An Inquiry-Based Approach to Abstract Algebra

Dana C. Ernst, PhD
Northern Arizona University

Fall 2016

Here is a partial list of people that I need to thank for supplying content, advice, and feedback.

- Ben Woodruff
- Josh Wiscons
- Dave Richeson
- Nathan Carter
- Appendix B: The Elements of Style for Proofs is a blending of work by Anders Hendrickson and Dave Richeson.
- Dave Richeson is the original author of Appendix C: Fancy Mathematical Terms and Appendix D: Definitions in Mathematics.

Contents

Chapter 1

Introduction

1.1 What is Abstract Algebra?

Abstract algebra is the subject area of mathematics that studies algebraic structures, such as groups, rings, fields, modules, vector spaces, and algebras. This course is an introduction to abstract algebra. We will spend most of our time studying groups. Group theory is the study of symmetry, and is one of the most beautiful areas in all of mathematics. It arises in puzzles, visual arts, music, nature, the physical and life sciences, computer science, cryptography, and of course, throughout mathematics. This course will cover the basic concepts of group theory, and a special effort will be made to emphasize the intuition behind the concepts and motivate the subject matter. In the last few weeks of the semester, we will also introduce rings and fields.

1.2 An Inquiry-Based Approach

In a typical course, math or otherwise, you sit and listen to a lecture. (Hopefully) These lectures are polished and well-delivered. You may have often been lured into believing that the instructor has opened up your head and is pouring knowledge into it. I absolutely love lecturing and I do believe there is value in it, but I also believe that in reality most students do not learn by simply listening. You must be active in the learning process. I'm sure each of you have said to yourselves, "Hmmm, I understood this concept when the professor was going over it, but now that I am alone, I am lost." In order to promote a more active participation in your learning, we will incorporate ideas from an educational philosophy called inquiry-based learning (IBL).

Loosely speaking, IBL is a student-centered method of teaching mathematics that engages students in sense-making activities. Students are given tasks requiring them to solve problems, conjecture, experiment, explore, create, communicate. Rather than showing facts or a clear, smooth path to a solution, the instructor guides and mentors students via well-crafted problems through an adventure in mathematical discovery. Effective IBL courses encourage deep engagement in rich mathematical activities and provide opportunities to collaborate with peers (either through class presentations or group-oriented work).

Perhaps this is sufficiently vague, but I believe that there are two essential elements to IBL. Students should as much as possible be responsible for:

1. Guiding the acquisition of knowledge, and
2. Validating the ideas presented. That is, students should not be looking to the instructor as the sole authority.

For additional information, check out my blog post, What the Heck is IBL?

Much of the course will be devoted to students proving theorems on the board and a significant portion of your grade will be determined by how much mathematics you produce. I use the word "produce" because I believe that the best way to learn mathematics is by doing mathematics. Someone cannot master a musical instrument or a martial art by simply watching, and in a similar fashion, you cannot master mathematics by simply watching; you must do mathematics!

Furthermore, it is important to understand that proving theorems is difficult and takes time. You should not expect to complete a single proof in 10 minutes. Sometimes, you might have to stare at the statement for an hour before even understanding how to get started.

In this course, everyone will be required to

- read and interact with course notes on your own;
- write up quality proofs to assigned problems;
- present proofs on the board to the rest of the class;
- participate in discussions centered around a student's presented proof;
- call upon your own prodigious mental faculties to respond in flexible, thoughtful, and creative ways to problems that may seem unfamiliar on first glance.

As the semester progresses, it should become clear to you what the expectations are. This will be new to many of you and there may be some growing pains associated with it.

Lastly, it is highly important to respect learning and to respect other people's ideas. Whether you disagree or agree, please praise and encourage your fellow classmates. Use ideas from others as a starting point rather than something to be judgmental about. Judgement is not the same as being judgmental. Helpfulness, encouragement, and compassion are highly valued.

1.3 Rules of the Game

You should *not* look to resources outside the context of this course for help. That is, you should not be consulting the Internet, other texts, other faculty, or students outside of our course. On the other hand, you may use each other, the course notes, me, and your own intuition. In this class, earnest failure outweighs counterfeit success; you need not feel pressure to hunt for solutions outside your own creative and intellectual reserves. For more details, check out the Syllabus.

1.4 Structure of the Notes

As you read the notes, you will be required to digest the material in a meaningful way. It is your responsibility to read and understand new definitions and their related concepts. However, you will be supported in this sometimes difficult endeavor. In addition, you will be asked to complete exercises aimed at solidifying your understanding of the material. Most importantly, you will be asked to make conjectures, produce counterexamples, and prove theorems.

Most items in the notes are labelled with a number. The items labelled as **Definition** and **Example** are meant to be read and digested. However, the items labelled as **Exercise**, **Question**, **Theorem**, **Corollary**, and **Problem** require action on your part. In particular, items labelled as **Exercise** are typically computational in nature and are aimed at improving your understanding of a particular concept. There are very few items in the notes labelled as **Question**, but in each case it should be obvious what is required of you. Items with the **Theorem** and **Corollary** designation are mathematical facts and the intention is for you to produce a valid proof of the given statement. The main difference between a **Theorem** and **Corollary** is that corollaries are typically statements that follow quickly from a previous theorem. In general, you should expect corollaries to have very short proofs. However, that doesn't mean that you can't produce a more lengthy yet valid proof of a corollary. The items labelled as **Problem** are sort of a mixed bag. In many circumstances, I ask you to provide a counterexample for a statement if it is false or to provide a proof if the statement is true. Usually, I have left it to you to determine the truth value. If the statement for a problem is true, one could relabel it as a theorem.

It is important to point out that there are very few examples in the notes. This is intentional. One of the goals of the items labelled as **Exercise** is for you to produce the examples.

Lastly, there are many situations where you will want to refer to an earlier definition or theorem/corollary/problem. In this case, you should reference the statement by number. For example, you might write something like, "By Theorem 1.13, we see that. . . ."

1.5 Some Minimal Guidance

Especially in the opening sections, it won't be clear what facts from your prior experience in mathematics we are "allowed" to use. Unfortunately, addressing this issue is difficult and is something we will sort out along the way. However, in general, here are some minimal and vague guidelines to keep in mind.

First, there are times when we will need to do some basic algebraic manipulations. You should feel free to do this whenever the need arises. But you should show sufficient work along the way. You do not need to write down justifications for basic algebraic manipulations (e.g., adding 1 to both sides of an equation, adding and subtracting the same amount on the same side of an equation, adding like terms, factoring, basic simplification, etc.).

On the other hand, you do need to make explicit justification of the logical steps in a proof. When necessary, you should cite a previous definition, theorem, etc. by number.

Unlike the experience many of you had writing proofs in geometry, our proofs will be written in complete sentences. You should break sections of a proof into paragraphs and use proper grammar. There are some pedantic conventions for doing this that I will point out along the way. Initially, this will be an issue that most students will struggle with, but after a few weeks everyone will get the hang of it.

Ideally, you should rewrite the statements of theorems before you start the proof. Moreover, for your sake and mine, you should label the statement with the appropriate number. I will expect you to indicate where the proof begins by writing "*Proof.*" at the beginning. Also, we will conclude our proofs with the standard "proof box" (i.e., □ or ■), which is typically right-justified.

Lastly, every time you write a proof, you need to make sure that you are making your assumptions crystal clear. Sometimes there will be some implicit assumptions that we can omit, but at least in the beginning, you should get in the habit of stating your assumptions up front. Typically, these statements will start off "Assume..." or "Let...".

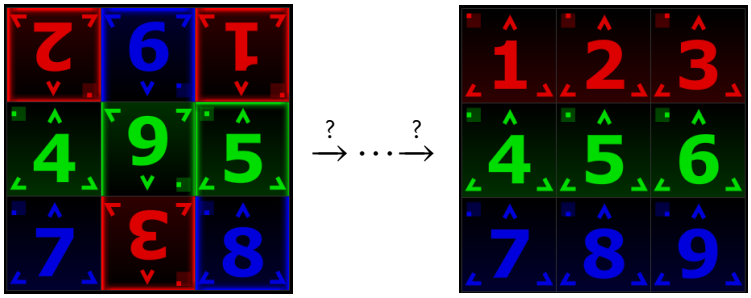
This should get you started. We will discuss more as the semester progresses. Now, go have fun and kick some butt!

Chapter 2

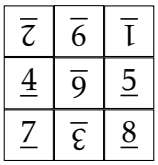
An Intuitive Approach to Groups

One of the major topics of this course is **groups**. The area of mathematics that is concerned with groups is called **group theory**. Loosely speaking, group theory is the study of symmetry, and in my opinion is one of the most beautiful areas in all of mathematics. It arises in puzzles, visual arts, music, nature, the physical and life sciences, computer science, cryptography, and of course, throughout mathematics.

Instead of starting with an abstract formal definition, we will begin our study of groups by developing some intuition about what groups actually are. To get started, we will be exploring the game Spinpossible^{TM}(which used to be available for iOS and Android devices). The game is played on a 3×3 board of scrambled tiles numbered 1 to 9, each of which may be right-side-up or up-side-down. The objective of the game is to return the board to the standard configuration where tiles are arranged in numerical order and right-side-up. This is accomplished by a sequence of "spins", where a spin consists of rotating an $m \times n$ subrectangle by 180°. The goal is to minimize the number of spins used. The following figure depicts a scrambled board on the left and the solved board on the right. The sequence of arrows is used to denote some sequence of spins that transforms the scrambled board into the solved board.

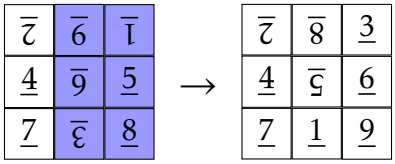


Example 2.1. Let's play with an example. Suppose we start with the following scrambled board.

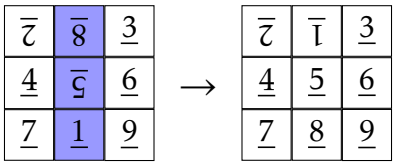


The underlines on the numbers are meant to help us tell whether a tile is right-side-up or up-side-down. Our goal is to use a sequence of spins to unscramble the board. Before we get started, let's agree on some conventions. When we refer to tile n, we mean the actual tile that is labeled by the number n regardless of its position and orientation on the board. On the other hand, position n will refer to the position on the board that tile n is supposed to be in when the board has been unscrambled. For example, in the board above, tile 1 is in position 3 and tile 7 happens to be in position 7.

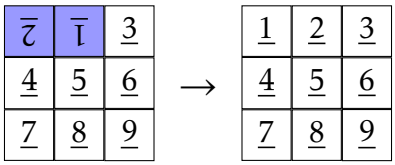
It turns out that there are multiple ways to unscramble this board., but I have one particular sequence in mind. First, let's spin the rectangle determined by the two rightmost columns. Here's what we get. I've shaded the subrectangle that we are spinning.



