + 4 x + 1;

In[2]:= PolynomialMod[PolynomialQuotientRemainder[a, x^2 - 1, x], 7]

Out[2]= {6 x + 4 x^3, 1 + 3 x}
```

Hence, $\alpha(x) = (x^2-1)\cdot(4x^3+6x)+(3x+1)$. Note how the degree of the remainder $3x+1$ is less than the degree of the divisor x^2-1. We then have $\alpha(x)+\langle x^2-1\rangle = (3x+1)+\langle x^2-1\rangle$, because $\alpha(x)-(3x+1) = (x^2-1)\cdot(4x^3+6x) \in \langle x^2-1\rangle$.

We emphasize that $\alpha(x) \neq 3x+1$ in $\mathbb{Z}_7[x]$; they're *different* polynomials. But they generate the *same* coset in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$; i.e., $\alpha(x)+\langle x^2-1\rangle = (3x+1)+\langle x^2-1\rangle$. As mentioned in Chapter 32, this is analogous to how $4{,}378 \neq 3$ in $\mathbb{Z}$; i.e., they're *different* integers. However, their cosets are the *same* in $\mathbb{Z}/5\mathbb{Z}$; i.e., $4{,}378+5\mathbb{Z} = 3+5\mathbb{Z}$, because $4{,}378-3 = 5\cdot 873 \in 5\mathbb{Z}$.

More generally, let $f(x) \in \mathbb{Z}_7[x]$. By the division algorithm, there exist $q(x), r(x) \in \mathbb{Z}_7[x]$ such that $f(x) = (x^2-1)\cdot q(x)+r(x)$ with either $r(x) = 0$ or $\deg r(x) < \deg(x^2-1)$. (In practice, we can find such $q(x)$ and $r(x)$ using long division.) Thus $r(x)$ has the form $r(x) = ax + b$ where $a, b \in \mathbb{Z}_7$. Then, $f(x) + \langle x^2 - 1\rangle = r(x) + \langle x^2 - 1\rangle$, because $f(x) - r(x) = (x^2 - 1) \cdot q(x) \in \langle x^2 - 1\rangle$. Thus, $f(x) + \langle x^2 - 1\rangle$ can be reduced to $(ax + b) + \langle x^2 - 1\rangle$ where $a, b \in \mathbb{Z}_7$.

Hence we have

$$\mathbb{Z}_7[x]/\langle x^2 - 1\rangle = \{(ax + b) + \langle x^2 - 1\rangle \mid a, b \in \mathbb{Z}_7\}$$

so that the quotient ring contains 7^2 or 49 cosets. Let's verify that they are actually distinct. Suppose for contradiction that $(5x + 3) + \langle x^2 - 1\rangle = (3x + 6) + \langle x^2 - 1\rangle$. Then we have $(5x + 3) - (3x + 6) \in \langle x^2 - 1\rangle$; i.e., $2x + 4 \in \langle x^2 - 1\rangle$. This implies $2x + 4$ is a multiple of $x^2 - 1$, which is a contradiction. A similar argument can be used to show that any pair of the 49 cosets above are, in fact, distinct.

In Chapter 32, we worked with the quotient ring $\mathbb{Z}_3[x]/\langle x^2\rangle$. For instance, we saw that

$$(2x^7 + x^5 + 2x^4 + \mathbf{2x + 1}) + \langle x^2\rangle = (\mathbf{2x + 1}) + \langle x^2\rangle,$$

since $(2x^7+x^5+2x^4+\mathbf{2x+1})-(\mathbf{2x+1}) = 2x^7+x^5+2x^4 = x^2\cdot(2x^5+x^3+2x^2) \in \langle x^2\rangle$. It seems much simpler to reduce an element in $\mathbb{Z}_3[x]/\langle x^2\rangle$. However, we were (implicitly) using the same approach taken in Example 33.3. Let $f(x) = 2x^7 + x^5 + 2x^4 + \mathbf{2x + 1}$. Then the division algorithm in $\mathbb{Z}_3[x]$ yields $f(x) = x^2 \cdot (2x^5 + x^3 + 2x^2) + (\mathbf{2x + 1})$, where the degree of the remainder $2x+1$ is less than the degree of the divisor x^2. Then $f(x) + \langle x^2\rangle = (2x + 1) + \langle x^2\rangle$ as before.

Example 33.4. Consider the following product in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$:

$$\begin{aligned}((x + 2) + \langle x^2 - 1\rangle) \cdot ((2x + 3) + \langle x^2 - 1\rangle) &= (x + 2) \cdot (2x + 3) + \langle x^2 - 1\rangle \\ &= (2x^2 + 7x + 6) + \langle x^2 - 1\rangle \\ &= (2x^2 + 6) + \langle x^2 - 1\rangle.\end{aligned}$$

To reduce $(2x^2 + 6) + \langle x^2 - 1\rangle$, note that $2x^2 + 6 = (x^2 - 1) \cdot 2 + 1$, where the quotient $q(x) = 2$ and remainder $r(x) = 1$ can be obtained via long division. Thus,

$$(2x^2 + 6) + \langle x^2 - 1\rangle = 1 + \langle x^2 - 1\rangle,$$

since $(2x^2 + 6) - 1 = (x^2 - 1) \cdot 2 \in \langle x^2 - 1\rangle$. Hence,

$$((x + 2) + \langle x^2 - 1\rangle) \cdot ((2x + 3) + \langle x^2 - 1\rangle) = 1 + \langle x^2 - 1\rangle,$$

so that $(x + 2) + \langle x^2 - 1\rangle$ and $(2x + 3) + \langle x^2 - 1\rangle$ are multiplicative inverses of each other, and thus units in the quotient ring $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$.

Example 33.5. Consider the following product in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$:

$$\begin{aligned}((x + 1) + \langle x^2 - 1\rangle) \cdot ((x - 1) + \langle x^2 - 1\rangle) &= (x + 1) \cdot (x - 1) + \langle x^2 - 1\rangle \\ &= (x^2 - 1) + \langle x^2 - 1\rangle \\ &= 0 + \langle x^2\rangle.\end{aligned}$$

Moreover, $(x + 1) + \langle x^2 - 1\rangle$ and $(x - 1) + \langle x^2 - 1\rangle$ are non-zero; i.e., they do not equal $0 + \langle x^2 - 1\rangle$, because neither $x + 1$ nor $x - 1$ is a multiple of $x^2 - 1$. Therefore, $(x+1)+\langle x^2-1\rangle$ and $(x-1)+\langle x^2-1\rangle$ are zero divisors in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$, which means they are not units. And since not every non-zero element of $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$ is a unit, $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$ is *not* a field.

33.2 Another way to reduce in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$

We saw how to use long division to reduce elements in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$. But long division can be a tedious process. Thus, we introduce a more efficient approach. Observe that $x^2+\langle x^2-1\rangle = 1+\langle x^2-1\rangle$ because $x^2-1 \in \langle x^2-1\rangle$. Using this equality of cosets, we have

$$\begin{aligned}(3\cdot x^2+5)+\langle x^2-1\rangle &= (3\cdot x^2+\langle x^2-1\rangle)+(5+\langle x^2-1\rangle)\\ &= (3+\langle x^2-1\rangle)\cdot(\boldsymbol{x^2}+\langle \boldsymbol{x^2-1}\rangle)+(5+\langle x^2-1\rangle)\\ &= (3+\langle x^2-1\rangle)\cdot(\mathbf{1}+\langle \boldsymbol{x^2-1}\rangle)+(5+\langle x^2-1\rangle)\\ &= (3\cdot 1+\langle x^2-1\rangle)+(5+\langle x^2-1\rangle)\\ &= (3\cdot 1+5)+\langle x^2-1\rangle.\end{aligned}$$

Thus $(3\cdot \boldsymbol{x^2}+5)+\langle x^2-1\rangle = (3\cdot \mathbf{1}+5)+\langle x^2-1\rangle$. In other words, we can treat x^2 and 1 to be the same *as coset representatives*. Whenever we see x^2 as part of a coset representative, we can replace it with 1.

Example 33.6. Consider $f(x)=4x^5+2x^3+4x+1 \in \mathbb{Z}_7[x]$ from Example 33.3. To apply the reduction method described above to $f(x)+\langle x^2-1\rangle$, we must isolate x^2 in the coset representative. Below, we write the term $4x^5$ as $4\cdot x^2\cdot x^2\cdot x$, so that each x^2 can be replaced with 1. We do likewise with $2x^3$:

$$\begin{aligned}f(x)+\langle x^2-1\rangle &= (4x^5+2x^3+4x+1)+\langle x^2-1\rangle\\ &= (4\cdot \boldsymbol{x^2}\cdot \boldsymbol{x^2}\cdot x+2\cdot \boldsymbol{x^2}\cdot x+4x+1)+\langle x^2-1\rangle\\ &= (4\cdot \mathbf{1}\cdot \mathbf{1}\cdot x+2\cdot \mathbf{1}\cdot x+4x+1)+\langle x^2-1\rangle\\ &= (4x+2x+4x+1)+\langle x^2-1\rangle\\ &= (3x+1)+\langle x^2-1\rangle,\end{aligned}$$

which is what we found before.

Remark. When using this new reduction method, we must be careful to replace x^2 with 1 *only in coset representatives*. Although we have the coset equality $x^2+\langle x^2-1\rangle = 1+\langle x^2-1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$, the polynomials x^2 and 1 are different in $\mathbb{Z}_7[x]$.

Example 33.7. Since $\mathbb{Z}_7[x]/\langle x^2-1\rangle$ is finite, every non-zero element must be either a unit or a zero divisor (i.e., it cannot be neither). Let's consider the element $(3x+5)+\langle x^2-1\rangle \in \mathbb{Z}_7[x]/\langle x^2-1\rangle$ and determine which type it is. Consider the following product in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$:

$$\begin{aligned}((3x+5)+\langle x^2-1\rangle)\cdot((ax+b)+\langle x^2-1\rangle) &= (3x+5)\cdot(ax+b)+\langle x^2-1\rangle\\ &= (3a\cdot \boldsymbol{x^2}+(3b+5a)x+5b)+\langle x^2-1\rangle\\ &= (3a\cdot \mathbf{1}+(3b+5a)x+5b)+\langle x^2-1\rangle\\ &= ((3b+5a)x+(3a+5b))+\langle x^2-1\rangle.\end{aligned}$$

Here, we used the reduction technique of replacing x^2 by 1 in the coset representative. We'll now find $a, b \in \mathbb{Z}_7$ such that one of the following is true:

- $((3b + 5a)x + (3a + 5b)) + \langle x^2 - 1\rangle = 1 + \langle x^2 - 1\rangle$; hence $(3x + 5) + \langle x^2 - 1\rangle$ is a unit.
- $((3b + 5a)x + (3a + 5b)) + \langle x^2 - 1\rangle = 0 + \langle x^2 - 1\rangle$; hence $(3x + 5) + \langle x^2 - 1\rangle$ is a zero divisor.

Let's consider the first possibility. As the coset representatives $(3b + 5a)x + (3a + 5b)$ and 1 both have degree less than 2, the cosets $((3b + 5a)x + (3a + 5b)) + \langle x^2 - 1\rangle$ and $1 + \langle x^2 - 1\rangle$ are equal in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$ precisely when their coset representatives are equal in $\mathbb{Z}_7[x]$. We'll seek $a, b \in \mathbb{Z}_7$ such that $(3b + 5a)x + (3a + 5b) = 1$ or $(3b + 5a)x + (3a + 5b) = 0x + 1$ in $\mathbb{Z}_7[x]$. Setting the coefficients equal, we obtain the system of equations $3b + 5a = 0$ and $3a + 5b = 1$. Solving this in $\mathbb{Z}_7$, we find that $a = 2$ and $b = 6$. Thus, $(3x + 5) + \langle x^2 - 1\rangle$ is a unit in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$ whose multiplicative inverse is $(2x + 6) + \langle x^2 - 1\rangle$.

For the second possibility, we obtain the system of equations $3b + 5a = 0$ and $3a + 5b = 0$. Solving this in $\mathbb{Z}_7$, we find that $a = 0$ and $b = 0$. Therefore $(ax+b)+\langle x^2-1\rangle$ is the zero element and we have

$$((3x + 5) + \langle x^2 - 1\rangle) \cdot (0 + \langle x^2 - 1\rangle) = 0 + \langle x^2 - 1\rangle,$$

which, while valid, does *not* imply that $(3x + 5) + \langle x^2 - 1\rangle$ is a zero divisor.

33.3 $F[x]/\langle g(x)\rangle$ is *not* a field

Thus far, we have seen two examples of quotient rings involving polynomials:

- $\mathbb{Z}_3[x]/\langle x^2\rangle$ in Section 32.3.
- $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$ in this chapter.

Each of these has a zero divisor and thus is *not* a field. Moreover, we found these zero divisors by factoring. With $\mathbb{Z}_3[x]/\langle x^2\rangle$, we consider the factorization $x^2 = x \cdot x$. The element $x + \langle x^2\rangle \in \mathbb{Z}_3[x]/\langle x^2\rangle$ is non-zero, because x is not a multiple of x^2. And we have

$$(x + \langle x^2\rangle) \cdot (x + \langle x^2\rangle) = x^2 + \langle x^2\rangle = 0 + \langle x^2\rangle,$$

so that $x + \langle x^2\rangle$ is a zero divisor in $\mathbb{Z}_3[x]/\langle x^2\rangle$. With $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$, we factor $x^2 - 1 = (x + 1) \cdot (x - 1)$ and proceed similarly to find zero divisors $(x + 1) + \langle x^2 - 1\rangle$ and $(x - 1) + \langle x^2 - 1\rangle$. Here is the generalization, whose proof is left for you as an exercise.

Theorem 33.8. *Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is factorable, then $F[x]/\langle g(x)\rangle$ is* not *a field.*

Example 33.9. Let $g(x) = x^3-8 \in \mathbb{Z}_{13}[x]$ and consider the quotient ring $\mathbb{Z}_{13}[x]/\langle g(x)\rangle$. We note that $g(2) = 2^3 - 8 = 0$. Then by the factor theorem, $g(x) = (x - 2) \cdot q(x)$ for some $q(x) \in \mathbb{Z}_{13}[x]$ of degree 2. Thus Theorem 33.8 implies that $\mathbb{Z}_{13}[x]/\langle g(x)\rangle$ is *not* a field. In fact, we have

$$((x-2) + \langle g(x)\rangle) \cdot (q(x) + \langle g(x)\rangle) = (x-2) \cdot q(x) + \langle g(x)\rangle = g(x) + \langle g(x)\rangle = 0 + \langle g(x)\rangle,$$

so that $(x - 2) + \langle g(x)\rangle$ and $q(x) + \langle g(x)\rangle$ are zero divisors in $\mathbb{Z}_{13}[x]/\langle g(x)\rangle$. Both are non-zero in $\mathbb{Z}_{13}[x]/\langle g(x)\rangle$, since $x-2$ and $q(x)$ are not multiples of $g(x)$, as their degrees are less than that of $g(x)$.

33.4 $F[x]/\langle g(x)\rangle$ is a field

Based on Theorem 33.8, we might conjecture the following:

> *Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

Below are some examples that support the conjecture.

Example 33.10. Consider the polynomial $g(x) = x^2-2$, which is unfactorable in $\mathbb{Q}[x]$. (See Example 30.21.) Do we also know that the quotient ring $\mathbb{Q}[x]/\langle x^2-2\rangle$ is a field? In Example 32.13, we used the First Isomorphism Theorem to derive a ring isomorphism $\mathbb{Q}[x]/\langle x^2-2\rangle \cong \mathbb{Q}(\sqrt{2})$. Moreover, $\mathbb{Q}(\sqrt{2})$ is a field. (See Chapter 27, Exercise #14(c).) Therefore $\mathbb{Q}[x]/\langle x^2-2\rangle$ is a field, as desired.

Example 33.11. Consider $g(x) = x^2+1 \in \mathbb{Z}_3[x]$. We have $g(0) = 1$, $g(1) = 2$, and $g(2) = 2$, so that $g(x)$ does not have a root in $\mathbb{Z}_3$. And since $\deg g(x) = 2$, we conclude that $g(x)$ is unfactorable in $\mathbb{Z}_3[x]$ by Theorem 30.19. In the exercises, you'll verify that the quotient ring $\mathbb{Z}_3[x]/\langle x^2+1\rangle$ is indeed a field.

Example 33.12. Consider $g(x) = x^2+1 \in \mathbb{Z}_7[x]$. In the exercises, you'll verify the following:

- $g(x)$ does not have a root in $\mathbb{Z}_7$, so that $g(x)$ is unfactorable in $\mathbb{Z}_7[x]$.
- The quotient ring $\mathbb{Z}_7[x]/\langle x^2+1\rangle$ is a field.

Hence this example also supports our conjecture.

Big picture stuff

Theorem 33.8 above says that if $g(x)$ is factorable, then $F[x]/\langle g(x)\rangle$ is *not* a field. Let's see through an example whether an analogous claim can be made with the integers. We consider $n = 12$, which is composite (i.e., a factorable integer). Then can we conclude that $\mathbb{Z}/\langle 12\rangle$ (or, equivalently, $\mathbb{Z}/12\mathbb{Z}$) is *not* a field? Since 12 factors as $12 = 3\cdot 4$, we have

$$(3+\langle 12\rangle)\cdot(4+\langle 12\rangle) = 3\cdot 4+\langle 12\rangle = 12+\langle 12\rangle = 0+\langle 12\rangle.$$

Note that $3+\langle 12\rangle$ and $4+\langle 12\rangle$ are non-zero in $\mathbb{Z}/\langle 12\rangle$, as 3 and 4 are not multiples of 12. Thus, $3+\langle 12\rangle$ and $4+\langle 12\rangle$ are zero divisors in $\mathbb{Z}/\langle 12\rangle$, which implies that $\mathbb{Z}/\langle 12\rangle$ is *not* a field. In other words, the same argument used in $F[x]/\langle g(x)\rangle$ can be applied to $\mathbb{Z}/\langle n\rangle$ as well.

Exercises

1. (a) Consider $a(x), b(x) \in \mathbb{Z}_7[x]$ where $a(x) = 3x^3+5x^2+4x+2$ and $b(x) = 5x^3+5x^2+4$. Recall from Example 33.2 that $a(x)$ is a multiple of x^2-1, while $b(x)$ is not. Verify that *both* 1 and -1 in $\mathbb{Z}_7$ are roots of $a(x)$, while this is not the case for $b(x)$.

   (b) Find two more polynomials $f(x), g(x) \in \mathbb{Z}_7[x]$ such that $f(x)$ is a multiple of x^2-1, while $g(x)$ is not. Verify that *both* 1 and -1 in $\mathbb{Z}_7$ are roots of $f(x)$, while this is not the case for $g(x)$.

2. **Prove:** Let $f(x) \in \mathbb{Z}_7[x]$. Then $f(x) \in \langle x^2 - 1\rangle$ if and only if $f(1) = 0$ and $f(-1) = 0$.