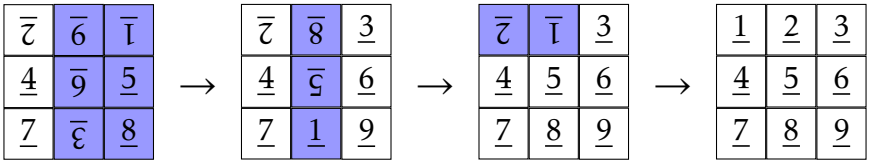
Okay, now let's spin the middle column.



Hopefully, you can see that we are really close to unscrambling the board. All we need to do is spin the rectangle determined by the tiles in positions 1 and 2.



Putting all of our moves together, here is what we have.



In this case, we were able to solve the scrambled board in 3 moves. It's not immediately obvious, but it turns out that there is no way to unscramble the board in fewer than 3 spins. However, there is at least one other solution that involves exactly 3 spins. We won't worry about proving this; right now we are just trying to gain some intuition.

Exercise 2.2. Without worrying about whether your solution is optimal, try to find a different sequence of spins that unscrambles the initial board in Example 2.1. Your answer should be a sequence of spins. Describe your sequence in a way that makes sense. Can you find a sequence of 3 spins that is different from the one described in Example 2.1 that unscrambles the board?

Exercise 2.3. How many scrambled 3×3 Spinpossible boards are there? To answer this question, you will need to rely on some counting principles such as factorials. *Note:* In this context, we want to include the solved board as one of the scrambled boards. It's just not very scrambled.

Exercise 2.4. A natural question to ask is whether every possible scrambling of a board in Spinpossible can be unscrambled using only spins. It turns out that the answer is yes. Justify this fact by describing an algorithm that will always unscramble a scrambled board. It does not matter whether your algorithm is efficient. That is, we don't care how many steps it takes to unscramble the board as long as it works in a finite number of steps. Also, if it didn't occur to you yet, we can always spin a single tile (referred to as *toggling* a tile).

Exercise 2.5. Does the order in which you apply spins matter? Does it always matter? Let's be as specific as possible. If the order in which we apply two spins does not matter, then we say that the spins **commute**. However, if the order does matter, then the spins do not commute. When will two spins commute? When will they not commute? Provide some specific examples.

Exercise 2.6. How many possible spins are there? We are referring to the moves you are allowed to do at any stage in the game. Don't forget that you are allowed to toggle a single tile.

In a 2011 paper, Alex Sutherland and Andrew Sutherland (a father and son team) present a number of interesting results about Spinpossible and list a few open problems. You can find the paper at http://arxiv.org/abs/1110.6645. As a side note, Alex is one of the developers of the game and his father, Andrew, is a mathematics professor at MIT. Using a brute-force computer algorithm, the Sutherlands verified that every scrambled 3×3 board can be solved in at most 9 moves. However, a human readable mathematical proof of this fact remains elusive. By the way, mathematics is chock full of open problems and you can often get to the frontier of what is currently known without too much trouble. Mathematicians are in the business of solving open problems.

At least for now, let's ignore the optimality requirement of the game. That is, let's not worry about how many spins it takes to solve a scrambled board. It turns out that we can "build" some spins from other spins. As an example, if I wanted to toggle the tile in position 2, I could first spin the rectangle determined by positions 1 and 2, then toggle the tile in position 1, and lastly spin the rectangle determined by positions 1 and 2 again. Of course, this is horribly inefficient, but it works. Also, it is important to point out that I was describing the tile positions we were spinning while not paying any attention to the tiles occupying the corresponding positions.

It's not too difficult to prove that we can build all of the possible spins by only using the following spins. I've listed some shorter names for these spins in parentheses.

1. Toggle position 1 (t),
2. Spin rectangle determined by positions 1 and 2 (s_1),
3. Spin rectangle determined by positions 2 and 3 (s_2),

4. Spin rectangle determined by positions 3 and 6 (s_3),
5. Spin rectangle determined by positions 6 and 5 (s_4),
6. Spin rectangle determined by positions 5 and 4 (s_5),
7. Spin rectangle determined by positions 4 and 7 (s_6),
8. Spin rectangle determined by positions 7 and 8 (s_7),
9. Spin rectangle determined by positions 8 and 9 (s_8).

We can describe any of the allowable spins in the game by writing down a sequence consisting of $t, s_1, s_2, \ldots, s_8$.

Example 2.7. Spinning the subrectangle determined by positions 1 and 4 is an allowable spin, but it's not on our list above. We can build this spin by using the following sequence of spins:

$$s_1 \to s_2 \to s_3 \to s_4 \to s_5 \to s_4 \to s_3 \to s_2 \to s_1.$$

Exercise 2.8. Toggling the tile in position 3 is an allowable spin. Try to find a sequence of spins involving $t, s_1, s_2, \ldots, s_8$ only that yields this toggle.

In addition to building all of the allowable spins, we can also describe any possible rearrangement of tiles (position and/or orientation) using just these 9 spins. For example, if we apply s_2, followed by s_3, and then s_2 again, the net result is swapping the tiles in positions 2 and 6 while maintaining their orientation. You should take the time to verify this. However, notice that the net action is *not* an allowable spin. That is, not every sequence of the 9 spins $t, s_1, s_2, \ldots, s_8$ results in an allowable spin.

Exercise 2.9. What is the net action of applying s_1, then s_2, and then s_1? Is the net action an allowable spin? How about s_2, then s_1, and then s_2?

We say that the set $\{t, s_1, \ldots, s_8\}$ **generates** all possible scramblings of the 3×3 board. In this case, we refer to $\{t, s_1, \ldots, s_8\}$ as a set of **generators**. It turns out that this generating set is minimal in the sense that if we tried to get rid of any one of $t, s_1, \ldots, s_8$, we would no longer be able to generate all scramblings. Note that there are other minimal generating sets and there are lots of sets that will generate all the scramblings that are not minimal.

We need to establish some conventions about how to write down sequences of spins involving the generators. Since we are doing spins on top of spins, we will follow the convention of function notation that says the function on the right goes first. For example, ts_1s_3 means do s_3 first, then do s_1, and lastly do t. This will take some getting used to, but just remember that it is just like function notation (stuff on the right goes first). We will refer to sequences like ts_1s_3 as **words** in the generators $t, s_1, \ldots, s_8$. We can also use exponents to abbreviate. For example, s_2^2 is the same as s_2s_2 (which in this case has the net action of doing nothing) and $(s_1s_2)^2$ is the same as $s_1s_2s_1s_2$.

Exercise 2.10. It turns out that there is an even simpler word (i.e., a shorter word) that yields the same net action as $(s_1s_2)^2$. Can you find one?

Exercise 2.11. Try to write the spin that rotates the entire top row (i.e., spin the top row) as a sequence of moves involving only $t, s_1, \ldots, s_8$.

Let's make a couple more observations. First, every spin is reversible (i.e., has an *inverse*). In this case, we could just apply the same spin again to undo it. For example, s_1^2 is the same as doing nothing. This means that the reverse of s_1, denoted s_1^{-1}, is s_1 itself. Symbolically, we write $s_1^{-1} = s_1$. *Warning:* Remember that we are exploring the game Spinpossible; it won't always be the case that repeating a generator will reverse the action. In the same vein, every sequence of spins is reversible. For example, if we apply $s_1 s_2$ (remember that's do s_2 first and then s_1) to some scrambled board, we could undo the net action by applying $s_2 s_1$. That is, the reverse (or inverse) of $s_1 s_2$ is $s_2 s_1$. Written symbolically, we have

$$(s_1 s_2)^{-1} = s_2^{-1} s_1^{-1} = s_2 s_1$$

since $s_2^{-1} = s_2$ and $s_1^{-1} = s_1$.

Exercise 2.12. Imagine we started with a scrambled board and you were then able to unscramble the board using some sequence from $t, s_1, \ldots, s_8$. In this case, you would have some word in $t, s_1, \ldots, s_8$ (with repeats allowed). Let's call it w. Now, imagine you have the solved board. How could you obtain the scrambled board that w unscrambled using only $t, s_1, \ldots, s_8$? How is this related to w^{-1}?

The upshot of the previous exercise is that the action of any sequence of generators can be reversed and is itself an action.

At this time, I think we are ready to summarize some of our observations of the game Spinpossible and to make a few general claims, which we will state as a list of rules.

Rule 1. There is a predefined list of actions that never changes.

Rule 2. Every action is reversible.*

Rule 3. Every action is deterministic.

Rule 4. Any sequence of consecutive actions is also an action.

Rule 1 states that we must start with some fixed set of actions. These are our generators. In the case of Spinpossible, we encountered two possible generating sets. First, there was the set of allowable spins, which you counted in Exercise 2.6. Second, we considered the set $\{t, s_1, \ldots, s_8\}$, which is a much smaller list of predefined actions.

Rule 2 tells us that every action given in Rule 1 has an inverse. In the case of Spinpossible, every predefined spin is its own inverse.

By deterministic, we mean that we know exactly what will happen when we apply an action. In contrast, pulling a card off the top of a shuffled deck of cards is not deterministic because we don't know which card we will end up with. Certainly, every spin is deterministic. For example, if we apply s_6, we know exactly what will happen.

Rule 4 provides us with a way to build new actions from the actions given in Rule 1. For example, if we are given $\{t, s_1, \ldots, s_8\}$ as our predefined list of actions (Rule 1), then Rule 4 guarantees that $s_1 s_2 s_3 t$ is also an action (but does not have to be a spin).

*Implicit in this rule is that the reverse of an action is also an action.

Exercise 2.13. Notice that there is no explicit rule that says that every sequence of consecutive actions is reversible. Is this a consequence of Rules 1–4? Explain your answer.

Alright, we are finally ready for our intuitive and unofficial definition of a group.

Intuitive Definition 2.14. A **group** is a system or collection of actions that satisfies Rules 1–4 above.

Our first example of a group is the set of actions that rearranges and reorients the tiles on the 3×3 Spinpossible board. Notice that I didn't say that the set of scrambled boards was a group. It turns out that there is a one-to-one correspondence between actions for Spinpossible and scrambled boards, but for now let's focus on the actions.

Exercise 2.15. Describe how the Rubik's Cube fits into the framework of Rules 1–4.

Exercise 2.16. Place two coins side by side on a table. Consider just one predefined action: swapping the positions of the two coins. Can we form a group of actions using this one action as our starting point? If so, completely describe the collection of actions. If not, explain why.

Exercise 2.17. Consider Exercise 2.16, but add a third coin to the right of the other two coins. The only predefined action is still the one from the previous exercise: swapping the positions of the two leftmost coins. Can we form a group of actions using this one action as our starting point? If so, completely describe the collection of actions. If not, explain why. How does your answer to this exercise compare to the previous?

Exercise 2.18. Consider your three coins from the previous exercise. Now, for your actions take all possible actions of rearranging the coins. It turns out that this is a group.

(a) One of the actions is to swap the second and third coins. What happens if you do this action twice? Is this an action?

(b) How many actions does this group have? Describe them all.

(c) Can you think of a small set of actions that would generate all the other actions? Can you find a minimal one (in the sense that removing one of your initial actions would result in a different group)? Write each of the actions of this group as a word in your generators? Do some actions have more than one word representing it?

In part (a) of the previous exercise you encountered the "do-nothing" action, which we will refer to as the **identity** of the group.

Exercise 2.19. Explain why every group has a do-nothing action (i.e., an identity).

Exercise 2.20. Imagine you have 10 coins in your left pocket. Consider two actions: (1) move a coin from your left pocket to your right pocket, and (2) move a coin from your right pocket to your left pocket. Is this a group? Explain your answer.

Exercise 2.21. Imagine you have a square puzzle piece that fits perfectly in a square hole. Consider these actions: pick up the square and rotate it an appropriate amount so that it fits back in the hole. Is this a group? Explain your answer. If it is a group, how many distinct actions are there?

Exercise 2.22. Can you describe a group that has exactly n actions for any natural number n?

Exercise 2.23. Can you describe a situation that satisfies Rules 1–3, but not Rule 4?

Exercise 2.24. Pick your favorite integer. Consider these actions: add any integer to the one you chose. This is an infinite set of actions. Is this a group? If so, how small a set of generators can you find?

Exercise 2.25. Consider the previous exercise, but this time multiply instead of add. Is this a group? Explain your answer.

Chapter 3

Cayley Diagrams

Recall that in the previous chapter we defined a group to be a set of actions that satisfies the following rules.

Rule 1. There is a predefined list of actions that never changes.

Rule 2. Every action is reversible.

Rule 3. Every action is deterministic.

Rule 4. Any sequence of consecutive actions is also an action.

It is important to point out that this is an intuitive starting point and does not constitute the official definition of a group. We'll continue to postpone a rigorous definition in this chapter and instead we will focus on developing more intuition about what groups are and what they "look like."

To get started, let's continue thinking about the game Spinpossible (see Chapter 2). In Exercise 2.3, we discovered that there are a total of $2^9 \cdot 9! = 185,794,560$ possible scrambled Spinpossible boards. Now, imagine we wanted to write a solution manual that would describe how to solve all these boards. There are likely many possible ways to construct such a solution manual, but here is one way.

The manual will consist of $185,794,560$ pages such that each page lists a unique scrambling of the 3×3 board. Don't forget that one of these scramblings is the solved board, which we will make page 1. Also, imagine that the book is arranged in such a way that it isn't too difficult to look up a given scrambled board. On each page below the scrambled board is a table that lists all possible spins. Next to each spin, the table indicates whether doing that particular spin will result in a board that is either closer to being solved or farther. In addition, the page number that corresponds to the resulting board is listed next to each spin.

In most cases, there will be many spins that take us closer to the solved board. Given a scrambled board, a solution would consist of following one possible sequence of pages through the book that takes us from the scrambled board to the solved board. There could be many such sequences. If we could construct such a solution manual, we would have an atlas or map for the game Spinpossible.

Note that even if we make a wrong turn (i.e., follow a page that takes us farther away from the solution), we can still get back on track by following page numbers that take us closer to the solved board. In fact, we can always flip back to the page we were on before taking a wrong turn. This page will be listed on our "wrong turn page" since doing the same spin twice has the net effect of doing nothing. If you were to actually do this, the number of pages we would need to visit would be longer than an optimal solution, but we'd get to the solved board nonetheless.

Let's get a little more concrete. Consider the game Spinpossible, except let's simplify it a little. Instead of playing on the 3×3 board, let's play on a 1×2 board consisting of a single row with tiles labeled 1 and 2. The rules of the game are what you would expect; we are restricted to spins involving just the tiles in positions 1 and 2 of the original board. A scrambling of the 1×2 Spinpossible board consists of any rearrangement of the tiles 1 and 2, where either of the tiles can be right-side-up or up-side-down.

Exercise 3.1. First, convince yourself that the set of actions corresponding to the 1×2 Spinpossible board satisfies our four rules of a group. We'll refer to this group as $\text{Spin}_{1\times 2}$.

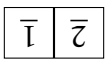
(a) How many scrambled boards are there for the 1×2 Spinpossible game? Don't forget to include the solved board.

(b) How many actions are there in $\text{Spin}_{1\times 2}$? Which of these actions are spins? *Hint:* There are actions that are not spins.

Let's try to make a map for $\text{Spin}_{1\times 2}$, but instead of writing a solution manual, we will draw a picture of the group called a **Cayley diagram**. The first thing we'll do is draw each of the scramblings that we found in the previous exercise. It doesn't matter how we arrange all of these drawings, as long as there is some space between them. Now, for each scrambling, figure out what happens when we do each of our allowable spins. For each of these spins, we'll draw an arrow from the initial scrambled board to the resulting board. Don't worry about whether doing each of these spins is a good idea or not. In fact, figure out what happens when we do our allowable spins to the solved board, as well. In this case, each of our scrambled boards will have 3 arrows heading out towards 3 distinct boards. Do you see why?

In order for us to keep straight what each arrow represents, let's color our arrows, so that doing a particular type of spin is always the same color. For example, we could color the arrows that toggle the tile in the first position as green. Recall that doing the same spin twice has the net effect of doing nothing, so we should just make all of our arrows point in both directions.

To make sure you are keeping up to speed, consider the following scrambled board.



This board is one of our 8 possible scrambled 1×2 boards. We have three possible spins we can do to this board: toggle position 1, toggle position 2, or spin the whole board. Each of these spins has a corresponding two-way arrow that takes us to three different scrambled boards. Figure 3.1 provides a visual representation of what we just discussed.

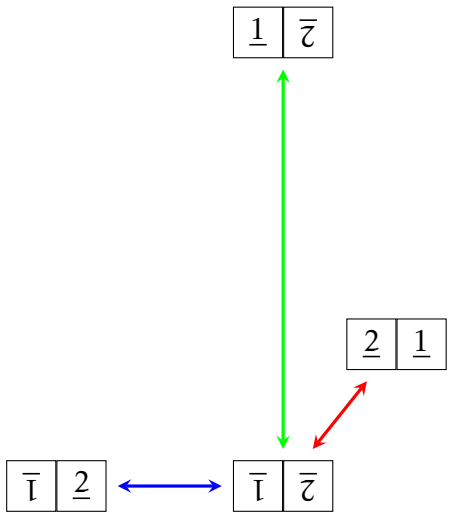


Figure 3.1. A portion of the Cayley diagram for $\text{Spin}_{1\times 2}$ with generating set $\{t_1, t_2, s\}$.

Note that I could have drawn the four scrambled boards in Figure 3.1 anywhere I wanted to, but I have a particular layout in mind. Also, notice we have three different colored arrows. Can you see what each of the colors corresponds to? In this case, a green arrow corresponds to toggling the tile in position 1, a blue arrow corresponds to toggling position 2, and a red arrow corresponds to spinning the whole board.

If we include the rest of the scrambled boards and all possible spins, we obtain Figure 3.2. Note that I've chosen a nice layout for the figure, but it's really the connections between the various boards that are important.

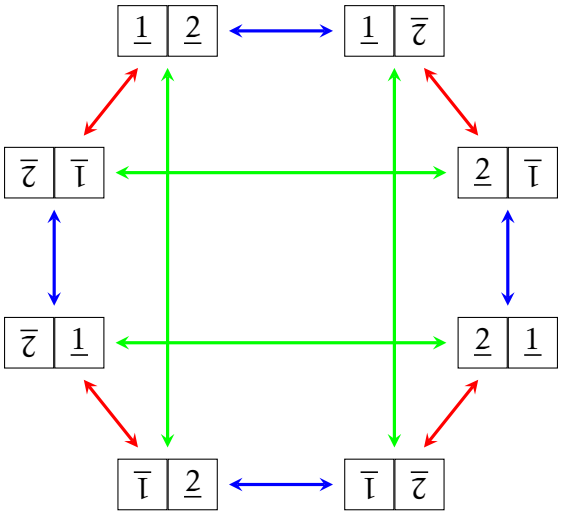


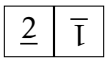
Figure 3.2. Cayley diagram for $\text{Spin}_{1\times 2}$ with generating set $\{t_1, t_2, s\}$.

In this case, the spins that correspond to the three arrow colors are the **generators** of $\text{Spin}_{1\times 2}$. What this means is that we can obtain all possible scramblings/unscramblings by using just these 3 spins. Let t_1 be the spin that toggles position 1, t_2 be the spin that

toggles position 2, and s be the spin that rotates the full board.

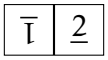
In order to obtain the Cayley diagram for $\text{Spin}_{1\times 2}$ (with the generators we have in mind), we need to identify each scrambled board in Figure 3.2 with an action from the group. The most natural choice is to identify the solved board with the do-nothing action, which we will denote by e. As soon as we make this choice, we can just follow the arrows around the diagram to determine which actions correspond to which scrambled boards.

For example, consider the following scrambled board.



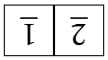
Looking at Figure 3.2, we see that one way to get to this board from the solved board is to follow a blue arrow and then a red arrow. This corresponds to the word st_2. (Recall that when we write down words, we should apply the actions from right to left, just like function composition.) However, it also corresponds to the word $t_2st_2t_1$ even though this is not an optimal solution. So, we can label the board in question with either st_2 or $t_2st_2t_1$ (there are other choices, as well).

As another example, consider the following scrambled board.



To get here from the solved board, we can simply follow a green arrow. So, this scrambled board corresponds to t_1. However, we could also follow a red arrow, then a blue arrow, and then a red arrow. Thus, we could also label the scrambled board by st_2s.

Exercise 3.2. Using Figure 3.2, find three distinct words in t_1, t_2, and s that correspond to the following scrambled board. Don't worry about whether your word is of optimal length or not.



Exercise 3.3. Label each of the remaining boards from Figure 3.2 with at least one appropriate word using t_1, t_2, and s. The diagram in Figure 3.2 together with your labels is the Cayley diagram for $\text{Spin}_{1\times 2}$ with generating set $\{t_1, t_2, s\}$.

It is important to point out that each word that corresponds to a given scrambled board tells you how to reach that scrambled board from the solved board (which is labeled by e, the do-nothing action).

Exercise 3.4. Given a word that corresponds to a scrambled board in Figure 3.2, how could we obtain a solution to the scrambled board? That is, how can we return to the solved board?