3. **Prove:** Let $f(x) \in F[x]$ and $a, b \in F$ where $a \neq b$. Then $(x-a)\cdot(x-b)$ is a factor of $f(x)$ if and only if $f(a) = 0$ and $f(b) = 0$.

   Note: This is a generalization of the statement in Exercise #2.

4. Use the polynomial $f(x) = x^2 - 6x + 8 \in \mathbb{Z}_{15}[x]$ to show how the statement in Exercise #3 is false when the coefficient ring F is *not* a field. (Also see Example 27.11.)

5. Fix $x^2 - 1 \in \mathbb{Z}_7[x]$ and define $\langle x^2 - 1\rangle = \{(x^2 - 1)\cdot q(x) \mid q(x) \in \mathbb{Z}_7[x]\}$.

   (a) Explain why $3x^5 + 3x^3 + 4x^2 + x + 3$ is contained in $\langle x^2 - 1\rangle$.
   (b) Explain why $2x^4 + 5x^3 + 6x^2 + 6x + 4$ is *not* contained in $\langle x^2 - 1\rangle$.
   (c) List three more elements of $\mathbb{Z}_7[x]$ that are contained in $\langle x^2 - 1\rangle$.
   (d) List three more elements of $\mathbb{Z}_7[x]$ that are *not* contained in $\langle x^2 - 1\rangle$.

6. Let $\alpha(x), \beta(x) \in \mathbb{Z}_7[x]$, where $\alpha(x) = 3x^5 + 3x^3 + 4x^2 + 6x + 4$ and $\beta(x) = 5x + 1$. Explain why the cosets $\alpha(x) + \langle x^2 - 1\rangle$ and $\beta(x) + \langle x^2 - 1\rangle$ are equal in the quotient ring $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$.

7. (a) Let $f(x) = 3x^7 + x^3 + 4 \in \mathbb{Z}_7[x]$. Find a polynomial $g(x) \in \mathbb{Z}_7[x]$ of the smallest degree such that $f(x) + \langle x^2 - 1\rangle = g(x) + \langle x^2 - 1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$.
   (b) Repeat part (a), this time with $f(x) = 2x^9 + 5x^7 + 4x^3 + 3 \in \mathbb{Z}_7[x]$.

8. In the quotient ring $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$, do the following:

   (a) Explain why $(5x + 2) + \langle x^2 - 1\rangle \neq (2x + 4) + \langle x^2 - 1\rangle$.
   (b) Explain why $(x + 2) + \langle x^2 - 1\rangle \neq (6x + 2) + \langle x^2 - 1\rangle$.
   (c) **Prove:** If $a_1x + a_0 \neq b_1x + b_0$ in $\mathbb{Z}_7[x]$, then $(a_1x + a_0) + \langle x^2 - 1\rangle \neq (b_1x + b_0) + \langle x^2 - 1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$.
   Note: This confirms that $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$ indeed has 49 distinct elements.

9. Consider the elements $(4x + 3) + \langle x^2 - 1\rangle$ and $(4x + 2) + \langle x^2 - 1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$. Determine if each is a unit or a zero divisor. Explain how you know.

10. Consider the quotient ring $\mathbb{Z}_{13}[x]/\langle x^3 - 8\rangle$ and let $f(x) = 5x^7 + 10x^6 + 2x^3 + 8x + 4 \in \mathbb{Z}_{13}[x]$.

    (a) Use long division to reduce the coset $f(x) + \langle x^3 - 8\rangle$ in $\mathbb{Z}_{13}[x]/\langle x^3 - 8\rangle$.
    (b) Explain why $x^3 + \langle x^3 - 8\rangle = 8 + \langle x^3 - 8\rangle$ in $\mathbb{Z}_{13}[x]/\langle x^3 - 8\rangle$.
    (c) Use your result in part (b) and the method in Section 33.2 to reduce $f(x) + \langle x^3 - 8\rangle$.

11. How many distinct elements does $\mathbb{Z}_{13}[x]/\langle x^3 - 8\rangle$ contain? Explain how you know.

12. Consider the quotient ring $\mathbb{Z}_{11}[x]/\langle x^3 + 4x + 6\rangle$ and let $f(x) = 5x^7 + 10x^6 + 2x^3 + 8x + 4 \in \mathbb{Z}_{11}[x]$.

    (a) Use long division to reduce the coset $f(x) + \langle x^3 + 4x + 6\rangle$ in $\mathbb{Z}_{11}[x]/\langle x^3 + 4x + 6\rangle$.
    (b) Explain why $x^3 + \langle x^3 + 4x + 6\rangle = (7x + 5) + \langle x^3 + 4x + 6\rangle$ in $\mathbb{Z}_{11}[x]/\langle x^3 + 4x + 6\rangle$.
    (c) Use your result in part (b) and the method in Section 33.2 to reduce $f(x) + \langle x^3 + 4x + 6\rangle$.

13. How many distinct elements does $\mathbb{Z}_{11}[x]/\langle x^3+4x+6\rangle$ contain? Explain how you know.

14. Let $g(x) = x^3+4x+6 \in \mathbb{Z}_{11}[x]$.

    (a) Find a root of $g(x)$ in $\mathbb{Z}_{11}$.

    (b) Show that $\mathbb{Z}_{11}[x]/\langle g(x)\rangle$ is *not* a field by finding zero divisors.

15. Consider the quotient ring $\mathbb{Z}_3[x]/\langle x^2+1\rangle$.

    (a) Find all distinct elements of $\mathbb{Z}_3[x]/\langle x^2+1\rangle$, i.e., distinct cosets $f(x)+\langle x^2+1\rangle$ where $f(x) \in \mathbb{Z}_3[x]$.

    (b) For each non-zero element $a(x)+\langle x^2+1\rangle$, determine if it's a unit or a zero divisor.

    (c) Moreover, find a non-zero $b(x)+\langle x^2+1\rangle$ such that the following are true:

    - For a unit: $(a(x)+\langle x^2+1\rangle)\cdot(b(x)+\langle x^2+1\rangle) = 1+\langle x^2+1\rangle$.
    - For a zero divisor: $(a(x)+\langle x^2+1\rangle)\cdot(b(x)+\langle x^2+1\rangle) = 0+\langle x^2+1\rangle$.

    (d) Is $\mathbb{Z}_3[x]/\langle x^2+1\rangle$ a field? (It should be!) Why or why not?

16. Consider again the quotient ring $R = \mathbb{Z}_3[x]/\langle x^2+1\rangle$. Define $R^* = \{\alpha \in R \mid \alpha \text{ is a unit}\}$ and recall that R^* is a multiplicative group. (See Theorem 26.24.) Verify that R^* is a cyclic group, generated by the element $\alpha = (x+1)+\langle x^2+1\rangle$.

17. Define a function $\theta : \mathbb{Z}_3[i] \to \mathbb{Z}_3[x]/\langle x^2+1\rangle$ where $\theta(a+bi) = (a+bx)+\langle x^2+1\rangle$ for all $a+bi \in \mathbb{Z}_3[i]$. For example, $\theta(1+2i) = (1+2x)+\langle x^2+1\rangle$. Show that θ is a ring isomorphism.

    Note: In Example 27.14, we saw that $\mathbb{Z}_3[i]$ is a field. Therefore, the above isomorphism is another way to verify that $\mathbb{Z}_3[x]/\langle x^2+1\rangle$ is a field.

18. Consider the function $\varphi : \mathbb{Z}_3[x] \to \mathbb{Z}_3[i]$ where $\varphi(f(x)) = f(i)$ for all $f(x) \in \mathbb{Z}_3[x]$. Prove the following:

    (a) φ is a ring homomorphism.

    (b) **(Challenge)** $\ker\varphi = \langle x^2+1\rangle$.

    (c) $\operatorname{im}\varphi = \mathbb{Z}_3[i]$.

    Note: Thus the First Isomorphism Theorem implies that $\mathbb{Z}_3[x]/\langle x^2+1\rangle \cong \mathbb{Z}_3[i]$.

    (This exercise is referenced in Chapter 34, Exercise #17.)

19. Let p be a prime number. Show that $\mathbb{Z}_p[x]/\langle x^2-6x+8\rangle$ is *not* a field.

20. Prove Theorem 33.8.

21. Consider the polynomial $g(x) = x^2+1 \in \mathbb{Z}_7[x]$.

    (a) Verify that $g(x)$ does not have a root in $\mathbb{Z}_7$.

    (b) Explain why $g(x)$ is unfactorable in $\mathbb{Z}_7[x]$.

22. **Prove:** Let $a, b \in \mathbb{Z}_7$. Then $a^2+b^2 = 0$ in $\mathbb{Z}_7$ if and only if $a = 0$ and $b = 0$.

23. Consider the quotient ring $\mathbb{Z}_7[x]/\langle x^2+1\rangle$. Since x^2+1 is unfactorable in $\mathbb{Z}_7[x]$ (see Exercise #21(b)), our conjecture in Section 33.4 would imply that $\mathbb{Z}_7[x]/\langle x^2+1\rangle$ is a field.

    (a) How many distinct elements does $\mathbb{Z}_7[x]/\langle x^2+1\rangle$ contain? Explain how you know.
    (b) Find a multiplicative inverse of the element $(4x+3)+\langle x^2+1\rangle \in \mathbb{Z}_7[x]/\langle x^2+1\rangle$.
    (c) Same as part (b), but with $(2x+5)+\langle x^2+1\rangle$.
    (d) Same as part (b), but with $(6x+1)+\langle x^2+1\rangle$.

24. **(Challenge)** Let $(ax+b)+\langle x^2+1\rangle$ be a non-zero element of $\mathbb{Z}_7[x]/\langle x^2+1\rangle$. Thus, a and b are not both zero, though one of them could be. Find a multiplicative inverse of $(ax+b)+\langle x^2+1\rangle$.

    Hint: Your result in Exercise #22 should help.

25. **(Challenge)** How many units and zero divisors does $\mathbb{Z}_7[x]/\langle x^2-1\rangle$ have? Can you determine the answer *without* actually considering each element of the quotient ring one by one?

34

Quotient Ring $\mathbb{R}[x]/\langle x^2+1\rangle$

We will study another quotient ring involving polynomials, namely $\mathbb{R}[x]/\langle x^2+1\rangle$. However, the highlight of this chapter is an *isomorphism* $\mathbb{R}[x]/\langle x^2+1\rangle \cong \mathbb{C}$ where the coset $(a+bx)+\langle x^2+1\rangle \in \mathbb{R}[x]/\langle x^2+1\rangle$ corresponds to the complex number $a+bi \in \mathbb{C}$. Intuitively, to say that the rings $\mathbb{R}[x]/\langle x^2+1\rangle$ and $\mathbb{C}$ are isomorphic means that they are essentially the same. Thus, a calculation done in one ring must have a corresponding (and equivalent) calculation in the isomorphic ring.

In particular, we'll work on the task of finding the multiplicative inverse of an element in $\mathbb{R}[x]/\langle x^2+1\rangle$, say $(3+4x)+\langle x^2+1\rangle$. This involves a somewhat tedious calculation. Thus, we'll consider the corresponding problem in $\mathbb{C}$ of finding the multiplicative inverse of $3+4i$, which happens to be a much simpler task. Once we find the multiplicative inverse in $\mathbb{C}$, we'll translate the result back to $\mathbb{R}[x]/\langle x^2+1\rangle$ to obtain the desired multiplicative inverse of $(3+4x)+\langle x^2+1\rangle$. Indeed, translating a task in one setting to an equivalent and simpler task in another is a powerful application of an isomorphism.

34.1 Reducing elements in $\mathbb{R}[x]/\langle x^2+1\rangle$

Consider the polynomial ring $\mathbb{R}[x]$ and a subset $\langle x^2+1\rangle = \{(x^2+1)\cdot q(x) \mid q(x) \in \mathbb{R}[x]\}$, i.e., the *principal ideal* generated by x^2+1. Note that $\langle x^2+1\rangle$ is the set of all multiples of x^2+1. Then the quotient ring $\mathbb{R}[x]/\langle x^2+1\rangle$ contains cosets of the form $a(x)+\langle x^2+1\rangle$ where $a(x) \in \mathbb{R}[x]$.

Consider the polynomial $f(x) = 5x^4+x^3-3x^2+4x-3 \in \mathbb{R}[x]$. We will illustrate two different approaches to reduce the coset $f(x)+\langle x^2+1\rangle$ in $\mathbb{R}[x]/\langle x^2+1\rangle$, in other words, to find $g(x) \in \mathbb{R}[x]$ of the smallest degree such that $f(x)+\langle x^2+1\rangle = g(x)+\langle x^2+1\rangle$.

Method #1: Division algorithm. We use long division to divide $f(x)$ by x^2+1, either by hand or using a software such as *Mathematica*. (See Example 33.2.) We find that $f(x) = (x^2+1)\cdot(5x^2+x-8)+(5+3x)$. Note how the degree of the remainder

$5+3x$ is less than the degree of the divisor x^2+1. Therefore,

$$f(x)+\langle x^2+1\rangle = (5+3x)+\langle x^2+1\rangle,$$

because $f(x)-(5+3x) = (x^2+1)\cdot(5x^2+x-8) \in \langle x^2+1\rangle$.

More generally, let $f(x) \in \mathbb{R}[x]$. By the division algorithm, there exist $q(x), r(x) \in \mathbb{R}[x]$ such that $f(x) = (x^2+1)\cdot q(x)+r(x)$ with either $r(x) = 0$ or $\deg r(x) < \deg(x^2+1)$. (In practice, we can find such $q(x)$ and $r(x)$ using long division.) Therefore, $r(x)$ has the form $r(x) = a+bx$ where $a, b \in \mathbb{R}$. Then, $f(x)+\langle x^2+1\rangle = r(x)+\langle x^2+1\rangle$, because $f(x)-r(x) = (x^2+1)\cdot q(x) \in \langle x^2+1\rangle$. Thus, $f(x)+\langle x^2+1\rangle$ can be reduced to $(a+bx)+\langle x^2+1\rangle$ where $a, b \in \mathbb{R}$.

Therefore, we have

$$\mathbb{R}[x]/\langle x^2+1\rangle = \{(a+bx)+\langle x^2+1\rangle \mid a, b \in \mathbb{R}\}.$$

Below is a theorem which states that these cosets of the form $(a+bx)+\langle x^2+1\rangle$ are distinct. Its proof is left for you as an exercise at the end of the chapter.

Theorem 34.1. *If $a_0+a_1x \neq b_0+b_1x$ in $\mathbb{R}[x]$, then $(a_0+a_1x)+\langle x^2+1\rangle \neq (b_0+b_1x)+\langle x^2+1\rangle$ in $\mathbb{R}[x]/\langle x^2+1\rangle$.*

Remark. We've written the reduced cosets in the form $(a+bx)+\langle x^2+1\rangle$ rather than the more customary $(ax+b)+\langle x^2+1\rangle$. This is done to emphasize the correspondence $(a+bx)+\langle x^2+1\rangle \leftrightarrow a+bi$ in the isomorphism $\mathbb{R}[x]/\langle x^2+1\rangle \cong \mathbb{C}$ that we'll see later in the chapter.

Method #2: Coset representative. We have $x^2+\langle x^2+1\rangle = -1+\langle x^2+1\rangle$ because $x^2-(-1) = x^2+1 \in \langle x^2+1\rangle$. Thus, *as coset representatives*, we can treat x^2 and -1 to be the same. Whenever we see x^2 as part of a coset representative, we can replace it with -1. This is the same technique we employed in $\mathbb{Z}_7[x]/\langle x^2-1\rangle$, where we treated x^2 and 1 to be the same as coset representatives. (See Section 33.2 for more details.)

Let $f(x) = 5x^4+x^3-3x^2+4x-3 \in \mathbb{R}[x]$ again. To apply the reduction method described above, we isolate x^2 in the coset representative. For instance, the term $5x^4$ is written as $5\cdot x^2\cdot x^2$, so that each occurrence of x^2 in the coset representative can be replaced with -1:

$$\begin{aligned} f(x)+\langle x^2+1\rangle &= (5x^4+x^3-3x^2+4x-3)+\langle x^2+1\rangle \\ &= (5\cdot \boldsymbol{x^2}\cdot \boldsymbol{x^2}+\boldsymbol{x^2}\cdot x-3\cdot \boldsymbol{x^2}+4x-3)+\langle x^2+1\rangle \\ &= (5\cdot(\boldsymbol{-1})\cdot(\boldsymbol{-1})+(\boldsymbol{-1})\cdot x-3\cdot(\boldsymbol{-1})+4x-3)+\langle x^2+1\rangle \\ &= (5+3x)+\langle x^2+1\rangle, \end{aligned}$$

which is the same reduction as before.

34.2 Field of complex numbers

Consider the set of complex numbers $\mathbb{C} = \{a+bi \mid a, b \in \mathbb{R}\}$, where $i = \sqrt{-1}$ so that $i^2 = -1$. In an exercise, you'll verify that $\mathbb{C}$ is a commutative ring. In fact, we'll soon see that $\mathbb{C}$ is a field.

Example 34.2. Here is how we add and multiply in $\mathbb{C}$:

- $(2+7i)+(4+3i) = (2+4)+(7+3)i = 6+10i$.
- $$\begin{aligned}(2+7i)\cdot(4+3i) &= 2\cdot 4 + 2\cdot 3i + 7i\cdot 4 + 7i\cdot 3i \\ &= 8 + 34i + 21\cdot \boldsymbol{i^2} \\ &= 8 + 34i + 21\cdot(\boldsymbol{-1}) \\ &= -13 + 34i.\end{aligned}$$

Note how we can replace i^2 with -1 to simplify the product.

Example 34.3. For comparison, let's add and multiply in $\mathbb{R}[x]/\langle x^2+1\rangle$:

- $$\begin{aligned}((2+7x)+\langle x^2+1\rangle)+((4+3x)+\langle x^2+1\rangle) &= ((2+7x)+(4+3x))+\langle x^2+1\rangle \\ &= ((2+4)+(7+3)x)+\langle x^2+1\rangle \\ &= (6+10x)+\langle x^2+1\rangle.\end{aligned}$$
- $$\begin{aligned}((2+7x)+\langle x^2+1\rangle)\cdot((4+3x)+\langle x^2+1\rangle) &= ((2+7x)\cdot(4+3x))+\langle x^2+1\rangle \\ &= (2\cdot 4+2\cdot 3x+7x\cdot 4+7x\cdot 3x)+\langle x^2+1\rangle \\ &= (8+34x+21\cdot \boldsymbol{x^2})+\langle x^2+1\rangle \\ &= (8+34x+21\cdot(\boldsymbol{-1}))+\langle x^2+1\rangle \\ &= (-13+34x)+\langle x^2+1\rangle.\end{aligned}$$

Note how we can replace x^2 with -1 in the coset representative.