Exercise 3.5. Consider the Cayley diagram for $\text{Spin}_{1\times 2}$ in Figure 3.2, but remove all the red arrows. This corresponds to forbidding the spin that rotates the full 1×2 board. Can we obtain all of the scrambled boards from the solved board using only blue and green arrows?

Exercise 3.6. Repeat the previous exercise, but this time remove only the green arrows. What about the blue arrows?

In general, a **Cayley diagram** for a group G is a digraph having the set of actions of G as its vertices and the directed edges (i.e., arrows) correspond to the generators of the group. Following an arrow forward corresponds to applying the corresponding action. Recall that the generators are a potentially smaller set of actions from which you can derive all the actions of the group. The way you can derive new actions is by forming words in the generators (i.e., follow a sequence of arrows). Rule 2 guarantees that every action is reversible, so we also allow the use of a generator's reversal in our words (i.e., follow an arrow backwards).

If a generator is its own reversal, then the arrows corresponding to that generator are two-way arrows. It is always true that following an arrow backwards corresponds to a generator's reversal. That is, if an arrow corresponds to the action a, then the inverse of a, namely a^{-1}, corresponds to the reverse arrow.

Notice that all the arrows in the Cayley diagram for $\text{Spin}_{1\times 2}$ given in Figure 3.2 are two-way arrows. This means that every generator is its own inverse (in the case of $\text{Spin}_{1\times 2}$). It's important to point out that this is not true in general (i.e., we may have one-way arrows that correspond to generators that are not their own inverses).

We need a way to tell our arrow types apart. One way to do this is to color them. Another way would be to label the arrows by their corresponding generator.

Remember that in any group there is always a do-nothing action and one of the vertices should be labeled by this action. From this point forward, unless someone says otherwise, let's use e to denote our do-nothing action for a group. Each vertex is labeled with a word that corresponds to the sequence of arrows that we can follow from the do-nothing action to the particular vertex. Since there are possibly many sequences of arrows that could take us from the do-nothing vertex to another, each vertex could be labeled with many different words.

In our Cayley diagram for $\text{Spin}_{1\times 2}$, our vertices were fancy pictures of scrambled 1×2 Spinpossible boards. This wasn't necessary, but is convenient and appealing for aesthetic reasons. After labeling the solved board with the do-nothing action, e, in Exercise 3.3 you labeled each remaining vertex of the diagram with a word that corresponds to a sequence of arrows from the solved board to the vertex in question.

The next two exercises may be too abstract for you at the moment. Give them a shot and if you can't do them now, come back to them after you've constructed a few Cayley diagrams.

Exercise 3.7. Assume G is a group of actions and S is a set of generators for G. Suppose we draw the Cayley diagram for G using the actions of S as our arrows and we color the arrows according to which generator they correspond to. Assume that each vertex is labeled with a word in the generators or their reversals. If the arrows are not labeled, how can you tell which generator they correspond to?

Exercise 3.8. Assume G is a group. Suppose that S and S' are two different sets that generate G. If you draw the Cayley diagram for G using S and then draw the Cayley diagram for G using S', what features of the two graphs are the same and which are potentially different?

Let's build a few more Cayley diagrams to further our intuition.

Exercise 3.9. Consider the group consisting of the actions that rearranges two coins (but we won't flip them over), say a penny and a nickel. Let's assume we start with the penny on the left and the nickel on the right. Let's call this group S_2.

(a) Write down all possible actions using verbal descriptions. *Hint:* There aren't that many of them.

(b) Let s be the action that swaps the left and right coins. Does s generate S_2? That is, can we write all of the actions of S_2 as words in s (or its reversal)?

(c) Decide on a simple generating set for S_2 and draw a Cayley diagram for S_2 using your generating set. Label all the vertices and arrows appropriately. Recall that above we said that we will use e to denote the do-nothing action unless someone says otherwise.

Exercise 3.10. Consider a square puzzle piece that fits perfectly into a square hole. Let R_4 be the group of actions consisting of rotating the square by an appropriate amount so that it fits back into the hole.

(a) Write down all possible actions using verbal descriptions. Are there lots of ways to describe each of your actions?

(b) Let r be the action that rotates the puzzle piece by 90° clockwise. Does r generate R_4? If so, write down all of the actions of R_4 as words in r.

(c) Which of your words above is the reversal of r? That is, can we describe r^{-1} using r?

(d) Draw the Cayley diagram for R_4 using r as the generator. Be sure to label the vertices and arrows. Are your arrows one-way or two-way arrows?

We will refer to R_4 as the group of rotational symmetries of a square. In general, R_n is the group of rotational symmetries of a regular n-gon.

Exercise 3.11. Consider a puzzle piece like the one in the previous exercise, except this time, let's assume that the piece and the hole are an equilateral triangle. Let D_3 be the group of actions that allow the triangle to fit back in the hole. In addition to rotations, we will also allow the triangle to be flipped over. To give us a common starting point, let's assume the triangle and hole are positioned so that one of the tips of the triangle is pointed up. Also, let's label both the points of the hole and the points of the triangle with the numbers 1, 2, and 3. Assume the labeling on the hole starts with 1 on top and then continues around in the obvious way going clockwise. Label the puzzle piece in the same way and let's assume that the triangle starts in the position that has the labels matching (i.e., the point of the triangle labeled 1 is in the corner of the hole labeled 1, etc.).

(a) How many actions are there? Can you describe them? One way to do this would be to indicate where the labels of the triangle are in the hole.

(b) Let r be rotation by 120° in the clockwise direction. Does r generate D_3? That is, can you write each of your actions from part (a) as words in r?

(c) What is the reversal of r? That is, what is r^{-1}? Can you write it as a word in r?

(d) Let s be the flip (or we could call it a reflection) that swaps the corners of the puzzle piece that are in the positions of the hole labeled by 2 and 3 (this leaves the corner in position 1 of the hole in the same spot). Does s generate D_3?

(e) What is the reversal of s? That is, what is s^{-1}? Can you write it as a word?

(f) Can we generate all of D_3 using both r and s? If so, write all the actions of D_3 as words in r and s (or their reversals/inverses).

(g) Draw the Cayley diagram for D_3 using r and s as your arrows. *Hints:* One of your arrow types is one-way and the other is two-way. I suggest putting half the vertices in a circle and then the other half in a concentric circle outside your first half. Label one of the vertices on the inner circle as e and first think about applying consecutive actions of r. Try to stay on the inner circle of vertices as you do this. Now, starting at e, apply s and go to one of the vertices in the outer circle. Try to label the remaining vertices using both r and s. There are multiple ways to label each of the vertices.

Exercise 3.12. Repeat the above exercise, but do it for a square instead of a triangle. You'll need to make some modifications to r and s. The resulting group is called D_4.

The groups D_3 and D_4 are the group of symmetries (rotations and reflections) of an equilateral triangle and a square, respectively. In general, D_n is the group of symmetries of a regular n-gon and is referred to as the **dihedral group** of order $2n$. In this case, the word "order" simply means the number of actions in the group. We will encounter a formal definition of order in Chapter 4. Why does D_n consist of $2n$ actions?

Exercise 3.13. Consider the group from Exercise 2.24. Using "add 1" (or simply 1) as the generator*, describe what the Cayley diagram for this group would look like. Draw a chunk of the Cayley diagram. Can you think of another generating set? What will the Cayley diagram look like in this case?

Now that you've constructed a few examples for yourself, you should have a pretty healthy understanding of Cayley diagrams. There are still lots of properties to discover and opportunities to gain more intuition. If you weren't able to complete exercises 2.24 and 3.8, go give them another shot.

By the way, Cayley diagrams are named after their inventor Arthur Cayley, a nineteenth century British mathematician. We'll see his name pop up a couple more times in the course.

Not only are Cayley diagrams visually appealing, but they provide a map for the group in question. That is, they provide a method for navigating the group. Following sequences of arrows tells us how to do an action. However, each Cayley diagram very much

*Recall that Rule 2 guarantees that every action is reversible. So, if we have "add 1", we also have "add –1."

depends on the set of generators that are chosen to generate the group. If we change the generating set, we may end up with a very different looking Cayley diagram. This was the point of Exercise 3.8. It's important to drive this point home, so let's construct an explicit example.

Exercise 3.14. In Exercise 3.11, you constructed the Cayley diagram for the group called D_3. In this case, you used the generators r and s. Now, let s' be the reflection that swaps the corners of the triangle that are in the corners of the hole labeled by 1 and 2.

(a) Justify that s and s' generate all of D_3. *Hint:* Is it enough to generate r with s and s'?

(b) Construct the Cayley diagram for D_3 using s and s' as your generators. Did you get a different diagram than you did in Exercise 3.11?

Let's do a few more exercises involving Cayley diagrams.

Exercise 3.15. Consider the Cayley diagram for the group that we will call R_6 given in Figure 3.3.

(a) Assuming e is the do-nothing action, which action is the generator of the group?

(b) Describe the inverse of each of the 6 actions as a word in r.

(c) Can you find a shorter word to describe r^8?

(d) Does r^2 generate the group? How about r^5? Explain your answers.

(e) Describe a concrete collection of actions that would yield this Cayley diagram.

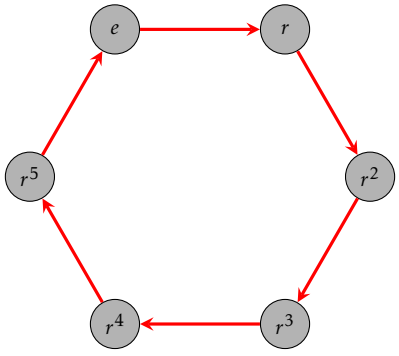


Figure 3.3. Cayley diagram for R_6 with generator r.

We haven't explicitly defined what a Cayley diagram actually is yet. So, it's not completely obvious that the diagram in the previous exercise is actually a diagram for a group. But rest assured; this Cayley diagram truly does correspond to a group. It's important to point out that we can't just throw together a digraph willy nilly and expect it to be a Cayley diagram.

Exercise 3.16. Consider the diagram given in Figure 3.4. Explain why the diagram cannot possibly be a Cayley diagram for a group. How many reasons can you come up with?

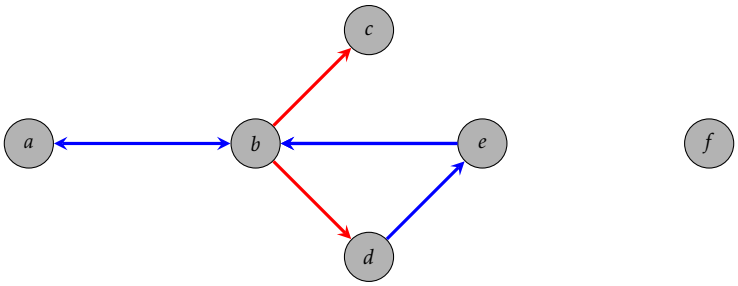


Figure 3.4

Exercise 3.17. Let G be a group of actions and suppose S is a set of generators for G. Suppose we draw the Cayley diagram for G using the actions of S as our arrows and we color the arrows according to which generator they correspond to.