Examples 34.2 and 34.3 demonstrate how addition and multiplication in $\mathbb{C}$ and in $\mathbb{R}[x]/\langle x^2+1\rangle$ behave essentially in the same manner. More precisely, the two rings are isomorphic, with the complex number $a+bi \in \mathbb{C}$ corresponding to the coset $(a+bx)+\langle x^2+1\rangle$ in $\mathbb{R}[x]/\langle x^2+1\rangle$.

Theorem 34.4. *Consider the function $\theta : \mathbb{C} \to \mathbb{R}[x]/\langle x^2+1\rangle$, where $\theta(a+bi) = (a+bx)+\langle x^2+1\rangle$ for all $a+bi \in \mathbb{C}$. Then θ is a ring isomorphism.*

PROOF. Below, we will show that θ is a bijection. The proof that θ preserves addition and multiplication is left for you as an exercise at the end of the chapter.

First, we will show that θ is onto. Let $f(x)+\langle x^2+1\rangle \in \mathbb{R}[x]/\langle x^2+1\rangle$. We've seen that $f(x)+\langle x^2+1\rangle$ can be reduced to $(a+bx)+\langle x^2+1\rangle$ where $a, b \in \mathbb{R}$. Thus for $a+bi \in \mathbb{C}$, we have

$$\theta(a+bi) = (a+bx)+\langle x^2+1\rangle = f(x)+\langle x^2+1\rangle,$$

so that θ is onto.

To show that θ is one-to-one, suppose $\theta(a+bi) = \theta(c+di)$ where $a+bi, c+di \in \mathbb{C}$. Then

$$(a+bx)+\langle x^2+1\rangle = (c+dx)+\langle x^2+1\rangle.$$

By the contrapositive of Theorem 34.1, we conclude that $a+bx = c+dx$ in $\mathbb{R}[x]$. This implies that $a=c$ and $b=d$ in $\mathbb{R}$, so that $a+bi = c+di$ in $\mathbb{C}$. Therefore, θ is one-to-one. ■

The upshot of this isomorphism is that we can apply our knowledge of $\mathbb{C}$ when working with $\mathbb{R}[x]/\langle x^2+1\rangle$. For instance, finding multiplicative inverses is simpler to do in $\mathbb{C}$ than in $\mathbb{R}[x]/\langle x^2+1\rangle$. The next couple of examples will illustrate this calculation.

Example 34.5. Let $4+3i \in \mathbb{C}$. Then its *complex conjugate* is $4-3i$. Note also that

$$\begin{aligned}(4+3i)\cdot(4-3i) &= 4\cdot 4 + 4\cdot(-3i) + 3i\cdot 4 + 3i\cdot(-3i)\\ &= 4^2 - 3^2\cdot \boldsymbol{i^2} \qquad \leftarrow \text{ the terms } -12i \text{ and } 12i \text{ cancel each other}\\ &= 4^2 - 3^2\cdot(-\mathbf{1})\\ &= 4^2+3^2,\end{aligned}$$

so that $(4+3i)\cdot(4-3i) = 4^2+3^2$.

Example 34.6. Again, let $4+3i \in \mathbb{C}$. Since $4+3i$ is a non-zero (complex) number, its reciprocal $\frac{1}{4+3i}$ is the multiplicative inverse of $4+3i$. To apply the isomorphism $\mathbb{C} \cong \mathbb{R}[x]/\langle x^2+1\rangle$, however, we must rewrite $\frac{1}{4+3i}$ in the form $a+bi$, where $a, b \in \mathbb{R}$. This is where the conjugate plays a role:

$$(4+3i)^{-1} = \frac{1}{4+3i}\cdot\frac{4-3i}{4-3i} = \frac{4}{4^2+3^2} + \frac{-3}{4^2+3^2}\,i.$$

Therefore, $(4+3i)^{-1} = \frac{4}{25} - \frac{3}{25}i$. We'll leave it to you as an exercise to verify that $(4+3i)\cdot\left(\frac{4}{25}-\frac{3}{25}i\right) = 1$ and $\left(\frac{4}{25}-\frac{3}{25}i\right)\cdot(4+3i) = 1$.

The calculation in Example 34.6 can be generalized to show that every non-zero element of $\mathbb{C}$ has a multiplicative inverse. Thus we obtain the following theorem whose proof is left to you.

Theorem 34.7. *$\mathbb{C}$ is a field.*

Since $\mathbb{C}$ is a field, $\mathbb{R}[x]/\langle x^2+1\rangle$ is also a field. Thus every non-zero element in $\mathbb{R}[x]/\langle x^2+1\rangle$ is a unit. Below, we find the multiplicative inverse of $(4+3x)+\langle x^2+1\rangle$ in two different ways.

Method #1: System of equations. We must find $(a+bx)+\langle x^2+1\rangle \in \mathbb{R}[x]/\langle x^2+1\rangle$ such that

$$((4+3x)+\langle x^2+1\rangle)\cdot((a+bx)+\langle x^2+1\rangle) = 1+\langle x^2+1\rangle.$$

Expanding the left side of the above equation, we get

$$\begin{aligned}(4+3x)\cdot(a+bx)+\langle x^2+1\rangle &= (4\cdot a + 4\cdot bx + 3x\cdot a + 3x\cdot bx)+\langle x^2+1\rangle\\ &= (4a+(4b+3a)x+3b\cdot \boldsymbol{x^2})+\langle x^2+1\rangle\\ &= (4a+(4b+3a)x+3b\cdot(-\mathbf{1}))+\langle x^2+1\rangle\\ &= ((4a-3b)+(4b+3a)x)+\langle x^2+1\rangle.\end{aligned}$$

Setting this equal to $1+\langle x^2+1\rangle$ implies $4a-3b=1$ and $4b+3a=0$ in $\mathbb{R}$. Solving this system of equations, we obtain $a=\frac{4}{25}$ and $b=-\frac{3}{25}$, so that the desired multiplicative inverse is $\left(\frac{4}{25}-\frac{3}{25}x\right)+\langle x^2+1\rangle$.

Method #2: Isomorphism with $\mathbb{C}$. We use the isomorphism $\mathbb{C} \cong \mathbb{R}[x]/\langle x^2+1\rangle$. The coset $(4+3x)+\langle x^2+1\rangle$ corresponds to $4+3i \in \mathbb{C}$, whose multiplicative inverse is $(4+3i)^{-1} = \frac{4}{25}-\frac{3}{25}i$. (See Example 34.6.) Translating back to $\mathbb{R}[x]/\langle x^2+1\rangle$, we obtain the multiplicative inverse $\left(\frac{4}{25}-\frac{3}{25}x\right)+\langle x^2+1\rangle$, as we found in Method #1.

34.3 $F[x]/\langle g(x)\rangle$ is a field revisited

In Section 33.4, we studied examples of quotient rings of the form $F[x]/\langle g(x)\rangle$, where $g(x)$ is unfactorable in $F[x]$. In all of these, $F[x]/\langle g(x)\rangle$ is a field. We include another example from this chapter, since x^2+1 is unfactorable in $\mathbb{R}[x]$ and $\mathbb{R}[x]/\langle x^2+1\rangle$ is a field. Here is the conjecture we had made:

> *Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

To start thinking about the proof of this conjecture, we recall how we determine that $F[x]/\langle g(x)\rangle$ is a field in each of our examples.

- For $\mathbb{Q}[x]/\langle x^2-2\rangle$, derive a ring isomorphism $\mathbb{Q}[x]/\langle x^2-2\rangle \cong \mathbb{Q}(\sqrt{2})$ using the First Isomorphism Theorem. And since $\mathbb{Q}(\sqrt{2})$ is a field, conclude that $\mathbb{Q}[x]/\langle x^2-2\rangle$ is also a field.
- Since $\mathbb{Z}_3[x]/\langle x^2+1\rangle$ contains only 9 elements, verify that each of its non-zero elements is a unit.
- For $\mathbb{Z}_7[x]/\langle x^2+1\rangle$, consider an arbitrary non-zero element $(ax+b)+\langle x^2+1\rangle$. Then compute its multiplicative inverse and show that it must exist.
- For $\mathbb{R}[x]/\langle x^2+1\rangle$, obtain a ring isomorphism $\mathbb{C} \cong \mathbb{R}[x]/\langle x^2+1\rangle$ via the correspondence $a+bi \leftrightarrow (a+bx)+\langle x^2+1\rangle$. And since $\mathbb{C}$ is a field, conclude that $\mathbb{R}[x]/\langle x^2+1\rangle$ is a field, too.

These four approaches are quite different from each other. Thus, they do not suggest a proof approach that can be applied to all cases. For that, we'll have to wait until the next chapter.

Exercises

1. Consider the quotient ring $\mathbb{Z}_{13}[x]/\langle x^3+7x+5\rangle$ and let $f(x) = 5x^7+10x^6+2x^3+8x+4 \in \mathbb{Z}_{13}[x]$.
   (a) Use long division to reduce the coset $f(x)+\langle x^3+7x+5\rangle$.
   (b) Explain why $x^3+\langle x^3+7x+5\rangle = (6x+8)+\langle x^3+7x+5\rangle$ in $\mathbb{Z}_{13}[x]/\langle x^3+7x+5\rangle$.
   (c) Use your result in part (b) and Method #2 in Section 34.1 to reduce $f(x)+\langle x^3+7x+5\rangle$.
2. How many distinct elements does $\mathbb{Z}_{13}[x]/\langle x^3+7x+5\rangle$ contain? Explain how you know.
3. Explain why $\mathbb{Z}_{13}[x]/\langle x^3+7x+5\rangle$ is *not* a field.
4. Consider again the quotient ring $\mathbb{Z}_{13}[x]/\langle x^3+7x+5\rangle$, and let $f(x) = x^2+1 \in \mathbb{Z}_{13}[x]$. It turns out that $f(x)+\langle x^3+7x+5\rangle$ is a unit in $\mathbb{Z}_{13}[x]/\langle x^3+7x+5\rangle$. Find its multiplicative inverse.
5. In the quotient ring $\mathbb{R}[x]/\langle x^2+1\rangle$, do the following:
   (a) Explain why $(3+5x)+\langle x^2+1\rangle \neq (-2+7x)+\langle x^2+1\rangle$.
   (b) Explain why $(4+10x)+\langle x^2+1\rangle \neq (-3+10x)+\langle x^2+1\rangle$.
   (c) Explain why $(\pi+3x)+\langle x^2+1\rangle \neq (\pi-8x)+\langle x^2+1\rangle$.

6. Prove Theorem 34.1.

7. Let $f(x) = 3x^7 + x^3 + 4 \in \mathbb{R}[x]$. Find a polynomial $g(x) \in \mathbb{R}[x]$ of the smallest degree such that $f(x) + \langle x^2+1\rangle = g(x) + \langle x^2+1\rangle$ in $\mathbb{R}[x]/\langle x^2+1\rangle$.

   Note: Do this *twice*, using the two different methods shown in Section 34.1.

8. Repeat Exercise #7, this time with $f(x) = 2x^9 + 5x^7 + 4x^3 + 3 \in \mathbb{R}[x]$.

9. Verify that the set $\mathbb{C} = \{a+bi \mid a,\ b \in \mathbb{R}\}$, with addition and multiplication defined in Example 34.2, satisfies the ring properties outlined in Definition 26.2.

10. Complete the proof of Theorem 34.4 by showing that θ preserves addition and multiplication.

11. Verify that $(4+3i)\cdot\left(\frac{4}{25} - \frac{3}{25}i\right) = 1$ and $\left(\frac{4}{25} - \frac{3}{25}i\right)\cdot(4+3i) = 1$. (See Example 34.6.)

12. Consider the quotient ring $\mathbb{R}[x]/\langle x^2+1\rangle$.

    (a) Use Method #1 in Section 34.2 to find the multiplicative inverse of $(2+7x) + \langle x^2+1\rangle$.

    (b) Use the isomorphism $\mathbb{C} \cong \mathbb{R}[x]/\langle x^2+1\rangle$ to find the multiplicative inverse of $(2+7x) + \langle x^2+1\rangle$.

    (c) Let $(a+bx) + \langle x^2+1\rangle$ be the multiplicative inverse you found in parts (a) and (b). Compute

    $$((2+7x) + \langle x^2+1\rangle)\cdot((a+bx) + \langle x^2+1\rangle)$$

    and

    $$((a+bx) + \langle x^2+1\rangle)\cdot((2+7x) + \langle x^2+1\rangle)$$

    and verify that these coset products actually equal $1 + \langle x^2+1\rangle$.

13. Repeat Exercise #12, with $(4-x) + \langle x^2+1\rangle$ in place of $(2+7x) + \langle x^2+1\rangle$.

14. Prove Theorem 34.7.

15. Let $f(x) = 4x^5 + 5x^4 + x^3 + 9x^2 - 3x + 4 \in \mathbb{R}[x]$.

    (a) Verify that $f(i) = 0$ where $i = \sqrt{-1} \in \mathbb{C}$.

    (b) Verify that $f(-i) = 0$ as well.

16. (a) Repeat Exercise #15, this time with $f(x) = 7x^{12} + 7x^{10} - 3x^5 - 3x^3 + 2x^2 + 2$.

    (b) **Prove:** Let $f(x) \in \mathbb{R}[x]$. If $i \in \mathbb{C}$ is a root of $f(x)$, then $-i$ is also a root of $f(x)$.

17. Consider the function $\varphi : \mathbb{R}[x] \to \mathbb{C}$ where $\varphi(f(x)) = f(i)$ for all $f(x) \in \mathbb{R}[x]$. Prove the following:

    (a) φ is a ring homomorphism.

    (b) $\ker\varphi = \langle x^2+1\rangle$. (**Hint:** Use Exercise #16(b).)

    (c) $\operatorname{im}\varphi = \mathbb{C}$.

    (d) What conclusion can you make using the First Isomorphism Theorem?

    Note: Compare with Chapter 33, Exercise #18.

18. Consider the quotient ring $\mathbb{Q}[x]/\langle x^2-2\rangle$, and let $f(x) = 5x^3 - 3x + 4 \in \mathbb{Q}[x]$.
    (a) Use long division to reduce the coset $f(x) + \langle x^2 - 2\rangle$.
    (b) Explain why $x^2 + \langle x^2 - 2\rangle = 2 + \langle x^2 - 2\rangle$ in $\mathbb{Q}[x]/\langle x^2 - 2\rangle$.
    (c) Use your result in part (b) and Method #2 in Section 34.1 to reduce $f(x) + \langle x^2 - 2\rangle$.

19. Explain why each $f(x)+\langle x^2-2\rangle \in \mathbb{Q}[x]/\langle x^2-2\rangle$ can be reduced to $(a+bx)+\langle x^2-2\rangle$, where $a, b \in \mathbb{Q}$.

    Note: Thus, we have the correspondence $(a + bx) + \langle x^2 - 2\rangle \leftrightarrow a + b\sqrt{2}$ in the ring isomorphism $\mathbb{Q}[x]/\langle x^2 - 2\rangle \cong \mathbb{Q}(\sqrt{2})$. (See Example 32.13.)

20. Find the multiplicative inverse of each element in $\mathbb{Q}(\sqrt{2})$. Be sure to write your answer in the form $a + b\sqrt{2}$, where $a, b \in \mathbb{Q}$. (**Hint:** See Example 34.6.)
    (a) $5 + 3\sqrt{2}$.
    (b) $-6 + 11\sqrt{2}$.
    (c) $1 + \sqrt{2}$.
    (d) $10 - 7\sqrt{2}$.
    (e) $a + b\sqrt{2}$ ← Assume that it's a non-zero element.

21. Consider the quotient ring $\mathbb{Q}[x]/\langle x^2 - 2\rangle$.
    (a) Use Method #1 in Section 34.2 to find the multiplicative inverse of $(5 + 3x) + \langle x^2 - 2\rangle$.
    (b) Use the isomorphism $\mathbb{Q}[x]/\langle x^2-2\rangle \cong \mathbb{Q}(\sqrt{2})$ to find the multiplicative inverse of $(5 + 3x) + \langle x^2 - 2\rangle$.
    (c) Let $(a + bx) + \langle x^2 - 2\rangle$ be the multiplicative inverse you found in parts (a) and (b). Compute
    $$((5 + 3x) + \langle x^2 - 2\rangle) \cdot ((a + bx) + \langle x^2 - 2\rangle)$$
    and
    $$((a + bx) + \langle x^2 - 2\rangle) \cdot ((5 + 3x) + \langle x^2 - 2\rangle)$$
    and verify that these coset products actually equal $1 + \langle x^2 - 2\rangle$.

22. Repeat Exercise #21, with $(-6 + 11x) + \langle x^2 - 2\rangle$ in place of $(5 + 3x) + \langle x^2 - 2\rangle$.

35

$F[x]/\langle g(x)\rangle$ Is/Isn't a Field, Part I

In Sections 33.4 and 34.3, we analyzed examples that support the following conjecture:

> *Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

The goal of this chapter is to (finally!) give its proof. What makes the proof interesting is not necessarily its content, but the process through which we develop the proof.

In recent chapters, we've found many structural similarities between the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$, where F is a field. We will now *use* these similarities to derive a proof of our conjecture. First, we find an analogous statement in $\mathbb{Z}$. Since unfactorable polynomials in $F[x]$ correspond to prime numbers in $\mathbb{Z}$, our conjecture, when written in the language of $\mathbb{Z}$, becomes the following:

Fix $p \in \mathbb{Z}$. If p is prime, then $\mathbb{Z}/\langle p\rangle$ is a field.

We'll prove this statement about $\mathbb{Z}$ and then translate that proof into the language of $F[x]$. As we'll see, all of the ingredients that go into the proof of the statement about $\mathbb{Z}$ are also valid in $F[x]$. Informally, we might say that the two proofs in $\mathbb{Z}$ and in $F[x]$ are "isomorphic" to each other. Indeed, this chapter gives yet another illustration of the power of *abstraction* in mathematics, i.e., the process of extracting structural similarities that occur in seemingly different scenarios.

35.1 Translate from $F[x]$ to $\mathbb{Z}$

Our main goal in this chapter is to complete the proof of the following theorem:

Theorem 35.1. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $F[x]/\langle g(x)\rangle$ is* not *a field.*

(b) *If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

Recall that part (a) of the theorem has been resolved already. (See Theorem 33.8.) Thus, our focus will be on part (b). As described in the introduction to this chapter,

we will rely on the structural similarities between $\mathbb{Z}$ and $F[x]$. To start, observe that unfactorable polynomials in $F[x]$ correspond to prime numbers in $\mathbb{Z}$. Therefore, part (b) of the theorem can be translated into the language of $\mathbb{Z}$ as follows:

Fix $p \in \mathbb{Z}$. If p is prime, then $\mathbb{Z}/\langle p\rangle$ is a field.

Now, $\mathbb{Z}/\langle p\rangle$ is the same as $\mathbb{Z}/p\mathbb{Z}$, because $\langle p\rangle = p\mathbb{Z}$. We also have a ring isomorphism $\mathbb{Z}/p\mathbb{Z} \cong \mathbb{Z}_p$ (see Chapter 32, Exercise #21), and we know that $\mathbb{Z}_p$ is a field when p is prime. Thus, $\mathbb{Z}/\langle p\rangle$ must be a field as well.

The following example reviews the argument for why $\mathbb{Z}_p$ is a field.

Example 35.2. Suppose $a \in \mathbb{Z}_{29}$ with $a \neq 0$. For instance, let $a = 8$. Since 29 is a prime number, we have $\gcd(8,\ 29) = 1$. Then by the GCD theorem (i.e., Theorem 3.9), there exist $x, y \in \mathbb{Z}$ such that $8x + 29y = 1$. In fact, we have $8 \cdot 11 + 29 \cdot (-3) = 1$. Then $8 \cdot 11 - 1 = 29 \cdot 3$, so that 29 is a divisor of $8 \cdot 11 - 1$. Thus $8 \cdot 11 = 1$ in $\mathbb{Z}_{29}$, so that 8 has a multiplicative inverse, namely 11. Hence, $8 \in \mathbb{Z}_{29}$ is a unit. Arguing similarly, we can show that every non-zero element of $\mathbb{Z}_{29}$ is a unit. Therefore, $\mathbb{Z}_{29}$ is a field.

We now generalize Example 35.2 by replacing 29 with a prime number p and using the quotient ring $\mathbb{Z}/\langle p\rangle$ instead of $\mathbb{Z}_p$.

Theorem 35.3. *If p is prime, then $\mathbb{Z}/\langle p\rangle$ is a field.*

PROOF. Let $a + \langle p\rangle$ be a non-zero element of $\mathbb{Z}/\langle p\rangle$, so that $a \in \mathbb{Z}$ with $a \notin \langle p\rangle$. We must show that $a + \langle p\rangle$ is a unit in $\mathbb{Z}/\langle p\rangle$. Since $a \notin \langle p\rangle$, the integer a is *not* a multiple of p. And since p is prime, we have $\gcd(a,\ p) = 1$. Hence by the GCD theorem, there exist $x, y \in \mathbb{Z}$ such that $ax + py = 1$.

We have $(a + \langle p\rangle) \cdot (x + \langle p\rangle) = ax + \langle p\rangle$. Moreover, $ax + \langle p\rangle = 1 + \langle p\rangle$, since $ax - 1 = p \cdot (-y) \in \langle p\rangle$. Thus, $(a + \langle p\rangle) \cdot (x + \langle p\rangle) = 1 + \langle p\rangle$, so that $a + \langle p\rangle$ is a unit in $\mathbb{Z}/\langle p\rangle$, with multiplicative inverse $x + \langle p\rangle$. Therefore $\mathbb{Z}/\langle p\rangle$ is a field, as desired. ■

Key to the above proof is the GCD theorem for integers (Theorem 3.9), which states:

Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) = 1$, then there exist integers x and y such that $ax + by = 1$.

The GCD theorem is proved in Appendix A. However, we'll give a different proof of the GCD theorem that can more easily translate back into the language of $F[x]$.

The following example illustrates a concept that is needed for our new proof of the GCD theorem.

Example 35.4. Define the set $\langle 8,\ 29\rangle = \{8x + 29y \mid x,\ y \in \mathbb{Z}\}$ containing all linear combinations of 8 and 29. In Example 35.2, we saw that $1 \in \langle 8,\ 29\rangle$, because $1 = 8 \cdot 11 + 29 \cdot (-3)$. We claim that $a \in \langle 8,\ 29\rangle$ for every $a \in \mathbb{Z}$. Multiplying both sides of the equation $1 = 8 \cdot 11 + 29 \cdot (-3)$ by a, we obtain $a = 8 \cdot (11a) + 29 \cdot (-3a)$, so that $a = 8x + 29y$ where $x = 11a$ and $y = -3a$. Thus, we conclude that $\langle 8,\ 29\rangle = \mathbb{Z}$.

Here is a definition that generalizes the set $\langle 8,\ 29\rangle$ in Example 35.4.

Definition 35.5. Let R be a commutative ring and fix elements $a_1, a_2, \ldots, a_n \in R$. The set

$$\langle a_1,\ a_2,\ \ldots,\ a_n\rangle = \{a_1 \cdot r_1 + a_2 \cdot r_2 + \cdots + a_n \cdot r_n \mid r_1, r_2, \ldots, r_n \in R\}$$

is called the *ideal generated by* $a_1, a_2, \ldots, a_n$.

This definition is also a generalization of the *principal ideal* $\langle a\rangle = \{a \cdot r \mid r \in R\}$, which is generated by a single fixed element $a \in R$. As its name suggests, the set $\langle a_1, a_2, \ldots, a_n\rangle$ is indeed an ideal of the ring R. We'll state that as a theorem and leave the proof to you as an exercise.

Theorem 35.6. *Let R be a commutative ring and fix elements $a_1, a_2, \ldots, a_n \in R$. The set $\langle a_1, a_2, \ldots, a_n\rangle$ is an ideal of R.*

Example 35.7. Consider the set A of all polynomials in $\mathbb{Z}[x]$ with an even constant term. In Chapter 31, Exercises #11, #12, and #13, we found that A is an ideal of $\mathbb{Z}[x]$, but not a principal ideal; i.e., there does not exist an element $\alpha(x) \in \mathbb{Z}[x]$ such that $A = \langle\alpha(x)\rangle$. In fact, we need *two* elements of $\mathbb{Z}[x]$ to generate the ideal A. We showed that $A = \langle x, 2\rangle$, where $\langle x, 2\rangle = \{x \cdot f(x) + 2 \cdot g(x) \mid f(x), g(x) \in \mathbb{Z}[x]\}$.

For the proof of our next theorem, recall that $\mathbb{Z}$ is a *principal ideal domain* (PID), i.e., an integral domain whose ideals are all principal. (See Theorem 31.33.)

Theorem 35.8. *Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) = 1$, then $\langle a, b\rangle = \mathbb{Z}$.*

PROOF. Assume $\gcd(a, b) = 1$. Since $\langle a, b\rangle$ is an ideal of $\mathbb{Z}$ and every ideal of $\mathbb{Z}$ is principal, we have $\langle a, b\rangle = \langle d\rangle$ for some $d \in \mathbb{Z}$. As explained in the remark below, we may assume that $d > 0$. Note that $a \in \langle a, b\rangle$, because $a = a \cdot 1 + b \cdot 0$. And since $\langle a, b\rangle = \langle d\rangle$, we find that $a \in \langle d\rangle$. Thus, $a = d \cdot r$ for some $r \in \mathbb{Z}$, so that d is a divisor of a. Similarly, d is also a divisor of b; thus d is a common divisor of a and b. But since $\gcd(a, b) = 1$, we have $d = 1$. As $\langle a, b\rangle = \langle d\rangle$ and $\langle d\rangle = \langle 1\rangle = \mathbb{Z}$, we obtain $\langle a, b\rangle = \mathbb{Z}$. ■

Remark. We have $d \neq 0$ in the above proof, because $\langle a, b\rangle = \langle 0\rangle$ would imply that a and b are both 0, which cannot occur since $\gcd(a, b) = 1$. Since $d \neq 0$, either d or $-d$ is positive. And you'll show in an exercise that $\langle d\rangle = \langle -d\rangle$. Thus, we may assume that d is positive in the equality $\langle a, b\rangle = \langle d\rangle$.

Using Theorem 35.8, we can reprove the GCD theorem for integers. We'll leave the details to you as an exercise.

35.2 Translate (back) from $\mathbb{Z}$ to $F[x]$

We'd like to state and prove the polynomial version of the GCD theorem. To do that, we must define what it means for two polynomials to be *relatively prime*. For insight, let's return to the ring of integers $\mathbb{Z}$. We say that 8 and 29 are relatively prime (or $\gcd(8, 29) = 1$), as they have no common divisor except 1 and -1. Moreover, observe that 1 and -1 are the only units in $\mathbb{Z}$. Now, the only units in $F[x]$ are the units of F (Theorem 28.18), which are all the non-zero elements of F since F is a field.

The above comparison with $\mathbb{Z}$ motivates the following definition.

Definition 35.9. Let F be a field, and let $f(x), g(x) \in F[x]$. Then $f(x)$ and $g(x)$ are *relatively prime* if they have no common factor except for the non-zero elements of F, i.e., the non-zero constant polynomials.

Example 35.10. Let $f(x) = 3x^2 + 4x + 1$ and $g(x) = x^3 + 2$ in $\mathbb{Z}_7[x]$. Since $g(x)$ has no root in $\mathbb{Z}_7$ (we'll leave the verification to you) and $\deg g(x) = 3$, Theorem 30.19 implies that $g(x)$ is unfactorable in $\mathbb{Z}_7[x]$. Also, $f(x)$ is *not* a multiple of $g(x)$, because $\deg f(x) < \deg g(x)$. Therefore, $f(x)$ and $g(x)$ are relatively prime.

The argument in Example 35.10 above seems sensible, especially when compared to its counterpart in $\mathbb{Z}$. For instance, since 29 is prime and 8 is *not* a multiple of 29, we conclude that 8 and 29 are relatively prime. But let's make a more rigorous argument, using the definition of *relatively prime* in $F[x]$. We'll start the proof of the next theorem and leave it for you to complete as an exercise.

Theorem 35.11. *Let F be a field, and let $f(x), g(x) \in F[x]$. If $g(x)$ is unfactorable and $f(x)$ is not a multiple of $g(x)$, then $f(x)$ and $g(x)$ are relatively prime.*

Proof know-how. To show that the polynomials $f(x)$ and $g(x)$ are relatively prime, let $d(x)$ be a common factor. Then show that $d(x)$ is a non-zero element of F, i.e., a non-zero constant polynomial.

PROOF. Assume $g(x)$ is unfactorable and $f(x)$ is *not* a multiple of $g(x)$. Let $d(x) \in F[x]$ be a common factor of $f(x)$ and $g(x)$. Thus, $f(x) = d(x) \cdot p(x)$ and $g(x) = d(x) \cdot q(x)$ for some $p(x), q(x) \in F[x]$. Since $g(x)$ is unfactorable, Theorem 30.8 implies that either $d(x)$ or $q(x)$ is a non-zero constant.

Suppose for contradiction that $q(x)$ is a non-zero constant; i.e., $q(x) = \alpha$ for some $\alpha \in F$ with $\alpha \neq 0$.

(We'll leave it up to you to obtain a contradiction here.)

Thus $q(x)$ cannot be a non-zero constant, and hence $d(x)$ must be a non-zero constant. Therefore $f(x)$ and $g(x)$ are relatively prime, as desired. ■

We now proceed as we did with $\mathbb{Z}$. Notice the stark similarity between the proof below and that of Theorem 35.8. And as with $\mathbb{Z}$, recall that $F[x]$ is a principal ideal domain. (See Theorem 31.34.)

Theorem 35.12. *Let F be a field, and let $f(x), g(x) \in F[x]$. If $f(x)$ and $g(x)$ are relatively prime, then $\langle f(x), g(x)\rangle = F[x]$.*

PROOF. Assume $f(x)$ and $g(x)$ are relatively prime. Since $\langle f(x), g(x)\rangle$ is an ideal of $F[x]$ and every ideal of $F[x]$ is principal, we have $\langle f(x), g(x)\rangle = \langle d(x)\rangle$ for some $d(x) \in F[x]$. Note that $f(x) \in \langle f(x), g(x)\rangle$, because $f(x) = f(x) \cdot 1 + g(x) \cdot 0$. And since $\langle f(x), g(x)\rangle = \langle d(x)\rangle$, we find that $f(x) \in \langle d(x)\rangle$. Thus, $f(x) = d(x) \cdot p(x)$ for some $p(x) \in F[x]$, so that $d(x)$ is a factor of $f(x)$. Similarly, $d(x)$ is also a factor of $g(x)$; thus $d(x)$ is a common factor of $f(x)$ and $g(x)$. But since $f(x)$ and $g(x)$ are relatively prime, $d(x)$ is a non-zero element of F. In other words, $d(x)$ is a unit of $F[x]$ and thus $\langle d(x)\rangle = F[x]$. Since $\langle f(x), g(x)\rangle = \langle d(x)\rangle$ and $\langle d(x)\rangle = F[x]$, we obtain $\langle f(x), g(x)\rangle = F[x]$. ■

The above proof contains the claim, "$d(x)$ is a unit of $F[x]$ and thus $\langle d(x)\rangle = F[x]$." This statement relies on the following theorem, whose proof is left for you as an exercise. Note that in the above proof, the ideal $A = \langle d(x)\rangle$ contains a unit of $R = F[x]$, namely the element $d(x)$; thus, Theorem 35.13 implies that $A = R$ or $\langle d(x)\rangle = F[x]$.

Theorem 35.13. *Let A be an ideal of a ring R. If A contains a unit of R, then $A = R$.*

To conclude this section, here is the GCD theorem for polynomials. Its proof is left to you as an exercise.

Theorem 35.14 (GCD theorem for polynomials)**.** *Let F be a field, and let $f(x)$, $g(x) \in F[x]$. If $f(x)$ and $g(x)$ are relatively prime, then there exist $p(x)$, $q(x) \in F[x]$ such that $f(x) \cdot p(x) + g(x) \cdot q(x) = 1$.*

Example 35.15. Let $f(x) = 3x^2 + 4x + 1$ and $g(x) = x^3 + 2$ in $\mathbb{Z}_7[x]$. We showed in Example 35.10 that $f(x)$ and $g(x)$ are relatively prime. Thus, there exist $p(x)$, $q(x) \in \mathbb{Z}_7[x]$ such that $f(x) \cdot p(x) + g(x) \cdot q(x) = 1$. In fact, you should verify that

$$f(x) \cdot (5x^2 + 5x + 1) + g(x) \cdot (-x) = 1.$$

35.3 Proof of Theorem 35.1(b)

Now we are ready to prove part (b) of Theorem 35.1. To show that $F[x]/\langle g(x)\rangle$ is a field, we must show that every non-zero element is a unit (i.e., has a multiplicative inverse). Notice how the proof below is essentially the same as the proof of Theorem 35.3.

Theorem 35.1(b). *Let F be a field, and fix $g(x) \in F[x]$. If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

PROOF. Let $\alpha(x)+\langle g(x)\rangle$ be a non-zero element of $F[x]/\langle g(x)\rangle$, so that $\alpha(x) \in F[x]$ with $\alpha(x) \notin \langle g(x)\rangle$. We must show that $\alpha(x) + \langle g(x)\rangle$ is a unit in $F[x]/\langle g(x)\rangle$. Since $\alpha(x) \notin \langle g(x)\rangle$, the polynomial $\alpha(x)$ is *not* a multiple of $g(x)$. And since $g(x)$ is unfactorable, Theorem 35.11 implies that $\alpha(x)$ and $g(x)$ are relatively prime. Hence by the GCD theorem for polynomials, there exist $p(x)$, $q(x) \in F[x]$ such that $\alpha(x) \cdot p(x) + g(x) \cdot q(x) = 1$.

We have $(\alpha(x) + \langle g(x)\rangle) \cdot (p(x) + \langle g(x)\rangle) = \alpha(x) \cdot p(x) + \langle g(x)\rangle$. Moreover, $\alpha(x) \cdot p(x) + \langle g(x)\rangle = 1 + \langle g(x)\rangle$, since $\alpha(x) \cdot p(x) - 1 = g(x) \cdot (-q(x)) \in \langle g(x)\rangle$. Thus, $(\alpha(x)+\langle g(x)\rangle)\cdot(p(x)+\langle g(x)\rangle) = 1+\langle g(x)\rangle$, so that $\alpha(x)+\langle g(x)\rangle$ is a unit in $F[x]/\langle g(x)\rangle$, with multiplicative inverse $p(x) + \langle g(x)\rangle$. Therefore $F[x]/\langle g(x)\rangle$ is a field, as desired. ■

Big picture stuff

The main goal of the chapter was to prove the following: *If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.* To develop its proof, we took advantage of the structural similarities between $\mathbb{Z}$ and $F[x]$. We began by stating and proving the analogous theorem for the integers: *If p is prime, then $\mathbb{Z}/\langle p\rangle$ is a field.* Then we translated the proof in $\mathbb{Z}$ back to our proof in $F[x]$.

In fact, these two statements can be further generalized as follows:

Let R be a principal ideal domain and fix $p \in R$. If p is irreducible, then $R/\langle p\rangle$ is a field.

Here, an *irreducible* element (in an integral domain) is a generalization of an unfactorable polynomial in $F[x]$ and a prime number in $\mathbb{Z}$.

Exercises

1. **(Review)** Fix $g(x) = x^2 - 4 \in \mathbb{Z}_{11}[x]$, and consider $\mathbb{Z}_{11}[x]/\langle g(x)\rangle$.
   (a) Let $\alpha(x) \in \mathbb{Z}_{11}[x]$. Explain *how* you'd find $\beta(x) \in \mathbb{Z}_{11}[x]$ of the smallest degree such that $\alpha(x) + \langle g(x)\rangle = \beta(x) + \langle g(x)\rangle$.
   (b) In part (a), describe how $\alpha(x)$ and $\beta(x)$ are related.
   (c) Describe all distinct elements of $\mathbb{Z}_{11}[x]/\langle g(x)\rangle$. How many are there?
   (d) Is $\mathbb{Z}_{11}[x]/\langle g(x)\rangle$ a field? Why or why not?

2. **(Review)** Find the multiplicative inverse of each element.
   (a) $(3x + 1) + \langle x^2 - 1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2 - 1\rangle$.
   (b) $(3x + 1) + \langle x^2 + 1\rangle$ in $\mathbb{Z}_7[x]/\langle x^2 + 1\rangle$.
   (c) $(7x + 4) + \langle x^2 + 1\rangle$ in $\mathbb{R}[x]/\langle x^2 + 1\rangle$.
   (d) $(7x + 4) + \langle x^2 - 2\rangle$ in $\mathbb{Q}[x]/\langle x^2 - 2\rangle$.

3. Find an integer solution (x, y) to the equation $8x + 41y = 1$. Use that solution to find the multiplicative inverse of 8 in $\mathbb{Z}_{41}$.

4. Find an integer solution (x, y) to the equation $7x + 19y = 1$. Use that solution to find the multiplicative inverse of 7 in $\mathbb{Z}_{19}$.

5. Let $g(x) = x^2 + 2 \in \mathbb{Z}_{13}[x]$.
   (a) Explain why $g(x)$ is unfactorable in $\mathbb{Z}_{13}[x]$.
   (b) How many elements does $\mathbb{Z}_{13}[x]/\langle g(x)\rangle$ contain? Explain how you know.