(a) Explain why there must be a sequence of arrows (forwards or backwards) from the vertex labeled e to every other vertex. Do you think this is true for every pair of vertices?

(b) Recall that G must satisfy Rule 1. What restriction does this put on our Cayley diagram?

(c) Since G must satisfy Rule 3, what constraints does this place on the Cayley diagram? Try to draw a diagram that is almost a Cayley diagram but violates Rule 3.

(d) Since G must satisfy Rule 2, what does this imply about the Cayley diagram? Can you construct a diagram that is almost a Cayley diagram but violates Rule 2? To do this, you may need to violate another one of our rules.

(e) What property does Rule 4 force the Cayley diagram to have? Can you construct a diagram that is almost a Cayley diagram but violates Rule 4?

In the previous exercise, you discovered several properties embodied by all Cayley diagrams. Unfortunately, not every diagram having these properties will yield a Cayley diagram. For example, the diagram in Figure 3.5 satisfies the properties you discovered in Exercise 3.17, but it turns out that this cannot be a diagram for any group (regardless of how we label the vertices).

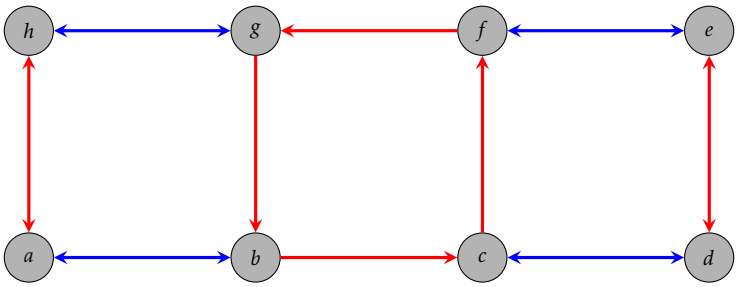


Figure 3.5

This fact exposes one of the weaknesses of our intuitive definition of a group and is one of the many reasons we will soon require a more rigorous definition.

Chapter 4

An Introduction to Subgroups and Isomorphisms

In this chapter, we'll continue to utilize our intuitive definition of a group. That is, a group G is a set of actions that satisfies the following rules.

Rule 1. There is a predefined list of actions that never changes.

Rule 2. Every action is reversible.

Rule 3. Every action is deterministic.

Rule 4. Any sequence of consecutive actions is also an action.

In the previous chapter, we constructed lots of Cayley diagrams for various groups. To construct a Cayley diagram for a group G, we need to first identify a set of generators, say S. Recall that our choice of generators is important as changing the generators can result in a different Cayley diagram.

In the Cayley diagram for G using S, all the actions of G are represented by the vertices of the graph. Each vertex corresponds to a unique action. This does not imply that there is a unique way to obtain a given action from the generators. Each of the generators determines an arrow type in the diagram. One way to distinguish the different arrow types is by using different colors. An arrow of a particular color always represents the same generator.

One of the vertices in the diagram is labeled by the do-nothing action, often denoted by e. Each of the other vertices are labeled by words that correspond to following arrows (forwards or backwards) from e to a given vertex. There may be many ways to do this as each sequence of arrows corresponds to a unique word. So, a vertex could be potentially labeled by many words. Also, one potentially confusing item is that we read our words from right to left. That is, the first arrow we follow out of e is the rightmost generator in the word.

4.1 Subgroups

Exercise 4.1. Recall the definition of "subset." What do you think "subgroup" means? Try to come up with a potential definition. Try not to read any further before doing this.

Before continuing, gather up the following Cayley diagrams:

- $\text{Spin}_{1\times 2}$. There are 3 of these. I drew one for you in Chapter 3 and you discovered two more in Exercise 3.6.
- S_2. See Exercise 3.9.
- R_4. See Exercise 3.10.
- D_3. There are two of these. See exercises 3.11 and 3.14.
- D_4. See Exercise 3.12.

Exercise 4.2. Examine your Cayley diagrams for D_4 and R_4 and make some observations. How are they similar and how are they different? Can you reconcile the similarities and differences by thinking about the actions of each group?

Hopefully, one of the things you noticed in the previous exercise is that we can "see" R_4 inside of D_4 (and hopefully you didn't just read that before completing the exercise). You may have used different colors in each case and maybe even labeled the vertices with different words, but the overall structure of R_4 is there nonetheless.

Exercise 4.3. If you just pay attention to the configuration of arrows, it appears that there are two copies of the Cayley diagram for R_4 in the Cayley diagram for D_4. Isolate these two copies by ignoring the edges that correspond to the generator s. Paying close attention to the words that label the vertices from the original Cayley diagram for D_4, are either of these groups in their own right?

Recall that the do-nothing action must always be one of the actions included in a group. If this didn't occur to you when doing the previous exercise, you might want to go back and rethink your answer. Just like in the previous exercise, we can often "see" smaller groups living inside larger groups. These smaller groups are called **subgroups**.

Intuitive Definition 4.4. Let G be a group of actions and let $H \subseteq G$. We say that H is a **subgroup** if and only if H is a group in its own right. In this case, we write $H \leq G$.

In light of Exercise 4.3, we would write $R_4 \leq D_4$. The second sub-diagram of D_4 that resembles R_4 cannot be a subgroup because it does not contain the do-nothing action. However, since it looks a lot like R_4, we call it a **clone** of R_4. For convenience, we may also say that a subgroup is a clone of itself.

The next theorem* tells us that if we already have a subset of a group, we only need to check two of our rules instead of four.

*Perhaps we should call this an "Intuitive Theorem" since we are using an intuitive definition of a group.

Exercise 4.5. Let G be a group of actions and let $H \subseteq G$. If we wanted to determine whether H is a subgroup of G or not, can we skip checking any of the four rules? Which rules must we verify?

There are a couple subgroups that every group has.

Theorem 4.6. Let G be a group of actions and suppose that e is the do-nothing action. Then $\{e\} \leq G$.

Exercise 4.7. Let G be a group and suppose that e is the do-nothing action. What does the Cayley diagram for the subgroup $\{e\}$ look like?

Earlier, we referred to subgroups as being "smaller." However, our definition does not imply that this has to be the case.

Theorem 4.8. Let G be a group of actions. Then $G \leq G$.

We refer to subgroups that are strictly smaller than the whole group as **proper subgroups**.

Lots of groups have been given formal names (e.g., D_4, R_4, etc.). However, not every group or subgroup has a name. In this case, it's useful to have notation to refer to specific subgroups.

Definition 4.9. Let G be a group of actions and let $g_1, \ldots, g_n$ be distinct actions from G. We define $\langle g_1, \ldots, g_n \rangle$ to be the smallest subgroup containing $g_1, \ldots, g_n$. In this case, we call $\langle g_1, \ldots, g_n \rangle$ the **subgroup generated by** $g_1, \ldots, g_n$.

For example, consider $r, s, s' \in D_3$ (as defined in exercises 3.11 and 3.14). Then $\langle r, s \rangle = \langle s, s' \rangle = D_3$. Recall that R_4 is the subgroup of D_4 consisting of rotations of the square. Similarly, the group of rotations of an equilateral triangle is called R_3. Then using the r from D_3, we have $\langle r \rangle = R_3$, which is a subgroup of D_3.

Note that in Definition 4.9, we used a finite number of generators. There's no reason we have to do this. That is, we can consider groups/subgroups generated by infinitely many elements.

Exercise 4.10. Suppose $\{g_1, \ldots, g_n\}$ is a generating set for a group G.

(a) Explain why $\{g_1^{-1}, \ldots, g_n^{-1}\}$ is also a generating set for G.

(b) How does the Cayley diagram for G with generating set $\{g_1, \ldots, g_n\}$ compare to the Cayley diagram with generating set $\{g_1^{-1}, \ldots, g_n^{-1}\}$?

Exercise 4.11. Consider $\text{Spin}_{1\times 2}$.

(a) Can you find the Cayley diagram for $\langle t_1 \rangle$ as a subgroup of $\text{Spin}_{1\times 2}$?

(b) Write down all the actions of the subgroup $\langle t_1, t_2 \rangle$. Write them as words in t_1 and t_2. Can you find the Cayley diagram for $\langle t_1, t_2 \rangle$ as a subgroup of $\text{Spin}_{1\times 2}$? Can you find a clone for $\langle t_1, t_2 \rangle$?

One of the benefits of Cayley diagrams is that they are useful for visualizing subgroups. However, recall that if we change our set of generators, we might get a very different looking Cayley diagram. The upshot of this is that we may be able to see a subgroup in one Cayley diagram for a given group, but not be able to see it in a Cayley diagram with a different set of arrows.

Exercise 4.12. We currently have two different Cayley diagrams for D_3 (see Exercises 3.11 and 3.14).

(a) Can you find the Cayley diagram for $\langle e \rangle$ as a subgroup of D_3? Can you see it in both Cayley diagrams for D_3? Can you find all the clones?

(b) Can you find the Cayley diagram for $\langle r \rangle = R_3$ as a subgroup of D_3? Can you see it in both Cayley diagrams? Can you find all the clones?

(c) Find the Cayley diagrams for $\langle s \rangle$ and $\langle s' \rangle$ as subgroups of D_3. Can you see them in both Cayley diagrams for D_3? Can you find all the clones?

Exercise 4.13. Consider D_4. Let h be the action that reflects (i.e., flips over) the square over the horizontal midline and let v be the action that reflects the square over the vertical midline. Also, recall that r^2 is shorthand for the action rr that does r twice in a row. Which of the following are subgroups of D_4? In each case, justify your answer. If a subset is a subgroup, try to find a minimal set of generators. Also, determine whether you can see the subgroups in our Cayley diagram for D_4.

(a) $\{e, r^2\}$

(b) $\{e, h\}$

(c) $\{e, h, v\}$

(d) $\{e, h, v, r^2\}$

The subgroup in Exercise 4.13(d) is often referred to as the **Klein four-group** and is denoted by V_4.

Exercise 4.14. Draw the Cayley diagram for V_4 using $\{v, h\}$ as our set of generators.

Let's introduce a group we haven't seen yet. We define the **quaternion group** to be the group $Q_8 = \{1, -1, i, -i, j, -j, k, -k\}$ having the Cayley diagram with generators $i, j, -1$ given in Figure 4.1. In this case, 1 is the do-nothing action.

Notice that I didn't mention what the actions actually do. For now, let's not worry about that. The relationship between the arrows and vertices tells us everything we need to know. Also, let's take it for granted that Q_8 actually is a group.

Exercise 4.15. Consider the Cayley diagram for Q_8 given in Figure 4.1.

(a) Which arrows correspond to which generators in our Cayley diagram for Q_8?