   (c) Is $\mathbb{Z}_{13}[x]/\langle g(x)\rangle$ a field? Why or why not?

6. Find a prime p and a polynomial $g(x) \in \mathbb{Z}_p[x]$ such that the quotient ring $\mathbb{Z}_p[x]/\langle g(x)\rangle$ is a field containing 121 elements. Explain your reasoning.

7. Find a prime p and a polynomial $g(x) \in \mathbb{Z}_p[x]$ such that the quotient ring $\mathbb{Z}_p[x]/\langle g(x)\rangle$ is a field containing 343 elements. Explain your reasoning. (This exercise is referenced in Example 37.4.)

8. Consider the quotient ring $\mathbb{R}[x]/\langle x^2 - 1\rangle$. Let $a, b \in \mathbb{R}$ such that $a^2 \neq b^2$. Show that the element $(ax + b) + \langle x^2 - 1\rangle$ is a unit in $\mathbb{R}[x]/\langle x^2 - 1\rangle$.

9. Prove Theorem 35.6.

10. Let R be a commutative ring and fix elements $a_1, a_2, \ldots, a_n \in R$. If I is an ideal of R that contains $a_1, a_2, \ldots, a_n$, then show that $\langle a_1, a_2, \ldots, a_n\rangle \subseteq I$.

    Note: In other words, $\langle a_1, a_2, \ldots, a_n\rangle$ is the *smallest* ideal containing the elements $a_1, a_2, \ldots, a_n$.

11. Let R be a commutative ring, and let $d \in R$. Prove that $\langle d\rangle = \langle -d\rangle$. (This exercise is needed in the proof of Theorem 35.8. It is also referenced in Example 36.12.)

12. Fix $4, 6 \in \mathbb{Z}$ and consider the ideal $\langle 4, 6\rangle = \{4x + 6y \mid x, y \in \mathbb{Z}\}$. Since every ideal of $\mathbb{Z}$ is principal, we have $\langle 4, 6\rangle = \langle d\rangle$ for some $d \in \mathbb{Z}$. Assuming that $d > 0$, find the value of d.

13. Repeat Exercise #12, but with each of the following. What conjecture do you have?

    (a) $\langle 8, 12\rangle$. (b) $\langle 10, 12\rangle$. (c) $\langle 15, 20\rangle$. (d) $\langle 21, 35\rangle$. (e) $\langle 0, 23\rangle$.

14. **Prove:** Let $a, b \in \mathbb{Z}$, not both zero. Then $\langle a, b\rangle = \langle d\rangle$, where $d = \gcd(a, b)$.

    Hint: There are a couple of possible approaches:

    - **Approach #1:** Let $d = \gcd(a, b)$. Then show that $\langle a, b\rangle \subseteq \langle d\rangle$ and $\langle d\rangle \subseteq \langle a, b\rangle$.
    - **Approach #2:** Let $g = \gcd(a, b)$. Proceeding as in the proof of Theorem 35.8, we have $\langle a, b\rangle = \langle d\rangle$ where d is a positive integer. Then show that $d \geq g$ and $g \geq d$, and thus $d = g$.

15. Fix $4, 6 \in \mathbb{Z}$ and consider the ideal $\langle 4\rangle \cap \langle 6\rangle = \{n \in \mathbb{Z} \mid n \in \langle 4\rangle \text{ and } n \in \langle 6\rangle\}$. (See Chapter 31, Exercise #21(b).) Since every ideal of $\mathbb{Z}$ is principal, we have $\langle 4\rangle \cap \langle 6\rangle = \langle d\rangle$ for some $d \in \mathbb{Z}$. Assuming that $d > 0$, find the value of d.

16. Repeat Exercise #15, but with each of the following. What conjecture do you have?

    (a) $\langle 8\rangle \cap \langle 12\rangle$. (b) $\langle 10\rangle \cap \langle 12\rangle$. (c) $\langle 15\rangle \cap \langle 20\rangle$. (d) $\langle 21\rangle \cap \langle 35\rangle$. (e) $\langle 0\rangle \cap \langle 23\rangle$.

17. **Prove:** Let $a, b \in \mathbb{Z}$. Then $\langle a\rangle \cap \langle b\rangle = \langle d\rangle$, where d is the least common multiple of a and b.

18. Let $f(x) = x^2 - 6x + 8$ and $g(x) = x^2 + 3x - 10$ in $\mathbb{R}[x]$. Given that every ideal of $\mathbb{R}[x]$ is principal, find $d(x) \in \mathbb{R}[x]$ such that the following hold:

    (a) $\langle f(x), g(x)\rangle = \langle d(x)\rangle$.
    (b) $\langle f(x)\rangle \cap \langle g(x)\rangle = \langle d(x)\rangle$.

19. Repeat Exercise #18 with the following pairs of polynomials in $\mathbb{R}[x]$.

    (a) $f(x) = x^2 + 1$ and $g(x) = x^3$.
    (b) $f(x) = x^2 + 1$ and $g(x) = x^4 - 1$.

20. Write a proof of the GCD theorem for integers that uses Theorem 35.8.

21. Complete the proof of Theorem 35.11 by obtaining a contradiction.

22. Prove Theorem 35.13.

    Hint: We have $A \subseteq R$ by definition. Thus, it suffices to show the other inclusion $R \subseteq A$.

23. Prove Theorem 35.14.

36

Maximal Ideals

G. H. Hardy, a British mathematician, wrote the following in his book *A Mathematician's Apology*:

> The mathematician's patterns, like the painter's or the poet's, must be beautiful; the ideas, like the colours or the words, must fit together in a harmonious way.

Indeed, aesthetics plays an important role in mathematics, and the purpose of the remaining two chapters of this textbook is to bring that beauty to the forefront. To set the context, recall Theorem 35.1 whose proof was completed in the last chapter.

Theorem 35.1. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $F[x]/\langle g(x)\rangle$ is* not *a field.*

(b) *If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

For part (a), we used the factorization of $g(x)$ to construct zero divisors in $F[x]/\langle g(x)\rangle$. To prove part (b), we used the GCD theorem for polynomials to find a multiplicative inverse for any non-zero element in $F[x]/\langle g(x)\rangle$. While these proofs were mathematically valid, they were two *separate* proofs that did not illuminate any connection between the two parts of this theorem.

The next two chapters will introduce an alternate proof to Theorem 35.1 that brings its two parts together, as if they're two sides of the same coin. Hardy might say that this new proof approach will highlight how the two parts of the theorem "fit together in a harmonious way."

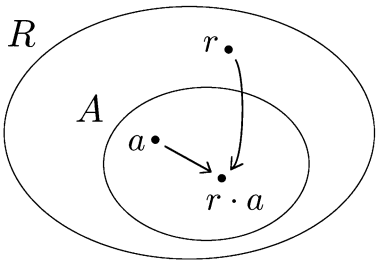
The key ingredient in this new proof will be the notion of *maximal* ideals, which is the focus of this chapter. We will continue to work with the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$, where F is a field. In particular, we will rely on the fact that each is a *principal ideal domain* (PID), i.e., an integral domain whose ideals are all principal. (See Theorems 31.33 and 31.34.)

36.1 Examples and definition

Recall that a subset A of a ring R is called an *ideal* of R if the following hold:

- A is an additive subgroup of R.
- A satisfies the product absorption property: If $r \in R$ and $a \in A$, then $r \cdot a \in A$.

Below is a visual depiction of product absorption. It's as if the element $a \in A$ *absorbs* the element $r \in R$ into the set A when they are multiplied together.

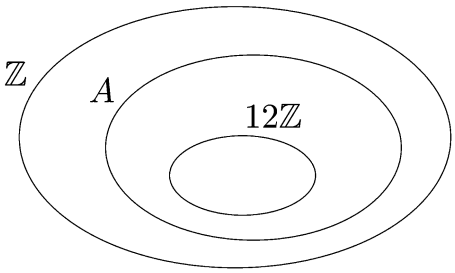


Example 36.1. The subset $12\mathbb{Z}$ is an ideal of $\mathbb{Z}$. We'll leave it up to you to verify that $12\mathbb{Z}$ is an additive subgroup of $\mathbb{Z}$. For product absorption, say $7 \in \mathbb{Z}$ and $24 = 12 \cdot 2 \in 12\mathbb{Z}$. Then $7 \cdot 24 = 168 = 12 \cdot 14 \in 12\mathbb{Z}$. More generally, let $r \in \mathbb{Z}$ and $a \in 12\mathbb{Z}$, so that $a = 12 \cdot k$ for some $k \in \mathbb{Z}$. Then

$$r \cdot a = r \cdot (12 \cdot k) = 12 \cdot (r \cdot k) \in 12\mathbb{Z}.$$

Thus, $12\mathbb{Z}$ satisfies the product absorption property. We can also write $12\mathbb{Z}$ as $\langle 12 \rangle = \{12 \cdot r \mid r \in \mathbb{Z}\}$, i.e., the principal ideal generated by 12. In fact, recall from Theorem 31.33 that *every* ideal of $\mathbb{Z}$ is principal.

Example 36.2. Suppose A is an ideal of $\mathbb{Z}$ such that $12\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$, which is a shorthand for two set inclusions: $12\mathbb{Z} \subseteq A$ ($12\mathbb{Z}$ is contained in ideal A) and $A \subseteq \mathbb{Z}$ (ideal A is contained in $\mathbb{Z}$). Intuitively, the ideal A is "sandwiched" between $12\mathbb{Z}$ and $\mathbb{Z}$, as shown by the diagram below:



One possibility is $A = 4\mathbb{Z}$. (You'll find the other possibilities in an exercise at the end of the chapter.) Certainly $4\mathbb{Z} \subseteq \mathbb{Z}$. To see that $12\mathbb{Z} \subseteq 4\mathbb{Z}$, let $\alpha \in 12\mathbb{Z}$ so that $\alpha = 12k$ for some $k \in \mathbb{Z}$. Then we have $\alpha = 12k = 4 \cdot (3k) \in 4\mathbb{Z}$, so that $\alpha \in 4\mathbb{Z}$. Thus $12\mathbb{Z} \subseteq 4\mathbb{Z}$, and therefore $12\mathbb{Z} \subseteq 4\mathbb{Z} \subseteq \mathbb{Z}$.

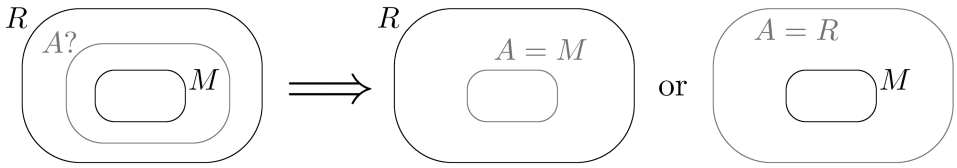
Example 36.3. Let A be an ideal of $\mathbb{Z}$ such that $5\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$. We claim that A must be equal to either $5\mathbb{Z}$ or $\mathbb{Z}$. Since every ideal in $\mathbb{Z}$ is principal, we have $A = \langle d \rangle$ for some positive integer d. (See the remark after Theorem 35.8 for why we can assume that

$d > 0$.) Thus $5\mathbb{Z} \subseteq \langle d \rangle$, and since $5 \in 5\mathbb{Z}$, we obtain $5 \in \langle d \rangle$. Therefore, $5 = d \cdot r$ for some $r \in \mathbb{Z}$ and hence d is a divisor of 5. The only positive divisors of 5 are 1 and 5. If $d = 1$, we get $A = \langle 1 \rangle = \mathbb{Z}$. If $d = 5$, we have $A = \langle 5 \rangle = 5\mathbb{Z}$. Hence A equals either $5\mathbb{Z}$ or $\mathbb{Z}$.

The ideal $5\mathbb{Z}$ in Example 36.3 is an instance of a *maximal* ideal. The name "maximal" is fitting, since there is no ideal (other than $\mathbb{Z}$) that is "bigger" than $5\mathbb{Z}$. When we tried to "sandwich" an ideal A between $5\mathbb{Z}$ and $\mathbb{Z}$, i.e., $5\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$, then we found that A must equal either $5\mathbb{Z}$ or $\mathbb{Z}$. Here is a generalization.

Definition 36.4 (Maximal ideal)**.** Let M be an ideal of a commutative ring R, with $M \neq R$. Then M is said to be a *maximal ideal* of R (or simply *maximal* in R) if for any ideal A such that $M \subseteq A \subseteq R$, we must have either $A = M$ or $A = R$.

The figure below is a visual depiction of this definition.



Example 36.5 (Non-example)**.** The ideal $12\mathbb{Z}$ in Example 36.2 is *not* maximal in $\mathbb{Z}$, since $12\mathbb{Z} \subseteq 4\mathbb{Z} \subseteq \mathbb{Z}$ where $4\mathbb{Z}$ does not equal either $12\mathbb{Z}$ or $\mathbb{Z}$. We write $12\mathbb{Z} \subsetneq 4\mathbb{Z}$ to mean that $12\mathbb{Z}$ is a subset of $4\mathbb{Z}$, but $12\mathbb{Z} \neq 4\mathbb{Z}$. Likewise, $4\mathbb{Z} \subsetneq \mathbb{Z}$ means $4\mathbb{Z}$ is a subset of $\mathbb{Z}$ but does not equal $\mathbb{Z}$. Combining these, we write $12\mathbb{Z} \subsetneq 4\mathbb{Z} \subsetneq \mathbb{Z}$ and we say that $4\mathbb{Z}$ is *strictly* between $12\mathbb{Z}$ and $\mathbb{Z}$.

Proof know-how. To show that M is *not* maximal in R, find an ideal A such that $M \subseteq A \subseteq R$ where A does not equal either M or R. Symbolically, we write $M \subsetneq A \subsetneq R$ and we say that A is *strictly* between M and R. Here, $M \subsetneq A$ means M is a subset of A but $M \neq A$. Similarly, $A \subsetneq R$ means $A \subseteq R$ but $A \neq R$.

Example 36.6. Consider the ring $\mathbb{Z}_{12}$. In Example 14.10, we found its additive subgroups:

$$\{0\},\ \{0, 6\},\ \{0, 4, 8\},\ \{0, 3, 6, 9\},\ \{0, 2, 4, 6, 8, 10\},\ \mathbb{Z}_{12}.$$

We'll leave it to you as an exercise to verify that each of these satisfies the product absorption property and hence is an ideal. Let's see which of these are maximal in $\mathbb{Z}_{12}$.

- $\{0\}$ is *not* maximal, because $\{0\} \subsetneq \{0, 6\} \subsetneq \mathbb{Z}_{12}$.
- $\{0, 6\}$ is *not* maximal, because $\{0, 6\} \subsetneq \{0, 3, 6, 9\} \subsetneq \mathbb{Z}_{12}$.
- $\{0, 4, 8\}$ is *not* maximal, because $\{0, 4, 8\} \subsetneq \{0, 2, 4, 6, 8, 10\} \subsetneq \mathbb{Z}_{12}$.
- $\{0, 3, 6, 9\}$ is maximal, because there is no ideal strictly between $\{0, 3, 6, 9\}$ and $\mathbb{Z}_{12}$. Thus if there is an ideal A such that $\{0, 3, 6, 9\} \subseteq A \subseteq \mathbb{Z}_{12}$, then we must have $A = \{0, 3, 6, 9\}$ or $A = \mathbb{Z}_{12}$.

- $\{0, 2, 4, 6, 8, 10\}$ is maximal, because there is no ideal strictly between $\{0, 2, 4, 6, 8, 10\}$ and $\mathbb{Z}_{12}$.
- $\mathbb{Z}_{12}$ is *not* maximal. By definition, a maximal ideal M (of a ring R) must be different from R.

Thus $\mathbb{Z}_{12}$ has two maximal ideals, namely $\{0, 3, 6, 9\}$ and $\{0, 2, 4, 6, 8, 10\}$.

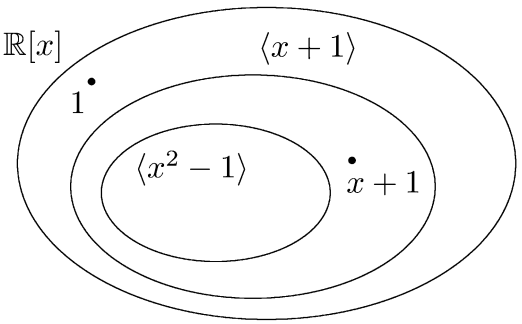
Example 36.7. Consider the ring $\mathbb{Z}_7$. Its ideals are $\{0\}$ and $\mathbb{Z}_7$ only. (Do you see why?) Then $\{0\}$ is a maximal ideal of $\mathbb{Z}_7$. If there is an ideal A such that $\{0\} \subseteq A \subseteq \mathbb{Z}_7$, then A must equal either $\{0\}$ or $\mathbb{Z}_7$, since those are the only ideals in $\mathbb{Z}_7$.

36.2 Maximality of $\langle g(x)\rangle$

Example 36.8 (Non-example)**.** Consider $x^2 - 1 \in \mathbb{R}[x]$. We'll show that the ideal $\langle x^2 - 1\rangle$ is *not* maximal in $\mathbb{R}[x]$ by finding an ideal A that is strictly between $\langle x^2 - 1\rangle$ and $\mathbb{R}[x]$; i.e., $\langle x^2-1\rangle \subsetneq A \subsetneq \mathbb{R}[x]$. Noting that x^2-1 factors as $x^2-1 = (x+1)\cdot(x-1)$, we'll let $A = \langle x + 1\rangle$.

First, we'll show that $\langle x^2 - 1\rangle \subsetneq A$. If $\alpha(x) \in \langle x^2 - 1\rangle$, then $\alpha(x) = (x^2 - 1) \cdot q(x)$ for some $q(x) \in \mathbb{R}[x]$. Then $\alpha(x) = ((x+1)\cdot(x-1))\cdot q(x) = (x+1)\cdot((x-1)\cdot q(x))$, so that $\alpha(x)$ is a multiple of $x+1$; i.e., $\alpha(x) \in A$. Hence, $\langle x^2 - 1\rangle \subseteq A$. Moreover, $x+1$ is a multiple of $x+1$, but not a multiple of x^2-1. Therefore, $x+1 \in A$, but $x+1 \notin \langle x^2-1\rangle$. This show that $\langle x^2 - 1\rangle \neq A$, so that $\langle x^2 - 1\rangle \subsetneq A$.

Certainly, $A \subseteq \mathbb{R}[x]$. But $A \neq \mathbb{R}[x]$, as $1 \in \mathbb{R}[x]$ but $1 \notin A$. (Note that the only constant polynomial in $A = \langle x+1\rangle$ is 0.) Thus, we obtain $A \subsetneq \mathbb{R}[x]$. Therefore, we find that $\langle x^2 - 1\rangle \subsetneq A \subsetneq \mathbb{R}[x]$, so that $A = \langle x + 1\rangle$ is an ideal that is strictly between $\langle x^2 - 1\rangle$ and $\mathbb{R}[x]$. We conclude that $\langle x^2-1\rangle$ is *not* maximal. The picture below illustrates these set inclusions:



Proof know-how. To show that $M \subsetneq A$, we must show that $M \subseteq A$ and $M \neq A$. Showing $M \subseteq A$ can be done in the familiar manner: Consider an element $m \in M$ and show that $m \in A$. To show that $M \neq A$ (i.e., M and A are different sets), one approach is to find an element $a \in A$ such that $a \notin M$. In Example 36.8, for instance, we showed that $\langle x^2 - 1\rangle \neq A$ by showing that $x + 1 \in A$, but $x + 1 \notin \langle x^2 - 1\rangle$.

Here is the generalization of Example 36.8, whose proof is left for you as an exercise.

Theorem 36.9. *Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is factorable, then $\langle g(x)\rangle$ is* not *maximal in $F[x]$.*

Based on this theorem, we might conjecture the following:

Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is unfactorable, then $\langle g(x)\rangle$ is maximal in $F[x]$.

Let's first verify this with an example.

Example 36.10. Consider the polynomial $x^2 + 1$, which is unfactorable in $\mathbb{R}[x]$. To show that the ideal $\langle x^2+1\rangle$ is maximal in $\mathbb{R}[x]$, consider an ideal A such that $\langle x^2+1\rangle \subseteq A \subseteq \mathbb{R}[x]$. We will show that either $A = \langle x^2+1\rangle$ or $A = \mathbb{R}[x]$. Since every ideal of $\mathbb{R}[x]$ is principal (Theorem 31.34), we have $A = \langle p(x)\rangle$ for some $p(x) \in \mathbb{R}[x]$. Therefore, $\langle x^2+1\rangle \subseteq \langle p(x)\rangle \subseteq \mathbb{R}[x]$.

Since $x^2 + 1 \in \langle x^2 + 1\rangle$ and $\langle x^2 + 1\rangle \subseteq \langle p(x)\rangle$, we have $x^2 + 1 \in \langle p(x)\rangle$. Thus $x^2 + 1 = p(x) \cdot q(x)$ for some $q(x) \in \mathbb{R}[x]$. As $x^2 + 1$ is unfactorable in $\mathbb{R}[x]$, either $\deg p(x) = 0$ or $\deg q(x) = 0$ by Theorem 30.8. We'll consider each case separately.

If $\deg p(x) = 0$, then $p(x)$ is a non-zero constant polynomial in $\mathbb{R}[x]$. Thus, $p(x)$ is a non-zero real number, which is a unit in $\mathbb{R}[x]$. Then by Theorem 35.13, we find that $\langle p(x)\rangle = \mathbb{R}[x]$. Hence $A = \mathbb{R}[x]$ in this case. Next, suppose $\deg q(x) = 0$. Then $q(x)$ is a unit of $\mathbb{R}[x]$, and so $\langle p(x)\rangle = \langle p(x) \cdot q(x)\rangle$ (see the remark after the proof), which implies that $A = \langle x^2 + 1\rangle$ for this case. In either case, we showed that $A = \langle x^2 + 1\rangle$ or $A = \mathbb{R}[x]$. Therefore, $\langle x^2 + 1\rangle$ is maximal in $\mathbb{R}[x]$.

Proof know-how. To show that an ideal M is maximal in a commutative ring R, first consider an ideal A such that $M \subseteq A \subseteq R$. Then show that either $A = M$ or $A = R$.

Remark. The above proof contains the line, "Then $q(x)$ is a unit of $\mathbb{R}[x]$, and so $\langle p(x)\rangle = \langle p(x) \cdot q(x)\rangle$." This step requires the use of the following theorem, which is left for you to prove as an exercise.

Theorem 36.11. *Let R be an integral domain and suppose $a, b \in R$ where $a, b \neq 0$. Then $\langle a\rangle = \langle a \cdot b\rangle$ if and only if b is a unit of R.*

Example 36.12. In Chapter 35, Exercise #11, you showed that $\langle d\rangle = \langle -d\rangle$ in a commutative ring. Writing $-d = d \cdot (-1)$, we have $\langle d\rangle = \langle d \cdot (-1)\rangle$, where -1 is a unit in any ring as $(-1) \cdot (-1) = 1$. Hence, this example illustrates Theorem 36.11 with $a = d$ and $b = -1$.

Let's state our conjecture as a theorem.

Theorem 36.13. *Let F be a field and fix $g(x) \in F[x]$. If $g(x)$ is unfactorable, then $\langle g(x)\rangle$ is maximal in $F[x]$.*

PROOF. Assume $g(x)$ is unfactorable in $F[x]$. As every ideal of $F[x]$ is principal, consider an ideal $\langle p(x)\rangle$ for some $p(x) \in F[x]$ such that $\langle g(x)\rangle \subseteq \langle p(x)\rangle \subseteq F[x]$. We must show that $\langle p(x)\rangle = \langle g(x)\rangle$ or $\langle p(x)\rangle = F[x]$. Since $g(x) \in \langle g(x)\rangle$ and $\langle g(x)\rangle \subseteq \langle p(x)\rangle$, we have $g(x) \in \langle p(x)\rangle$ and hence $g(x) = p(x) \cdot q(x)$ for some $q(x) \in F[x]$. And because $g(x)$ is unfactorable, we must have either $\deg p(x) = 0$ or $\deg q(x) = 0$.

Suppose $\deg p(x) = 0$; i.e., $p(x)$ is a non-zero constant polynomial. Thus $p(x)$ is a non-zero element of the field F, which makes it a unit in $F[x]$. Thus $\langle p(x)\rangle = F[x]$ by Theorem 35.13. Next, suppose $\deg q(x) = 0$, so that $q(x)$ is a unit of $F[x]$. Then $\langle p(x)\rangle = \langle p(x) \cdot q(x)\rangle$ by Theorem 36.11, so that $\langle p(x)\rangle = \langle g(x)\rangle$. In either case, we showed that $\langle p(x)\rangle = \langle g(x)\rangle$ or $\langle p(x)\rangle = F[x]$. Thus $\langle g(x)\rangle$ is maximal in $F[x]$. ■

The following theorem serves as a summary of the two theorems that we proved above.

Theorem 36.14. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $\langle g(x) \rangle$ is* not *maximal in $F[x]$.*

(b) *If $g(x)$ is unfactorable, then $\langle g(x) \rangle$ is maximal in $F[x]$.*

Big picture stuff

By using the contrapositive of its first statement, Theorem 36.14 can be restated as:

$$\langle g(x) \rangle \textit{ is maximal in } F[x] \textit{ if and only if } g(x) \textit{ is unfactorable.}$$

Exercise #5 at the end of this chapter is about the ideal $\langle n \rangle$ (or $n\mathbb{Z}$) in $\mathbb{Z}$. It states the following:

$$\langle n \rangle \textit{ is maximal in } \mathbb{Z} \textit{ if and only if } n \textit{ is prime.}$$

Notice how these two statements are essentially the same. This is yet another instance of the structural similarity between the ring of integers $\mathbb{Z}$ and the polynomial ring $F[x]$ where F is a field.

Exercises

1. Determine whether each ideal of $\mathbb{Z}$ is maximal. What conjecture do you have?

   (a) $6\mathbb{Z}$. (b) $28\mathbb{Z}$. (c) $17\mathbb{Z}$. (d) $41\mathbb{Z}$. (e) $375\mathbb{Z}$.

2. (a) Find all ideals A of the ring $\mathbb{Z}$ such that $12\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$. (See Example 36.2.)
   (b) Find all ideals A of the ring $\mathbb{Z}$ such that $10\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$.
   (c) Find all ideals A of the ring $\mathbb{Z}$ such that $13\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$.
   (d) Find all ideals A of the ring $\mathbb{Z}$ such that $p\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$, where p is prime.
   (e) What conjecture do you have?

3. **Prove:** Let $a, b \in \mathbb{Z}$. Then $b \mid a$ if and only if $\langle a \rangle \subseteq \langle b \rangle$.
   Note: Compare this with Chapter 31, Exercise #20.

4. Use the statement in Exercise #3 to describe all ideals A of the ring $\mathbb{Z}$ such that $n\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$.

5. **Prove:** Let n be a positive integer. The ideal $n\mathbb{Z}$ is maximal in $\mathbb{Z}$ if and only if n is prime.

6. Verify that each additive subgroup of $\mathbb{Z}_{12}$ satisfies the product absorption property and hence is an ideal. (See Example 36.6.)

7. **Prove:** Every additive subgroup of $\mathbb{Z}_n$ satisfies the product absorption property and thus is an ideal.

8. Find all maximal ideals in the following:

   (a) $\mathbb{Z}_{24}$. (b) $\mathbb{Z}_{10}$. (c) $\mathbb{Z}_{32}$. (d) $\mathbb{Z}_{101}$.

9. Describe all maximal ideals in $\mathbb{Z}_n$.

10. Suppose a ring R has 40 elements and an ideal M has 8 elements. Explain why M is a maximal ideal.

11. In Example 36.6, we saw that $\mathbb{Z}_{12}$ has exactly two maximal ideals.

    (a) Verify that $\mathbb{Z}_{20}$ has exactly two maximal ideals.
    (b) Verify that $\mathbb{Z}_{28}$ has exactly two maximal ideals.
    (c) Verify that $\mathbb{Z}_{18}$ has exactly two maximal ideals.
    (d) Find a few more values of n for which $\mathbb{Z}_n$ has exactly two maximal ideals.
    (e) What conjectures do you have?

12. Find a ring that has exactly three maximal ideals.

13. (a) Verify that 9 is *not* a unit in $\mathbb{Z}_{24}$. Then find a maximal ideal of $\mathbb{Z}_{24}$ containing 9.
    (b) Verify that 10 is *not* a unit in $\mathbb{Z}_{35}$. Then find a maximal ideal of $\mathbb{Z}_{35}$ containing 10.
    (c) Find a (non-zero) non-unit in $\mathbb{Z}_{30}$ and a maximal ideal of $\mathbb{Z}_{30}$ containing that non-unit element.
    (d) Find a (non-zero) non-unit in $\mathbb{Z}_{54}$ and a maximal ideal of $\mathbb{Z}_{54}$ containing that non-unit element.
    (e) What conjecture do you have?

14. Prove Theorem 36.9.

15. In $\mathbb{R}[x]$, explain why the principal ideals $\langle 3 \cdot (x^2+1)\rangle$ and $\langle x^2+1\rangle$ are equal.

    Note: This is a set equality proof. Show that $\langle 3 \cdot (x^2+1)\rangle \subseteq \langle x^2+1\rangle$ and $\langle x^2+1\rangle \subseteq \langle 3 \cdot (x^2+1)\rangle$.

16. Prove Theorem 36.11.

17. (a) Let F be a field. Prove that $\{0\}$ is the only maximal ideal of F.
    (b) Is the statement in part (a) still true if we replace a field F with an integral domain R? If it's true, prove it. If it's false, provide a counterexample.

18. Let R be a commutative ring, and let M be an ideal with $M \neq R$. Suppose every element of R which is *not* in M is a unit. Prove that M is the only maximal ideal of R.

    Note: This is a generalization of Exercise #17(a). You must show that (1) M is maximal and (2) there is no other maximal ideal of R.

19. The principal ideal $\langle 2\rangle$ refers to different sets, depending on the ring in which it resides.

    (a) Let $\langle 2\rangle$ be an ideal of the ring $\mathbb{Z}$. Describe the elements in $\langle 2\rangle$.
    (b) Let $\langle 2\rangle$ be an ideal of the ring $\mathbb{Z}[x]$. Describe the elements in $\langle 2\rangle$.

(c) Is $\langle 2 \rangle$ a maximal ideal of $\mathbb{Z}$? Why or why not?

(d) Is $\langle 2 \rangle$ a maximal ideal of $\mathbb{Z}[x]$? Why or why not?

20. Explain why the ideal $\langle x \rangle = \{x \cdot f(x) \mid f(x) \in \mathbb{Z}[x]\}$ is *not* maximal in $\mathbb{Z}[x]$. (This exercise is referenced in Chapter 37, Exercise #15.)

21. Consider the ideal $\langle x, 2 \rangle = \{x \cdot f(x) + 2 \cdot g(x) \mid f(x), g(x) \in \mathbb{Z}[x]\}$ in $\mathbb{Z}[x]$. Recall from Chapter 31, Exercise #13 that $\langle x, 2 \rangle$ is equal to the set of all polynomials in $\mathbb{Z}[x]$ with an even constant term. Prove that $\langle x, 2 \rangle$ is a maximal ideal of $\mathbb{Z}[x]$. (This exercise is referenced in Chapter 37, Exercise #13.)

37

$F[x]/\langle g(x)\rangle$ Is/Isn't a Field, Part II

We will finish the work started in Chapter 36, namely to give an alternate proof to the following:

Theorem 35.1. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $F[x]/\langle g(x)\rangle$ is* not *a field.*

(b) *If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

Towards that goal, the focus of this chapter is to prove the following theorem:

Theorem 37.2. *Let M be an ideal of a commutative ring R.*

(a) *If M is* not *maximal in R, then R/M is* not *a field.*

(b) *If M is maximal in R, then R/M is a field.*

The proof of Theorem 37.2 is fairly complicated, which is perhaps fitting for the last chapter of this book. In the proof of part (a), there are two possible paths that we could take, but only one of them leads to the desired conclusion. For part (b), the proof requires the use of an ideal that might seem as if it was pulled out of thin air. As we've done throughout this book, we will give a detailed breakdown of the proof of each part and describe how one might come up with such proof ideas on one's own.

37.1 Maximal ideals and quotient rings

Example 37.1. Recall that $12\mathbb{Z}$ is *not* a maximal ideal of $\mathbb{Z}$, because $12\mathbb{Z} \subseteq 4\mathbb{Z} \subseteq \mathbb{Z}$ where $4\mathbb{Z}$ does not equal either $12\mathbb{Z}$ or $\mathbb{Z}$. In other words, $4\mathbb{Z}$ is an ideal that is *strictly* between $12\mathbb{Z}$ and $\mathbb{Z}$. On the other hand, $5\mathbb{Z}$ is a maximal ideal of $\mathbb{Z}$. If there is an ideal A such that $5\mathbb{Z} \subseteq A \subseteq \mathbb{Z}$, then we must have either $A = 5\mathbb{Z}$ or $A = \mathbb{Z}$. Thus, there is no ideal (other than $\mathbb{Z}$) that is "bigger" than $5\mathbb{Z}$.

We also note that the quotient ring $\mathbb{Z}/12\mathbb{Z}$ is *not* a field, because it has zero divisors:

$$(3 + 12\mathbb{Z}) \cdot (4 + 12\mathbb{Z}) = 3 \cdot 4 + 12\mathbb{Z} = 0 + 12\mathbb{Z}.$$

But $\mathbb{Z}/5\mathbb{Z}$ is a field. In particular, it is ring isomorphic to $\mathbb{Z}_5$.

The example above motivates the following theorem.

Theorem 37.2. *Let M be an ideal of a commutative ring R.*

(a) *If M is* not *maximal in R, then R/M is* not *a field.*

(b) *If M is maximal in R, then R/M is a field.*

We will start by proving part (a) of the theorem. (Actually, we will prove its contrapositive.)

Theorem 37.2(a). *If R/M is a field, then M is maximal in R.*

PROOF. Assume R/M is a field. Let A be an ideal of R such that $M \subseteq A \subseteq R$. We must show that $A = M$ or $A = R$. We will suppose $A \neq M$ and then show that $A = R$.

Since $M \subseteq A$ but $M \neq A$, there exists an element $a \in A$ such that $a \notin M$. Thus, $a + M \neq 0 + M$ in the quotient ring R/M. And as R/M is a field, $a + M$ is a unit and thus $(a + M) \cdot (b + M) = 1 + M$ for some $b + M \in R/M$. Then, $a \cdot b + M = 1 + M$, so that $1 - a \cdot b \in M$.

Let $m = 1 - a \cdot b$ where $m \in M$. Note that $m \in A$, because $M \subseteq A$. Then $1 = a \cdot b + m$ is in A, because a, $m \in A$ and A is an ideal. (More specifically, $a \cdot b \in A$ by product absorption; and $a \cdot b + m \in A$, because A is an additive subgroup of R.) But $1 \in A$ implies $A = R$ by Theorem 35.13. Hence, M is maximal in R. ■

We now give an in-depth breakdown of this proof. Early in the proof, we were faced with an "or" statement; i.e., we had to prove that (1) $A = M$ or (2) $A = R$. Recalling the Proof know-how for Theorem 20.15, there are two possible paths that we could have pursued:

- **Path #1:** $A = M$ is either true or false. If it's true, then we're done with the proof. Thus, we assume that (1) $A = M$ is false and prove that (2) $A = R$ is true.

- **Path #2:** $A = R$ is either true or false. If it's true, then we're done with the proof. Thus, we assume that (2) $A = R$ is false and prove that (1) $A = M$ is true.

We took the first path; i.e., we assumed that $A \neq M$ and showed that $A = R$. This was a more natural path to take, because we knew that R/M is a field. Assuming $A \neq M$ allowed us to find an element $a \in A$ such that $a + M$ is a non-zero element in R/M. And since R/M is a field, there exists $b + M \in R/M$ such that $(a + M) \cdot (b + M) = 1 + M$. Looking ahead, our goal was to show that $A = R$. In recent proofs (e.g., Theorem 36.13), we accomplished this by showing that $1 \in A$. Thus, we were left with the task of justifying the implication: $(a + M) \cdot (b + M) = 1 + M \implies 1 \in A$. While we still had some work remaining, it's mostly technical and symbolic in nature. The more challenging aspect of coming up with an overall framework of the proof had been completed.

Before moving on, we recall the following theorem, which was proved in Chapter 31, Exercise #21(a).

Theorem 37.3. *Let I and J be ideals of a ring R. Then $I + J = \{i + j \mid i \in I,\ j \in J\}$ is an ideal of R.*

Now we prove the second statement of Theorem 37.2.

Theorem 37.2(b). *If M is maximal in R, then R/M is a field.*

PROOF. Assume M is maximal in R. Let $a+M$ be a non-zero element of R/M. We must show that $a+M$ is a unit. Since $a+M \neq 0+M$, we know that $a \notin M$. Then, consider the sum of ideals

$$M + \langle a \rangle = \{m + \alpha \mid m \in M,\ \alpha \in \langle a \rangle\},$$

which, by Theorem 37.3 above, is also an ideal of R.

For each $m \in M$, we have $m = m + a \cdot 0 \in M + \langle a \rangle$. Thus, $M \subseteq M + \langle a \rangle$. Next, $a = 0 + a \cdot 1 \in M + \langle a \rangle$, but $a \notin M$, which implies that $M + \langle a \rangle \neq M$. Since $M \subseteq M + \langle a \rangle \subseteq R$ and $M + \langle a \rangle \neq M$, the maximality of M implies that $M + \langle a \rangle = R$.