(b) What is i^2 equal to? That is, what element of $\{1, -1, i, -i, j, -j, k, -k\}$ is i^2 equal to? How about i^3, i^4, and i^5?

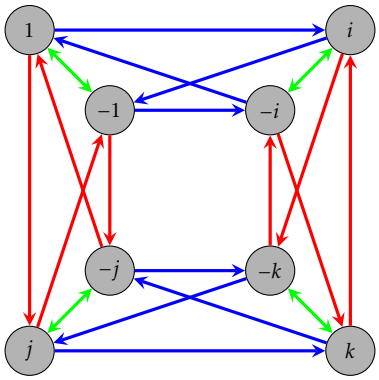


Figure 4.1. Cayley diagram for Q_8 with generating set $\{-1, i, j\}$.

Figure 4.2. Cayley diagram for S_2 with generator s.

(c) What are j^2, j^3, j^4, and j^5 equal to?

(d) What is $(-1)^2$ equal to?

(e) What is ij equal to? How about ji?

(f) Can you determine what k^2 and ik are equal to?

(g) Can you identify a generating set consisting of only two elements? Can you find more than one?

(h) What subgroups of Q_8 can you see in the Cayley diagram in Figure 4.1?

(i) Find a subgroup of Q_8 that you cannot see in the Cayley diagram.

4.2 Isomorphisms

By now you've probably seen enough examples of Cayley diagrams to witness some patterns appearing over and over again. One of the things you've probably noticed is that parts of some Cayley diagrams look just like parts of other Cayley diagrams.

Recall from Exercise 3.9 that S_2 is the group that acts on two coins by swapping their positions (but not flipping them over). We defined s to be the action that swaps the left and right coins and as usual e is the do-nothing action. The Cayley diagram for S_2 with generator s is given in Figure 4.2.

If you look back at all the Cayley diagrams you've encountered, you'll notice that many of them contained chunks that resemble the Cayley diagram for S_2 with generator s. In particular, in the Cayley diagrams for $\text{Spin}_{1\times 2}$, D_3, D_4, and Q_8 that we've seen, it is easy to identify the portions that "look like" S_2. For example, if you isolate the Cayley diagram for the subgroup $\langle -1 \rangle = \{1, -1\}$ in Q_8, we see that it looks just like the

Cayley diagram for S_2, except the labels are not identical. The clones of the subgroup $\langle -1 \rangle = \{1, -1\}$ in Q_8 look like S_2, as well, but they do not contain the do-nothing action.

The one thing that is different about the Cayley diagram for S_2 and the Cayley diagram for $\langle -1 \rangle$ is that the labels are different. If we set the Cayley diagram for S_2 on top of the Cayley diagram for $\langle -1 \rangle$ such that the do-nothing actions match up, then s and -1 would correspond to each other. In other words, the two Cayley diagrams are identical up to relabeling the vertices.

In this case, we say that S_2 and the subgroup $\langle -1 \rangle$ of Q_8 are **isomorphic** under the correspondence $e \leftrightarrow 1$ and $s \leftrightarrow -1$. This one-to-one correspondence between the two groups is called an **isomorphism**, which is depicted in Figure 4.3. Note that I've recolored the arrow in S_2 so that it matches the corresponding arrow color of $\langle -1 \rangle$. This isn't necessary, but it makes the correspondence more obvious.

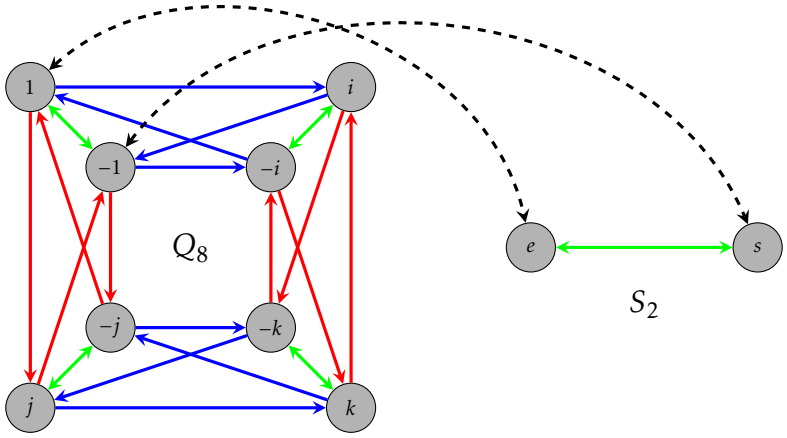


Figure 4.3. Isomorphism between $\langle -1 \rangle \leq Q_8$ and S_2.

What this means is that these two groups have the same structure and characteristics. Or, in other words, these two groups essentially do the "same kind" of thing. Clearly, the two do-nothing actions behave the same way. Also, s and -1 both have the property that doing the action twice results in having done nothing (i.e., each element is its own reverse). Since there are only two elements, there isn't anything else to check. In groups with more elements, things can get much more complicated.

It is important to point out that S_2 and $\langle -1 \rangle$ (in Q_8) are not equal. But they have the same structure. Identifying when two groups have the same structure (i.e., isomorphic) is an important pursuit in group theory.

If you look at the original Cayley diagram for $\text{Spin}_{1\times 2}$ (with generators s, t_1, t_2), we can see three subgroups that look like S_2; namely $\langle s \rangle$, $\langle t_1 \rangle$, and $\langle t_2 \rangle$. Each of these three subgroups is isomorphic to S_2.

There is one serious potential for confusion here. Notice that there is an s in S_2 and an s in $\text{Spin}_{1\times 2}$. Despite having identical names, they are not the same element. Since we only have 26 letters in our alphabet this sort of thing is unavoidable. Under the isomorphism between S_2 and the subgroup $\langle s \rangle$ in $\text{Spin}_{1\times 2}$, the two elements with the same name match up. That is, these two elements are the ones in each group with the same behavior.

Exercise 4.16. Can you find any other subgroups or groups that are isomorphic to S_2?

Let's write down an official definition of isomorphic.

Definition 4.17. Let G and G' be two groups. We say that G and G' are **isomorphic** if there exist generating sets S and S' for G and G', respectively, such that the corresponding Cayley diagrams are identical where we ignore the labels on the vertices and recolor the edges if necessary. In this case, we write $G \cong G'$. Otherwise, we say that G and G' are not isomorphic. If G and G' are isomorphic, then the one-to-one correspondence determined by matching up the corresponding generators and respecting arrow paths is called an **isomorphism**.

The last sentence in the definition above might be a bit much to handle at the moment, but as we construct more examples, the concept should become clear. The general idea is to take two identical Cayley diagrams (ignoring labels) for G and G' and then set one on top of the other so that the vertices and arrows of the same color match up. This should be done so that the do-nothing actions correspond to each other. Then it becomes clear which actions in G correspond to which actions in G'. There might be many ways to do this.

Consider the group R_4 with generator r (rotation by 90° clockwise). Now, take a look at the Cayley diagram for Q_8 with generators $i, j, -1$. It should be easy to convince yourself that R_4 is isomorphic to both $\langle i \rangle = \{1, i, -i, -1\}$ and $\langle j \rangle = \{1, j, -j, -1\}$. However, you have to do some rearranging of one of the diagrams to set one on top of the other. Let's just focus on $\langle i \rangle$.

How do R_4 and $\langle i \rangle$ match up? We want to pair elements in each group with an element in the other group that has the same behavior. Clearly, e and 1 match up since these are the two do-nothing actions. Also, the reason why we noticed these two groups were isomorphic is because their Cayley diagrams looked the same. Since each Cayley diagram only had one arrow type determined by r and i, we should pair these two elements. Now, following the arrows around the diagram, we see that r^2 must pair with $i^2 = -1$ and r^3 corresponds to $i^3 = -i$. In summary, the isomorphism between R_4 and $\langle i \rangle$ (in Q_8) is given by $e \leftrightarrow 1$, $r \leftrightarrow i$, $r^2 \leftrightarrow -1$, and $r^3 \leftrightarrow -i$, which is depicted in Figure 4.4. Note that this time we have not recolored the edges so that they match. Nonetheless, the correspondence should be clear.

Now, take a look at the Cayley diagram for D_4 with generating set $\{r, s\}$. As we noticed in Exercise 4.2, R_4 is a subgroup of D_4. We could say that this subgroup is isomorphic to R_4, but in this case, we can say something even stronger: they are equal!

Before continuing, we need to emphasize an important point. If the Cayley diagram for one group does not look like the Cayley diagram for another group, then that does *not* immediately imply that the groups are not isomorphic. The issue is that perhaps we could choose appropriate generating sets for each group so that the Cayley diagrams do look alike. For example, notice that our standard Cayley diagram for R_4 does not look like the Cayley diagram that you constructed for V_4 in Exercise 4.14. This does *not* imply that these two groups are not isomorphic. We would need to do some more work in order to determine whether the two groups are isomorphic or not. You will get a chance to do this in Exercise 4.21.

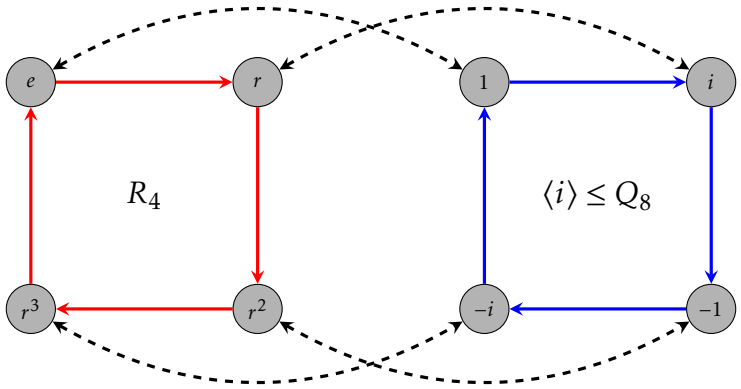


Figure 4.4. Isomorphism between $\langle i \rangle \leq Q_8$ and R_4.

It turns out that there is a fancy word for the size of a group.

Definition 4.18. If G is a group with n distinct actions, then we say that G has **order** n and write $|G| = n$. If G contains infinitely many elements, then we say G has infinite order and write $|G| = \infty$.

Exercise 4.19. Find the orders of the following groups: S_2, $\text{Spin}_{1\times 2}$, $\text{Spin}_{3\times 3}$, R_4, D_3, D_4, V_4, and Q_8.

Theorem 4.20. Suppose G and G' are two groups of actions such that $G \cong G'$. Then $|G| = |G'|$.

Unfortunately, the converse of the previous theorem is not true in general. That is, two groups that have the same order may or may not be isomorphic.

Loosely speaking, if one group has a property that the other does not have, then the two groups cannot be isomorphic. For example, if one group has the property that every pair of actions commutes (i.e., the order[†] of the actions does not matter), but another group has a pair of actions that do not commute, then the two groups cannot be isomorphic. Moreover, if one group contains an action that requires a minimum of k applications to get back to the do-nothing action, but a second group does not have such an element, then the two groups cannot be isomorphic.