Now, $1 \in R$ implies $1 \in M + \langle a \rangle$. Then, $1 = m + a \cdot b$ for some $m \in M$ and $b \in R$. Thus $1 - a \cdot b = m$, and since $m \in M$, we deduce that $1 - a \cdot b \in M$. Then we have $(a+M) \cdot (b+M) = a \cdot b + M = 1 + M$. Hence, $a+M$ is a unit whose multiplicative inverse is $b+M$. Therefore, R/M is a field, as desired. ∎

We'd like to understand the motivation behind constructing and using the ideal $M + \langle a \rangle$. Since our goal is to show that R/M is a field, it's natural to consider a non-zero element $a + M \in R/M$ with an aim of finding its multiplicative inverse $b + M \in R/M$ such that $(a+M) \cdot (b+M) = 1+M$. Working backwards (a proof technique that we've used often in this book), we obtain $a \cdot b + M = 1 + M$ so that $1 - a \cdot b \in M$. Letting $1 - a \cdot b = m$, we obtain $1 = m + a \cdot b$. Thus, our goal was twofold: (1) find an ideal that contains the element $m + a \cdot b$ and (2) show that this ideal contains 1. Our goal (1) was met by $M + \langle a \rangle$. Then the maximality of M allowed us to conclude that $M + \langle a \rangle = R$, so that $1 \in M + \langle a \rangle$, satisfying goal (2).

Remark. As usual, this "working backwards" process of starting with $(a + M) \cdot (b+M) = 1+M$ and obtaining $1 = m + a \cdot b$ is scratch work and must *not* be presented in the proof. In the actual proof, $(a+M) \cdot (b+M) = 1+M$ must be a conclusion, rather than an assumed starting point.

37.2 Putting it all together

We have the following theorems:

Theorem 36.14. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $\langle g(x) \rangle$ is* not *maximal in $F[x]$.*

(b) *If $g(x)$ is unfactorable, then $\langle g(x) \rangle$ is maximal in $F[x]$.*

Theorem 37.2. *Let M be an ideal of a commutative ring R.*

(a) *If M is* not *maximal in R, then R/M is* not *a field.*

(b) *If M is maximal in R, then R/M is a field.*

Combining these, while using $R = F[x]$ and $M = \langle g(x) \rangle$ in Theorem 37.2, we obtain the familiar theorem, which we had already proved. But this time, we proved it using the new tool of maximal ideals.

Theorem 35.1. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $F[x]/\langle g(x) \rangle$ is* not *a field.*

(b) *If $g(x)$ is unfactorable, then $F[x]/\langle g(x) \rangle$ is a field.*

37.3 Oh wait, but there's more!

Example 37.4 (Chapter 35, Exercise #7 revisited)**.** We will construct a field with 7^3 elements. Consider the polynomial $g(x) = x^3 + 2$ in $\mathbb{Z}_7[x]$. Then $\mathbb{Z}_7[x]/\langle g(x)\rangle = \{(ax^2 + bx + c) + \langle g(x)\rangle \mid a, b, c \in \mathbb{Z}_7\}$. With 7 choices for each of a, b, and c, this quotient ring contains $7^3 = 343$ elements. Since $g(x)$ has no root in $\mathbb{Z}_7$ (we'll leave the verification to you) and $\deg g(x) = 3$, Theorem 30.19 implies that $g(x)$ is unfactorable in $\mathbb{Z}_7[x]$. And because $g(x)$ is unfactorable, we conclude that $\mathbb{Z}_7[x]/\langle g(x)\rangle$ is a field.

Here are some beautiful facts about polynomials in $\mathbb{Z}_p[x]$, where p is prime:

- For every $n \geq 1$, there exists an unfactorable polynomial $g(x) \in \mathbb{Z}_p[x]$ with $\deg g(x) = n$.
- So, we can construct a field $\mathbb{Z}_p[x]/\langle g(x)\rangle$ with p^n elements, for all primes p and all integers $n \geq 1$.
- In fact, *every* finite field is obtained this way.

37.4 Prime ideals

In this final section of the textbook, we will examine *prime ideals*, which are a natural counterpart to maximal ideals. We focused on maximal ideals, since they were instrumental in the alternate proof of Theorem 35.1. However, prime ideals are also interesting objects of study and will play a prominent role in your further study of abstract algebra.

Example 37.5. Consider the ideal $12\mathbb{Z}$ of the ring of integers $\mathbb{Z}$. Let $a, b \in \mathbb{Z}$. If their product ab is in $12\mathbb{Z}$, must a or b be contained in $12\mathbb{Z}$? Not necessarily. Suppose $a = 6$ and $b = 4$. Then $ab = 24$, which is in $12\mathbb{Z}$. However, neither $a = 6$ nor $b = 4$ is a multiple of 12, and thus $a, b \notin 12\mathbb{Z}$.

Example 37.6. Consider the ideal $5\mathbb{Z}$ of $\mathbb{Z}$. Let $a, b \in \mathbb{Z}$. We claim the following: If $ab \in 5\mathbb{Z}$, then a or b is in $5\mathbb{Z}$. Assume $ab \in 5\mathbb{Z}$. Then $ab = 5k$ for some $k \in \mathbb{Z}$, so that 5 is a divisor of ab. Since 5 is prime, Theorem 37.7 below implies that $5 \mid a$ or $5 \mid b$. Therefore, $a \in 5\mathbb{Z}$ or $b \in 5\mathbb{Z}$.

Here's the theorem used in Example 37.6. We'll leave its proof to you as an exercise.

Theorem 37.7. *Let $a, b, p \in \mathbb{Z}$ where p is prime. If $p \mid ab$, then $p \mid a$ or $p \mid b$.*

The ideal $5\mathbb{Z}$ in Example 37.6 is an instance of a *prime* ideal. It has the property that if the product ab is in $5\mathbb{Z}$, then a or b (possibly both) must be in $5\mathbb{Z}$. Here is a generalization.

Definition 37.8 (Prime ideal)**.** Let P be an ideal of a commutative ring R, with $P \neq R$. Then P is said to be a *prime ideal* of R (or simply *prime* in R) if for any $a, b \in R$ such that $ab \in P$, we have $a \in P$ or $b \in P$.

Example 37.9 (Non-example)**.** The ideal $12\mathbb{Z}$ in Example 37.5 is *not* a prime ideal of $\mathbb{Z}$. In particular, we found that $6 \cdot 4 \in 12\mathbb{Z}$, even though $6, 4 \notin 12\mathbb{Z}$.

Proof know-how. To show that P is *not* prime in R, find elements $a, b \in R$ such that $ab \in P$ and $a, b \notin P$.

Example 37.10. As discussed in Example 36.6, the ideals of $\mathbb{Z}_{12}$ are

$$\{0\},\ \{0, 6\},\ \{0, 4, 8\},\ \{0, 3, 6, 9\},\ \{0, 2, 4, 6, 8, 10\},\ \mathbb{Z}_{12}.$$

Let's see which of these are prime in $\mathbb{Z}_{12}$.

- $\{0\}$ is *not* prime, because $3 \cdot 4 = 0 \in \{0\}$, even though $3, 4 \notin \{0\}$.
- $\{0, 6\}$ is *not* prime, because $3 \cdot 4 = 0 \in \{0, 6\}$, even though $3, 4 \notin \{0, 6\}$.
- $\{0, 4, 8\}$ is *not* prime, because $2 \cdot 6 = 0 \in \{0, 4, 8\}$, even though $2, 6 \notin \{0, 4, 8\}$.
- $\{0, 3, 6, 9\}$ is prime. In an exercise, you'll show that $ab \in \{0, 3, 6, 9\}$ implies a or b is in $\{0, 3, 6, 9\}$.
- $\{0, 2, 4, 6, 8, 10\}$ is prime. Again, you'll show this in an exercise.
- $\mathbb{Z}_{12}$ is *not* prime. By definition, a prime ideal P (of a ring R) must be different from R.

Thus $\mathbb{Z}_{12}$ has two prime ideals, namely $\{0, 3, 6, 9\}$ and $\{0, 2, 4, 6, 8, 10\}$.

Example 37.11 (Non-example)**.** Consider $x^2 - 1 \in \mathbb{R}[x]$. We'll show that the ideal $\langle x^2 - 1\rangle$ is *not* prime in $\mathbb{R}[x]$. Let $f(x) = x + 1$ and $g(x) = x - 1$, so that $f(x) \cdot g(x) = x^2 - 1 \in \langle x^2 - 1\rangle$. However, neither $x + 1$ nor $x - 1$ is a multiple of $x^2 - 1$. Thus $f(x)$, $g(x) \notin \langle x^2 - 1\rangle$. We conclude that $\langle x^2 - 1\rangle$ is *not* a prime ideal.

Example 37.12. Let $x^2 + 1 \in \mathbb{R}[x]$, which is unfactorable. To show that $\langle x^2 + 1\rangle$ is a prime ideal, consider $f(x), g(x) \in \mathbb{R}[x]$ such that $f(x) \cdot g(x) \in \langle x^2 + 1\rangle$. We'll show that $f(x) \in \langle x^2 + 1\rangle$ or $g(x) \in \langle x^2 + 1\rangle$. We have $f(x) \cdot g(x) = (x^2 + 1) \cdot q(x)$ for some $q(x) \in \mathbb{R}[x]$. Thus $x^2 + 1$ is a factor of $f(x) \cdot g(x)$. But since $x^2 + 1$ is unfactorable in $\mathbb{R}[x]$, Theorem 37.13 below implies that $(x^2 + 1) \mid f(x)$ or $(x^2 + 1) \mid g(x)$. Therefore, $f(x) \in \langle x^2 + 1\rangle$ or $g(x) \in \langle x^2 + 1\rangle$. (Compare with Example 37.6.)

Here's the theorem used in Example 37.12. We'll leave its proof to you as an exercise.

Theorem 37.13. *Let F be a field, and let $f(x), g(x), p(x) \in F[x]$ where $p(x)$ is unfactorable. If $p(x) \mid f(x) \cdot g(x)$, then $p(x) \mid f(x)$ or $p(x) \mid g(x)$.*

Comparing with the examples from Chapter 36, it seems that every maximal ideal is a prime ideal, and vice versa. Are the two concepts equivalent? Are they just two different ways of looking at the same thing? The answer is "No," as we'll see in the exercises.

Exercises

1. For Theorem 37.2(a), we proved the contrapositive of the statement: *If M is* not *maximal in R, then R/M is* not *a field.* Explore and describe what happens if we try to prove the original statement, rather than the contrapositive.

2. In the proof of Theorem 37.2(a), we assumed that $A \neq M$ and showed that $A = R$. Explore and describe what happens if we assume instead that $A \neq R$ and try to show that $A = M$.
3. Construct a field with 125 elements.
4. Construct a field with 1,331 elements.
5. Construct a field with 243 elements.
6. Determine if each of the following is a prime ideal of $\mathbb{Z}$:

   (a) $10\mathbb{Z}$. (b) $34\mathbb{Z}$. (c) $13\mathbb{Z}$. (d) $23\mathbb{Z}$. (e) $\{0\}$.

7. For each ideal of $\mathbb{Z}$ in Exercise #6, determine if it's a maximal ideal.
8. For each ring R, determine if $P = \{0\}$ is a prime ideal of R. (Thus, $\alpha \in P$ means $\alpha = 0$.)

   (a) $R = \mathbb{Z}_{20}$. (b) $R = \mathbb{Z}_7$. (c) $R = \mathbb{Z}_{33}$. (d) $R = \mathbb{Z}_{17}$. (e) $R = \mathbb{Z}$.

9. What's a name for a commutative ring R for which $\{0\}$ is a prime ideal? Explain your reasoning.
10. Prove Theorem 37.7.
11. Prove Theorem 37.13.
12. Complete Example 37.10 by showing the following:

    (a) $ab \in \{0, 3, 6, 9\}$ implies a or b is in $\{0, 3, 6, 9\}$.
    (b) $ab \in \{0, 2, 4, 6, 8, 10\}$ implies a or b is in $\{0, 2, 4, 6, 8, 10\}$.

13. Let A be the set of all polynomials in $\mathbb{Z}[x]$ with an even constant term. In Chapter 31, Exercise #11(a), we proved that A is an ideal of $\mathbb{Z}[x]$. Explain why A is a prime ideal of $\mathbb{Z}[x]$.

    Note: In Chapter 36, Exercise #21, we showed that A is maximal in $\mathbb{Z}[x]$.

14. In $\mathbb{Z}[x]$, let $\langle x\rangle = \{x \cdot q(x) \mid q(x) \in \mathbb{Z}[x]\}$ be the principal ideal generated by x. Prove that $\langle x\rangle = \{f(x) \in \mathbb{Z}[x] \mid f(0) = 0\}$.

    Note: Be careful. Since $\mathbb{Z}$ is not a field, you may *not* use the factor theorem here.

15. Explain why $\langle x\rangle$ is a prime ideal of $\mathbb{Z}[x]$.

    Note: In Chapter 36, Exercise #20, we showed that $\langle x\rangle$ is *not* maximal in $\mathbb{Z}[x]$.

16. **Prove:** Let R be a commutative ring. If A is maximal in R, then A is prime in R.
17. Give a counterexample to show that the converse of the statement in Exercise #16 is false.
18. **Prove:** Let R be a *finite* commutative ring. If A is prime in R, then A is maximal in R.
19. **Prove:** Let $n \in \mathbb{Z}$ with $n \geq 2$. Then $n\mathbb{Z}$ is a prime ideal of $\mathbb{Z}$ if and only if n is prime.
20. **Prove:** Let R be a commutative ring, and let P be an ideal of R with $P \neq R$. Then P is a prime ideal of R if and only if R/P is an integral domain.

Appendix A

Proof of the GCD Theorem

We will justify the GCD theorem, which was introduced in Chapter 3. (Note that "GCD" is an acronym for "greatest common divisor.")

Theorem 3.9 (GCD theorem)**.** *Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) = 1$, then there exist integers x and y such that $ax + by = 1$.*

In fact, we'll consider a more general statement where the GCD isn't necessarily 1.

Theorem A.1 (Generalized GCD theorem)**.** *Let $a, b \in \mathbb{Z}$. If $\gcd(a, b) = d$, then there exist integers x and y such that $ax + by = d$.*

To work towards a justification, we start with an interesting property of the GCD.

Example A.2. Consider the following pairs of GCDs:

- $\gcd(12, 30) = 6$ and $\gcd(12, 18) = 6$.
- $\gcd(156, 228) = 12$ and $\gcd(156, 72) = 12$.
- $\gcd(35, 21) = 7$ and $\gcd(14, 21) = 7$.
- $\gcd(182, 52) = 26$ and $\gcd(130, 52) = 26$.

These examples seem to suggest that $\gcd(a, b) = \gcd(a, b - a)$ and $\gcd(a, b) = \gcd(a - b, b)$.

We will prove that $\gcd(a, b) = \gcd(a, b - a)$. Reversing the roles of a and b yields $\gcd(a, b) = \gcd(a - b, b)$.

Theorem A.3. *Let $a, b \in \mathbb{Z}$. Then $\gcd(a, b) = \gcd(a, b - a)$.*

PROOF. Define the sets $S = \{d \in \mathbb{Z} \mid d \mid a \text{ and } d \mid b\}$ and $T = \{d \in \mathbb{Z} \mid d \mid a \text{ and } d \mid (b - a)\}$. We will show that the two sets are equal; i.e., $S = T$.

We first show $S \subseteq T$. Assume $d \in S$, so that $d \mid a$ and $d \mid b$. Thus, $a = dk$ and $b = dj$ for some integers k and j. Thus, $b - a = dj - dk = d(j - k)$ where $j - k \in \mathbb{Z}$. Hence, $d \mid (b - a)$, which shows that $d \in T$.

To show that $T \subseteq S$, assume $d \in T$ so that $d \mid a$ and $d \mid (b - a)$. Thus, $a = dk$ and $b - a = dj$ for some k, $j \in \mathbb{Z}$. Hence, $b = (b - a) + a = dj + dk = d(j + k)$ where $j + k \in \mathbb{Z}$. Hence, $d \mid b$, which show that $d \in S$.

Thus, we conclude that $S = T$. Note that S is the set of common divisors of a and b. Likewise, T is the set of common divisors of a and $b - a$. Since the two sets are equal, the greatest element from each set must be equal to each other. In other words, $\gcd(a, b) = \gcd(a, b - a)$. ■

The formulas $\gcd(a, b) = \gcd(a, b-a)$ and $\gcd(a, b) = \gcd(a-b, b)$ help simplify GCD computations, especially when working with larger numbers, by allowing us to reduce the pair of numbers for which to find the GCD. This is illustrated in Examples A.4 and A.5 below.

Example A.4. To compute $\gcd(861, 252)$, we use $\gcd(a, b) = \gcd(a - b, b)$ to replace 861 (i.e., the larger number) with the difference $861-252 = 609$. Thus, $\gcd(861, 252) = \gcd(609, 252)$. We use the formula a couple more times to obtain $\gcd(609, 252) = \gcd(357, 252) = \gcd(105, 252)$. Then we use $\gcd(a, b) = \gcd(a, b - a)$ to replace 252 (i.e., the larger number) with the difference $252-105$. In fact, we can apply $\gcd(a, b) = \gcd(a, b - a)$ twice to replace 252 with $252 - 105 \cdot 2 = 42$. Thus, $\gcd(105, 252) = \gcd(105, 42)$. Now we continue by replacing 105 with a smaller number.

Here is a complete calculation for $\gcd(861, 252)$:

$$\begin{aligned}
\gcd(861, 252) &= \gcd(105, 252) && 861 - 252 \cdot 3 = \mathbf{105},\\
&= \gcd(105, 42) && 252 - 105 \cdot 2 = \mathbf{42},\\
&= \gcd(21, 42) && 105 - 42 \cdot 2 = \mathbf{21},\\
&= \gcd(21, 0) && 42 - 21 \cdot 2 = \mathbf{0},\\
&= 21.
\end{aligned}$$

The numbers in **bold** are called *remainders*. For instance, **105** is the remainder when dividing 861 by 252. Notice that the last non-zero remainder, namely 21, is the desired GCD.

Example A.5. The calculation below shows that $\gcd(861, 253) = 1$, which is the last non-zero remainder:

$$\begin{aligned}
\gcd(861, 253) &= \gcd(102, 253) && 861 - 253 \cdot 3 = 102,\\
&= \gcd(102, 49) && 253 - 102 \cdot 2 = 49,\\
&= \gcd(4, 49) && 102 - 49 \cdot 2 = 4,\\
&= \gcd(4, 1) && 49 - 4 \cdot 12 = 1,\\
&= \gcd(0, 1) && 4 - 1 \cdot 4 = 0,\\
&= 1.