Justifying these two claims takes a bit of work and for now, we'll put that on hold. For the time being, if you don't see why these claims about when two groups are not isomorphic are true, just take them on faith and we will return to the issue in a later chapter. Feel free to use these ideas in the exercises that follow.

Problem 4.21. Determine whether R_4 and V_4 are isomorphic. Justify your answer. If they are isomorphic, specify the isomorphism by listing the correspondence of elements. If they are not isomorphic, explain why.

Problem 4.22. Consider the group given by the Cayley diagram for R_6 that was given in Exercise 3.15. We can think of R_6 as the rotation group for a regular hexagon. Determine

[†]Don't confuse the word "order" in this sentence with the order of a group.

whether R_6 and D_3 are isomorphic. Justify your answer. If they are isomorphic, specify the isomorphism by listing the correspondence of elements. If they are not isomorphic, explain why.

Exercise 4.23. Consider two light switches on a wall side by side. Consider the group of actions that consists of all possible actions that you can do to the two light switches. For example, one action is toggle the left light switch while leaving the right alone. Let's call this group L_2.

(a) How many distinct actions does L_2 have?

(b) Can you find a minimal generating set for L_2? If so, give these actions names and then write all of the actions of L_2 as words in your generator(s).

(c) Using your generators from part (b), draw a Cayley diagram for L_2.

Problem 4.24. Determine whether L_2 and V_4 are isomorphic. Justify your answer. If they are isomorphic, specify the isomorphism by listing the correspondence of elements. If they are not isomorphic, explain why.

Problem 4.25. Determine whether Q_8 and D_4 are isomorphic. Justify your answer. If they are isomorphic, specify the isomorphism by listing the correspondence of elements. If they are not isomorphic, explain why.

Problem 4.26. Determine whether $\mathrm{Spin}_{1\times 2}$ and D_4 are isomorphic. Justify your answer. If they are isomorphic, specify the isomorphism by listing the correspondence of elements. If they are not isomorphic, explain why.

Exercise 4.27. Consider the group that acts on three coins that are in a row by rearranging their positions (but not flipping them over). This group is called S_3. Number the positions of the coins (not the coins themselves) 1, 2, 3 from left to right. Let s_1 be the action that swaps the coins in positions 1 and 2 and let s_2 be the action that swaps the coins in positions 2 and 3.

(a) The group S_3 consists of 6 actions, which we can generate with s_1 and s_2. Write all 6 actions as words in s_1 and s_2.

(b) Using s_1 and s_2 as generators, draw a Cayley diagram for S_3.

Problem 4.28. Determine whether S_3 and D_3 are isomorphic. Justify your answer. If they are isomorphic, specify the isomorphism by listing the correspondence of elements. If they are not isomorphic, explain why. Don't forget that we've drawn two different Cayley diagrams for D_3.

Chapter 5

A Formal Approach to Groups

In this chapter we finally introduce the formal definition of a group. From this point on, our focus will shift from developing intuition to studying the abstract properties of groups. However, we should not abandon the intuition we have gained. As we progress, your intuitive understanding of groups will continue to improve and you should rely on this understanding as you try to make sense of the notions that follow. There has been plenty of intentional foreshadowing, so expect to revisit concepts you've already encountered. We'll also encounter plenty of new stuff, too.

It is important to point out that things are about to get quite a bit more difficult for most of you. Be patient and persistent!

5.1 Binary Operations

After learning to count as a child, you likely learned how to add, subtract, multiply, and divide with natural numbers. Loosely speaking, these operations are examples of binary operations since we are combining two objects to obtain a single object. More formally, we have the following definition.

Definition 5.1. A **binary operation** $*$ on a set A is a function from $A \times A$ into A. For each $(a, b) \in A \times A$, we denote the element $*(a, b)$ via $a * b$.

Remark 5.2. Don't misunderstand the use of $*$ in this context. We are not implying that $*$ is the ordinary multiplication of real numbers that you are familiar with. We use $*$ to represent a generic binary operation.

Remark 5.3. Notice that since the codomain of a binary operation on a set A is A, binary operations require that we yield an element of A when combining two elements of A. In this case, we say that A is **closed** under $*$. Binary operations have this closure property by definition. Also, since binary operations are functions, any attempt to combine two elements from A should result in a *unique* element of A. In this case, we say that $*$ is **well-defined**. Moreover, since the domain of $*$ is $A \times A$, it must be the case that $*$ is defined for *all* pairs of elements from A.

Example 5.4. Examples of binary operations include + (addition), − (subtraction), and · (multiplication) on the real numbers. However, ÷ (division) is not a binary operation on the set of real numbers because all elements of the form $(a, 0)$ are not in the domain $\mathbb{R} \times \mathbb{R}$ since we cannot divide by 0. Yet, ÷ is a suitable binary operation on $\mathbb{R} \setminus \{0\}$.

Example 5.5. Let C be the set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$. Then $\circ$ (function composition) is a binary operation on C.

Example 5.6. Consider the 6 actions of D_3. The composition of these actions is a binary operation on D_3. In fact, composition of actions for each of the groups that we have seen is a binary operation on the given group. Notice that we never used a symbol for these binary operations, but rather used juxtaposition (i.e., ab is the juxtaposition of a and b).

Example 5.7. Let $M_{2\times 2}(\mathbb{R})$ be the set of 2×2 matrices with real number entries. Then matrix multiplication is a binary operation on $M_{2\times 2}(\mathbb{R})$.

Exercise 5.8. Explain why composition of spins is not a binary operation on the set of allowable spins in $\text{Spin}_{3\times 3}$. *Hint:* Reread the paragraph below Exercise 2.8.

Exercise 5.9. Let $M(\mathbb{R})$ be the set of matrices (of any size) with real number entries. Is matrix addition a binary operation on $M(\mathbb{R})$? How about matrix multiplication? What if you restrict to square matrices of a fixed size $n \times n$?

Exercise 5.10. Determine whether $\cup$ (union) and $\cap$ (intersection) are binary operations on $\mathcal{P}(\mathbb{Z})$ (i.e., the power set of the integers).

Exercise 5.11. Consider the closed interval $[0, 1]$ and define $*$ on $[0, 1]$ via $a * b = \min\{a, b\}$ (i.e., take the minimum of a and b). Determine whether $*$ is a binary operation on $[0, 1]$.

Some binary operations have additional properties.

Definition 5.12. Let A be a set and let $*$ be a binary operation on A.

(a) We say that $*$ is **associative** if and only if $(a * b) * c = a * (b * c)$ for all $a, b, c \in A$.

(b) We say that $*$ is **commutative** if and only if $a * b = b * a$ for all $a, b \in A$.

Exercise 5.13. Provide at least one example of a binary operation on a set that is commutative. How about not commutative?

Theorem 5.14. Let A be a set and let F be the set of functions from A to A. Then function composition is an associative binary operation on F.

When the set A is finite, we can represent a binary operation on A using a table in which the elements of the set are listed across the top and the left side (in the same order). The entry in the ith row and jth column of the table represents the output of combining the element that labels the ith row with the element that labels the jth column (order matters).

Example 5.15. Consider the following table.

$*$	a	b	c
a	b	c	b
b	a	c	b
c	c	b	a

This table represents a binary operation on the set $A = \{a, b, c\}$. In this case, $a * b = c$ while $b * a = a$. This shows that $*$ is not commutative.

Exercise 5.16. What property must a table for a binary operation have in order for the operation to be commutative?

Exercise 5.17. Fill in the missing entries in the following table so that $*$ defines an associative binary operation on $\{a, b, c, d\}$.

$*$	a	b	c	d
a	a	b	c	d
b	b	a	c	d
c	c	d	c	d
d				

5.2 Groups

Without further ado, here is our official definition of a group.

Definition 5.18. A group $(G, *)$ is a set G together with a binary operation $*$ such that the following axioms hold.

(0) The set G is closed under $*$.

(1) The operation $*$ is associative.

(2) There is an element $e \in G$ such that for all $g \in G$, $e * g = g * e = g$. We call e the **identity**.

(3) Corresponding to each $g \in G$, there is an element $g' \in G$ such that $g * g' = g' * g = e$. In this case, g' is called the **inverse** of g, which we shall denote as g^{-1}.

Remark 5.19. A few comments are in order.

(a) Notice that a group has two parts to it, namely, a set and a binary operation. For simplicity, if $(G, *)$ is a group, we will often refer to G as being the group. However, you must remember that the binary operation is part of the structure.

(b) Axiom 2 forces G to be nonempty.

(c) In the generic case, even if $*$ is not actually multiplication, we will refer to $a * b$ as the **product** of a and b.

(d) We are not requiring $*$ to be commutative. If $*$ is commutative, then we say that G is **abelian**[*] (or **commutative**).

Exercise 5.20. Explain why Axiom 0 is unnecessary.

At this time, we have two definitions of a group. The first one was intended to provide an intuitive introduction and Definition 5.18 provides a rigorous mathematical definition. We should confirm that these two definitions are in fact compatible.

Exercise 5.21. Compare and contrast our two definitions of a group. How do the rules and axioms match up?

Exercise 5.22. Quickly verify that $\mathrm{Spin}_{1\times 2}$, S_2, R_4, D_3, D_4, V_4, and Q_8 are groups under composition of actions.

Exercise 5.23. Determine whether each of the following are groups. If the pair is a group, determine whether it is abelian and identify the identity. Explain your answers.

(a) $(\mathbb{Z}, +)$

(b) $(\mathbb{N}, +)$

(c) $(\mathbb{Z}, \cdot)$

(d) $(\mathbb{R}, +)$

(e) $(\mathbb{R}, \cdot)$

(f) $(\mathbb{R} \setminus \{0\}, \cdot)$

(g) $(M_{2\times 2}(\mathbb{R}), +)$

(h) $(M_{2\times 2}(\mathbb{R}), *)$, where $*$ is matrix multiplication.

(i) $(\{a, b, c\}, *)$, where $*$ is the operation determined by the table in Example 5.15.

(j) $(\{a, b, c, d\}, *)$, where $*$ is the operation determined by the table in Exercise 5.17.

Notice that in Axiom 2 of Definition 5.18, we said *the* identity and not *an* identity. Implicitly, this implies that the identity is unique.

Theorem 5.24. Let G be a group with binary operation $*$. Then there is a unique identity element in G. That is, there is only one element e in G such that $g * e = e * g = g$ for all $g \in G$.

The following theorem is crucial for proving many theorems about groups.

Theorem 5.25 (Cancellation Law)**.** Let $(G, *)$ be a group and let $g, x, y \in G$. Then $g * x = g * y$ if and only if $x = y$. Similarly, $x * g = y * g$ if and only if $x = y$.[†]

[*]Commutative groups are called abelian in honor of the Norwegian mathematician Niels Abel (1802–1829).
[†]You only need to prove one of these statements as the proof of the other is symmetric.