\end{aligned}$$

The approach for finding the GCD shown in Examples A.4 and A.5 is called the *Euclidean algorithm*. We will use the remainder calculation in this algorithm to understand why the generalized GCD theorem (i.e., Theorem A.1) is true. We'll leave a formal proof of the theorem up to you!

Example A.6. In Example A.4, we found that $\gcd(861, 252) = 21$. Thus, the generalized GCD theorem ensures the existence of integers x and y where $861x + 252y = 21$. We will find such integers using the remainder calculation from Example A.4:

$$\begin{aligned} 861 - 252 \cdot 3 &= 105, \\ 252 - 105 \cdot 2 &= 42, \\ 105 - 42 \cdot 2 &= 21, \\ 42 - 21 \cdot 2 &= 0. \end{aligned}$$

The second to last equation tells us that $21 = 105 - \mathbf{42} \cdot 2$. We substitute $\mathbf{42} = 252 - 105 \cdot 2$ from the previous equation to obtain

$$21 = 105 - (252 - 105 \cdot 2) \cdot 2 = 105 \cdot 5 - 252 \cdot 2.$$

Then we substitute $\mathbf{105} = 861 - 252 \cdot 3$ from the top equation into $21 = \mathbf{105} \cdot 5 - 252 \cdot 2$ to obtain

$$21 = (861 - 252 \cdot 3) \cdot 5 - 252 \cdot 2 = 861 \cdot 5 - 252 \cdot 17.$$

Thus, $(x, y) = (5, -17)$ is an integer solution to $861x + 252y = 21$.

This method is called *back-tracking*, working from the bottom to the top of the remainder calculation to write 21 as an integer linear combination of 861 and 252. Here is one more example.

Example A.7. We've seen that $\gcd(861, 253) = 1$. Thus, the generalized GCD theorem guarantees that there exist $x, y \in \mathbb{Z}$ with $861x + 253y = 1$. Here are the results from Example A.5:

$$\begin{aligned} 861 - 253 \cdot 3 &= 102, \\ 253 - 102 \cdot 2 &= 49, \\ 102 - 49 \cdot 2 &= 4, \\ 49 - 4 \cdot 12 &= 1, \\ 4 - 1 \cdot 4 &= 0. \end{aligned}$$

Starting with $1 = 49 - 4 \cdot 12$ and working up the remainder calculation, we obtain

$$\begin{aligned} 1 &= 49 - (102 - 49 \cdot 2) \cdot 12 \\ &= 49 \cdot 25 - 102 \cdot 12 \\ &= (253 - 102 \cdot 2) \cdot 25 - 102 \cdot 12 \\ &= 253 \cdot 25 - 102 \cdot 62 \\ &= 253 \cdot 25 - (861 - 253 \cdot 3) \cdot 62 \\ &= 253 \cdot 211 - 861 \cdot 62. \end{aligned}$$

Therefore, $(x, y) = (-62, 211)$ is an integer solution to $861x + 253y = 1$.

Appendix B

Composition Table for D_4

We saw in Chapter 5 that $D_4 = \{\varepsilon, r_{90}, r_{180}, r_{270}, h, v, d, d'\}$ is the set of symmetries of a square. Here is its composition table. For $\sigma, \tau \in D_4$, the "product" $\sigma \circ \tau$ is the entry in row σ and column τ. For example, the product $d \circ r_{90} = v$ is shown in bold letters.

$\circ$	ε	$\boldsymbol{r_{90}}$	r_{180}	r_{270}	h	v	d	d'
ε	ε	r_{90}	r_{180}	r_{270}	h	v	d	d'
r_{90}	r_{90}	r_{180}	r_{270}	ε	d'	d	h	v
r_{180}	r_{180}	r_{270}	ε	r_{90}	v	h	d'	d
r_{270}	r_{270}	ε	r_{90}	r_{180}	d	d'	v	h
h	h	d	v	d'	ε	r_{180}	r_{90}	r_{270}
v	v	d'	h	d	r_{180}	ε	r_{270}	r_{90}
$\boldsymbol{d}$	d	$\boldsymbol{v}$	d'	h	r_{270}	r_{90}	ε	r_{180}
d'	d'	h	d	v	r_{90}	r_{270}	r_{180}	ε

Appendix C

Symbols and Notations

The page number after each item indicates where the notation is introduced or defined.

$p \implies q$	if p, then q (i.e., implication), p. 35
$p \iff q$	p if and only if q, p. 200
$\mathbb{Z}$	integers, p. 14
$m\mathbb{Z}$	$\{mk \mid k \in \mathbb{Z}\}$, p. 17
$\mathbb{N}$	natural numbers, p. 15
$\mathbb{Q}$	rational numbers, p. 19
$\mathbb{R}$	real numbers, p. 75
$\mathbb{R}^*$	$\{a \in \mathbb{R} \mid a$ has a multiplicative inverse$\}$, p. 76
$\mathbb{C}$	complex numbers, p. 263
$\mathbb{Z}[i]$	$\{a + bi \mid a,\ b \in \mathbb{Z}\}$ where $i^2 = -1$, p. 320
$\mathbb{Z}_3[i]$	$\{a + bi \mid a,\ b \in \mathbb{Z}_3\}$ where $i^2 = -1$, p. 263
$\mathbb{Q}(\sqrt{2})$	$\{a + b\sqrt{2} \mid a,\ b \in \mathbb{Q}\}$, p. 274
$\mathbb{Z}_m$	$\{0, 1, 2, \ldots, m-1\}$ with addition/multiplication modulo m, p. 31
U_m	$\{a \in \mathbb{Z}_m \mid a$ has a multiplicative inverse in $\mathbb{Z}_m\}$, p. 36
$a \in S$	a is an element of set S, p. 14
$a \notin S$	a is *not* an element of set S, p. 14
$S \subseteq T$	S is a subset of T, p. 13
$S \subsetneq T$	S is a subset of T but $S \neq T$ (i.e., *strict* containment), p. 361
$\emptyset$	empty set (also denoted $\{\ \}$), p. 14
$S \cap T$	intersection of sets S and T, p. 20
$d \mid n$	d is a divisor of n, p. 21
$d \nmid n$	d is *not* a divisor of n, p. 21
$\gcd(a, b)$	greatest common divisor of a and b, p. 22
D_3	set of symmetries of an equilateral triangle, p. 48
D_4	set of symmetries of a square, p. 43
D_n	set of symmetries of a regular n-sided polygon, p. 48
S_n	set of permutations of $\{1, 2, 3, \ldots, n\}$, p. 52
ε	identity element of a group, p. 74

g^{-1}	multiplicative inverse of g, p. 76
$-g$	additive inverse of g, p. 75
$\mathrm{ord}(g)$	order of a group element g (also denoted $\lvert g\rvert$), p. 115
$Z(G)$	$\{z \in G \mid zg = gz \text{ for all } g \in G\}$, center of G, p. 109
$C(h)$	$\{g \in G \mid gh = hg\}$, centralizer of h in G, p. 110
$\langle g\rangle$	$\{g^k \mid k \in \mathbb{Z}\}$, cyclic subgroup generated by g (multiplicative), p. 130
$\langle g\rangle$	$\{k \cdot g \mid k \in \mathbb{Z}\}$, cyclic subgroup generated by g (additive), p. 131
$\det \alpha$	determinant of the matrix α, p. 66
$M(\mathbb{Z}_m)$	set of 2×2 matrices with entries in $\mathbb{Z}_m$, p. 62
$G(\mathbb{Z}_m)$	$\{\alpha \in M(\mathbb{Z}_m) \mid \alpha \text{ has a multiplicative inverse}\}$, p. 99
$S(\mathbb{Z}_m)$	$\{\alpha \in M(\mathbb{Z}_m) \mid \det \alpha = 1\}$, p. 100
$f : S \to T$	f is a function from domain S to codomain T, p. 148
$G \cong H$	G is isomorphic to H, p. 160
$g \leftrightarrow h$	$g \in G$ corresponds to $h \in H$ when $G \cong H$, p. 128
$\ker \theta$	kernel of the homomorphism θ, p. 180
$\mathrm{im}\, \theta$	image of the homomorphism θ, p. 182
aH	$\{ah \mid h \in H\}$, left coset of H generated by a, p. 196
Ha	$\{ha \mid h \in H\}$, right coset of H generated by a, p. 196
$a + H$	$\{a + h \mid h \in H\}$, left coset of H for an *additive* group, p. 196
$\#G$	size of G, i.e., the number of its elements, p. 205
$[G : H]$	index of H in G, p. 209
G/H	quotient group G mod H, p. 226
gHg^{-1}	$\{ghg^{-1} \mid h \in H\}$, conjugate of H, p. 124
$G \times H$	$\{(g, h) \mid g \in G,\ h \in H\}$, direct product, p. 82
$\triangle G$	$\{(g, g) \mid g \in G\}$, diagonal subgroup, p. 107
R^*	group of units of the ring R, p. 268
$R[x]$	polynomial ring with coefficients in R, p. 282
$\deg f(x)$	degree of the polynomial $f(x)$, p. 283
$g(x) \mid f(x)$	$g(x)$ is a factor of $f(x)$, p. 292
$\langle a\rangle$	$\{a \cdot r \mid r \in R\}$, principal ideal generated by a, p. 318
$\langle a_1, a_2, \ldots, a_n\rangle$	$\{a_1 \cdot r_1 + a_2 \cdot r_2 + \cdots + a_n \cdot r_n \mid r_1, r_2, \ldots, r_n \in R\}$, p. 352
R/A	quotient ring, p. 324

Appendix D

Essential Theorems

Theorem 3.9 (GCD theorem for integers)**.** *Let a, $b \in \mathbb{Z}$. If $\gcd(a, b) = 1$, then there exist integers x and y such that $ax + by = 1$.* (**Note:** Its converse is also true. See Theorem 3.18.)

Theorem 4.19 (Multiplicative inverses in $\mathbb{Z}_m$)**.** *Let $a \in \mathbb{Z}_m$. Then $a \in U_m$ if and only if $\gcd(a, m) = 1$.*

Theorem 8.11 (Socks-shoes property)**.** *Let a and b be elements of a group. Then $(ab)^{-1} = b^{-1}a^{-1}$.*

Theorem 8.18 (Left cancellation)**.** *Let a, b, c be elements of a group. If $ab = ac$, then $b = c$.*

Theorem 9.10 (Sudoku property)**.** *Let G be a group. In each row or column of its group table, every element of G shows up exactly once.*

Theorem 12.16 (Division algorithm in $\mathbb{Z}$)**.** *Let a and b be integers, with $b > 0$. Then there exist q, $r \in \mathbb{Z}$ such that $a = b \cdot q + r$ with $0 \leq r < b$.*

Theorem 12.18. *Let g be an element of a group with $\operatorname{ord}(g) = n$. Then $n \mid k$ if and only if $g^k = \varepsilon$.*

Theorem 13.17. *Let g be an element of a group with $\operatorname{ord}(g) = n$. Then $\langle g \rangle = \{g^k \mid k \in \mathbb{Z}\}$ contains n distinct elements; namely $\langle g \rangle = \{\varepsilon, g^1, g^2, g^3, \ldots, g^{n-1}\}$, where $\varepsilon = g^0$.*

Theorem 14.12 (Subgroups of cyclic groups)**.** *Let G be a cyclic group, and let H be a subgroup of G. Then H is also cyclic.*

Theorem 14.15 (Subgroups of *finite* cyclic groups)**.** *Suppose G is cyclic with n elements. Then G has a unique subgroup of size d for every divisor d of n, and those are the only subgroups of G.*

Theorem 15.17. *Consider a function $f : S \to T$, where S and T are finite sets of the same size. Then f is one-to-one if and only if f is onto.*

Theorem 17.9. *Let $\theta : G \to H$ be a group homomorphism. Then θ maps the identity element of G to the identity element of H; i.e., $\theta(\varepsilon_G) = \varepsilon_H$.*

Theorem 17.13. *Let $\theta : G \to H$ be a group homomorphism. Then $\theta(g^k) = \theta(g)^k$ for all $g \in G$ and $k \in \mathbb{Z}$.*

Theorem 18.6 (Kernel is a subgroup)**.** *Let $\theta : G \to H$ be a group homomorphism. Then $K = \ker\theta$ is a subgroup of the domain G.*

Theorem 18.11 (Image is a subgroup)**.** *Let $\theta : G \to H$ be a group homomorphism. Then $I = \operatorname{im}\theta$ is a subgroup of the codomain H.*

Theorem 19.14. *Let G be a group, H a subgroup of G, and $a \in G$. Then:*

- *(Multiplicative) $aH = H$ if and only if $a \in H$.*
- *(Additive) $a + H = H$ if and only if $a \in H$.*

Theorem 19.16. *Let G be a group, H a subgroup of G, and $a, b \in G$. Then:*

- *(Multiplicative) $aH = bH$ if and only if $b^{-1} \cdot a \in H$ and $a^{-1} \cdot b \in H$.*
- *(Additive) $a + H = b + H$ if and only if $a - b \in H$ and $b - a \in H$.*

Theorem 20.5 (Lagrange's theorem)**.** *Let H be a subgroup of a finite group G. Then $\#H$ is a divisor of $\#G$.*

Theorem 20.12. *Let G be a finite group, and let $g \in G$. Then $\operatorname{ord}(g)$ is a divisor of $\#G$.*

Theorem 22.8. *Let G be a group, and let H be its subgroup. Assume G/H satisfies the coset multiplication shortcut. Given $aH \in G/H$, we have $(aH)^n = a^nH$ for all integer exponents n.*

Theorem 24.15 (Normal subgroup test)**.** *Let G be a group, and let H be a subgroup of G. Then H is normal in G if and only if $gHg^{-1} \subseteq H$ for all $g \in G$.*

Theorem 24.17 (Kernel is a normal subgroup)**.** *Let $\theta : G \to H$ be a group homomorphism, with $K = \ker\theta = \{a \in G \mid \theta(a) = \varepsilon_H\}$. Then K is normal in G.*

Theorem 25.3 (First Isomorphism Theorem for groups)**.** *Let $\theta : G \to H$ be a group homomorphism with $K = \ker\theta$. Then $G/K \cong \operatorname{im}\theta$, where $gK \in G/K$ corresponds to $\theta(g) \in \operatorname{im}\theta$.*

Theorem 27.6 (Cancellation in an integral domain)**.** *Let a, b, c be elements of an integral domain R and suppose $a \neq 0$. If $ab = ac$, then $b = c$.*

Theorem 28.14 (Degree of a product)**.** *Suppose R is an integral domain. Let $f(x)$, $g(x) \in R[x]$ with $f(x), g(x) \neq 0$. Then $\deg f(x) \cdot g(x) = \deg f(x) + \deg g(x)$.*

Theorem 29.5 (Division algorithm in $F[x]$)**.** *Let F be a field. Suppose $f(x), g(x) \in F[x]$ with $g(x) \neq 0$. Then there exist $q(x), r(x) \in F[x]$ such that $f(x) = g(x) \cdot q(x) + r(x)$ with $r(x) = 0$ or $\deg r(x) < \deg g(x)$.*

Theorem 29.9 (Factor theorem)**.** *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $f(a) = 0$ if and only if $(x - a) \mid f(x)$.*

Theorem 30.8. *Let F be a field. Suppose $f(x) \in F[x]$ with $\deg f(x) \geq 1$. Then $f(x)$ is unfactorable if and only if $f(x)$ satisfies the following property: If $f(x) = p(x) \cdot q(x)$, then $\deg p(x) = 0$ or $\deg q(x) = 0$.*

Theorem 30.16. *Let F be a field, and let $f(x) \in F[x]$ with $\deg f(x) \geq 2$. If $f(x)$ has a root $\alpha \in F$, then $f(x)$ is factorable in $F[x]$.*

Theorem 30.19. *Let F be a field, and let $f(x) \in F[x]$ with $\deg f(x) = 2$ or 3. If $f(x)$ has no root in F, then $f(x)$ is unfactorable in $F[x]$.*

Theorem 31.24 (Kernel is an ideal)**.** *Let $\theta : R \to S$ be a ring homomorphism with $K = \ker \theta$. Then K is an ideal of the domain R.*

Theorem 31.33. *Every ideal of $\mathbb{Z}$ is a principal ideal.*

Theorem 31.34. *Let F be a field. Then every ideal of $F[x]$ is a principal ideal.*

Theorem 32.11 (First Isomorphism Theorem for rings)**.** *Let $\theta : R \to S$ be a ring homomorphism with $K = \ker \theta$. Then there is a ring isomorphism $R/K \cong \operatorname{im} \theta$, where $r + K \in R/K$ corresponds to $\theta(r) \in \operatorname{im} \theta$.*

Theorem 35.1. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $F[x]/\langle g(x)\rangle$ is* not *a field.*

(b) *If $g(x)$ is unfactorable, then $F[x]/\langle g(x)\rangle$ is a field.*

Theorem 35.14 (GCD theorem for polynomials)**.** *Let F be a field, and let $f(x), g(x) \in F[x]$. If $f(x)$ and $g(x)$ are relatively prime, then there exist $p(x), q(x) \in F[x]$ such that $f(x) \cdot p(x) + g(x) \cdot q(x) = 1$.*

Theorem 36.14. *Let F be a field and fix $g(x) \in F[x]$.*

(a) *If $g(x)$ is factorable, then $\langle g(x)\rangle$ is* not *maximal in $F[x]$.*

(b) *If $g(x)$ is unfactorable, then $\langle g(x)\rangle$ is maximal in $F[x]$.*

Theorem 37.2. *Let M be an ideal of a commutative ring R.*

(a) *If M is* not *maximal in R, then R/M is* not *a field.*

(b) *If M is maximal in R, then R/M is a field.*

Index of Terms

AMS / MAA | TEXTBOOKS

A Friendly Introduction to Abstract Algebra offers a new approach to laying a foundation for abstract mathematics. Prior experience with proofs is not assumed, and the book takes time to build proof-writing skills in ways that will serve students through a lifetime of learning and creating mathematics.

The author's pedagogical philosophy is that when students abstract from a wide range of examples, they are better equipped to conjecture, formalize, and prove new ideas in abstract algebra. Thus, students thoroughly explore all concepts through illuminating examples before formal definitions are introduced. The instruction in proof writing is similarly grounded in student exploration and experience. Throughout the book, the author carefully explains where the ideas in a given proof come from, along with hints and tips on how students can derive those proofs on their own.

Readers of this text are not just consumers of mathematical knowledge. Rather, they are learning mathematics by *creating* mathematics. The author's gentle, helpful writing voice makes this text a particularly appealing choice for instructors and students alike. The book's website has companion materials that support the active-learning approaches in the book, including in-class modules designed to facilitate student exploration.

ISBN 978-1-4704-6881-1
9 781470 468811
TEXT/72

For additional information and updates on this book, visit
www.ams.org/bookpages/text-72

AMS
www.ams.org