Exercise 5.26. Show that $(\mathbb{R},\cdot)$ fails the Cancellation Law (confirming the fact that it is not a group).

Corollary 5.27. Let G be a group with binary operation $*$. Then each $g \in G$ has a unique inverse.

Theorem 5.28. Let $(G,*)$ be a group and let $g,h \in G$. Then the equations $g * x = h$ and $y * g = h$ have unique solutions for x and y in G.

While proving the previous few theorems, hopefully one of the things you realized is that you can multiply both sides of a group equation by the same element but that you have to do it on the same side of each half. That is, since a group may or may not be abelian, if I multiply one side of an equation on the left by a group element, then we must multiply the other side of the equation on the left by the same group element.

Despite the fact that a group may or may not be abelian, if one product is equal to the identity, then reversing the order yields the same result.

Theorem 5.29. Let G be a group with binary operation $*$. If $g * h = e$, then $h * g = e$.

The upshot of the previous theorem is if we have a "left inverse" then we automatically have a "right inverse" (and vice versa).

The next theorem should not be surprising.

Theorem 5.30. Let $(G,*)$ be a group and let $g \in G$. Then $(g^{-1})^{-1} = g$.

Definition 5.31. Let $(G,*)$ be a group and let $g \in G$. Then for $n \in \mathbb{N}$, we define

$$g^n = \underbrace{g * g * \cdots * g}_{n \text{ factors}}$$

and

$$g^{-n} = \underbrace{g^{-1} * g^{-1} * \cdots * g^{-1}}_{n \text{ factors}}.$$

Moreover, we define $g^0 = e$.

The good news is that the rules of exponents you are familiar with still hold for groups.

Theorem 5.32. Let $(G,*)$ be a group and let $g \in G$. For $n,m \in \mathbb{Z}$, we have the following:

(a) $g^n * g^m = g^{n+m}$,

(b) $(g^n)^{-1} = g^{-n}$.

5.3 Group Tables

Recall that we could represent a binary operation on a finite set using a table. Since groups have binary operations at their core, we can represent a finite group (i.e., a group with finitely many elements) using a table, called a **group table** (or **Cayley table**). For example, below are group tables for D_3 and V_4, respectively.

$*$	e	r	r^2	s	sr	sr^2
e	e	r	r^2	s	sr	sr^2
r	r	r^2	e	sr^2	s	sr
r^2	r^2	e	r	sr	sr^2	s
s	s	sr	sr^2	e	r	r^2
sr	sr	sr^2	s	r^2	e	r
sr^2	sr^2	s	sr	r	r^2	e

$*$	e	v	h	vh
e	e	v	h	vh
v	v	e	vh	h
h	h	vh	e	v
vh	vh	h	v	e

Our convention will be that if x appears in row i and y appears in column j, then row i "times" column j will result in the element determined by xy, where as usual we follow our right to left convention. That is, xy means we apply y first and then x (as in function composition).

Exercise 5.33. Verify that V_4 is an abelian group. What feature of the table makes this clear?

Given an arbitrary group G, we should probably say, "a group table for G" and not "the group table for G." The reason for this is that if we chose a different order of the elements (e.g., swap rows 1 and 4—which swaps columns 1 and 4, as well), then the table would look slightly different. Also, if we had chosen a different generating set, then the names of the elements would look different. Regardless, the table still captures the same information about the binary operation. Because every possible table for a given group conveys the same information about the architecture of the group, people may refer to any table for the group as "the" table.

Exercise 5.34. Create group tables for the following groups: S_2, R_3, R_4, D_3, S_3, D_4, and Q_8. Which groups are abelian?

Perhaps you noticed when creating the tables above that each element of the group appeared exactly once in each row and column, respectively. This is true, in general. Use Theorem 5.28, to prove the following theorem.

Theorem 5.35. Let $(G,*)$ be a finite group. Then each element of G appears exactly once in each row and each column, respectively, in any group table for G.

We can also use tables to define groups. For example, consider the following table on the set $A = \{e, a, b, c\}$.

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Is this a table for a group? First, we see that the binary operation determined by the table is closed. Second, we see that e is acting as the identity. Since every row and column has the identity element e appearing, we know that every element has an inverse (do you see why that follows?). The only thing left to check is associativity. Imagine for a moment what this entails. It's messy right?! And this is only for a group of order 4.

Thankfully, we can rely on some prior knowledge to help out with associativity. It turns out that if you look closely, the group table for V_4 looks the "same" as the table above. What do we mean by "same" here? The names for elements are different (except for e), but

the product of corresponding elements yields the corresponding result.

To see what I mean, let's color both tables with white, red, blue, and green in such a way that each element corresponds to a unique color. If we choose our colors wisely, it is easy to see that both tables have the same structure.

$*$	e	v	h	vh
e	e	v	h	vh
v	v	e	vh	h
h	h	vh	e	v
vh	vh	h	v	e

$\longleftrightarrow$

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Since we already know that V_4 is a group, we know that the binary operation for V_4 is associative.

Exercise 5.36. Explain why the discussion above implies that the binary operation determined by the table on the right above must be associative. Have we shown that $(A,*)$ is a group?

It is important to point out that if we had not chosen our colors wisely, then perhaps the colorings of the two tables would not agree. Moreover, if we had made the same color choices for elements, but then rearranged columns and rows of one table, the colorings of the two tables would not agree. This doesn't imply anything. The point is whether we *can* get the tables to match.

Exercise 5.37. Draw the Cayley diagram for $(A,*)$ with generators a and b. Explain why this implies that V_4 and A (under their respective binary operations) are isomorphic.

Exercise 5.38. Is it possible to color the group table for R_4 so that it matches the coloring of V_4? Explain your answer.

Problem 5.39. Let $(G,*)$ and $(G',\circ)$ be two finite groups. Suppose we can arrange the rows and columns and color elements in such a way that the colorings for the two group tables agree. Explain why this implies that the two groups are isomorphic.

Problem 5.40. Suppose we have a table for $(G,*)$, where G is finite. Further suppose that (i) there is an identity element, and (ii) every element appears exactly once in each row and column, respectively. Explain why the only thing we need to verify in order for $(G,*)$ to be a group is that $*$ is associative.

Problem 5.41. Suppose that $(G,*)$ is a group. Theorem 5.24 guarantees that there is a unique identity in G. When creating the group table for G, what goes wrong if you try to include two different identity elements?

Consider the class of all possible groups. It turns out that "isomorphic" ($\cong$) determines an equivalence relation. That is, under this relation two groups are related if and only if they are isomorphic. We'll prove this formally later when we have a more rigorous definition of isomorphic.

Problem 5.42. Explain why all groups with a single element are isomorphic.

In this case, we say that "up to isomorphism" there is only one group with a single element.

Problem 5.43. Consider a group $(G,*)$ of order 2. Suppose that $G = \{e, a\}$. Complete the following group table for G.

$*$	e	a
e		
a		

Explain why every group with 2 elements must be isomorphic to S_2.

The previous problem implies that up to isomorphism, there is only one group of order 2.

Problem 5.44. Consider a group $(G,*)$ of order 3. Suppose that $G = \{e, a, b\}$. Complete the following group table for G.

$*$	e	a	b
e			
a			
b			

Explain why every group with 3 elements must be isomorphic to R_3.

Problem 5.45. Consider a group $(G,*)$ of order 4. Suppose that $G = \{e, a, b, c\}$. Assuming that e is the identity, the first row and first column of the corresponding group table must be completed as follows.

$*$	e	a	b	c
e	e	a	b	c
a	a	?		
b	b			
c	c			

The cell with the question mark cannot be filled with an a. So, this entry must be either e, b, or c. However, it should be easy to see the cases with b and c are symmetric. Thus, there are two cases: (i) the entry with the question mark is filled with e, (ii) the entry with the question mark is (without loss of generality) filled with b. Complete the group table in each of these two cases. Recall that we've seen two non-isomorphic groups of order 2, namely R_4 and V_4. What conclusion can you make about groups of order 4?

So far we've seen that there are unique groups up to isomorphism of orders 1, 2, and 3, but that there are two groups up to isomorphism of order 4. A general question we will want to address is, how many groups are there of order n?

In a future chapter we will be able to prove that there is only one group up to isomorphism of order 5, namely those groups isomorphic to R_5 (i.e., rotation group of a regular pentagon).

We've seen three groups of order 6, namely R_6, D_3, and S_3. However, $D_3 \cong S_3$ (see Problem 4.28) while R_6 is not isomorphic to either of these (see Problem 4.22). So, we can conclude that there are at least two groups up to isomorphism of order 6. But are there others? It turns out that the answer is yes, but why?

The group R_7 is the group of rotations of a regular 7-sided polygon. This group has order 7. Are there other groups of order 7 that are not isomorphic to R_7?

We've encountered four groups of order 8, namely D_4, $\mathrm{Spin}_{1\times 2}$, Q_8, and R_8. Of these, only D_4 and $\mathrm{Spin}_{1\times 2}$ are isomorphic. Thus, there are at least three groups up to isomorphism of order 8. However, are these the only ones? It turns out that the answer is no. What are the missing ones?

5.4 Revisiting Cayley Diagrams and Our Original Definition of a Group

Let's begin with a couple of exercises.

Exercise 5.46. Consider the diagram given in Figure 5.1. This is identical to the diagram that appeared in Figure 3.5 that we saw at the end of Chapter 3.

(a) Consider Rules 1–4 of our original definition of a group (see Definition 2.14). Does the diagram in Figure 5.1 satisfy Rules 1–4?

(b) Try to convert this diagram into a group table. Does the table represent a group? What goes wrong?

Exercise 5.47. Consider the diagram given in Figure 5.2

(a) Does the diagram in Figure 5.2 satisfy Rules 1–4 of Definition 2.14?

(b) Try to convert this diagram into a group table. Does the table represent a group? What goes wrong?

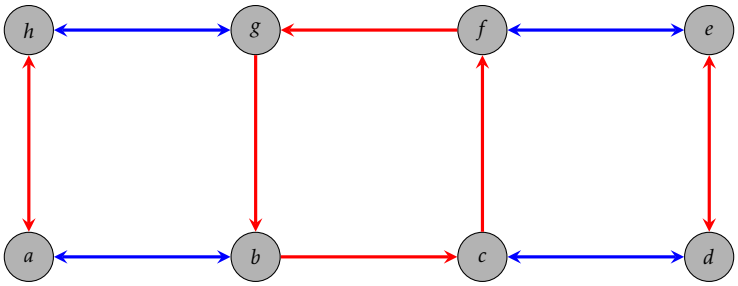


Figure 5.1. Diagram for Exercise 5.46.

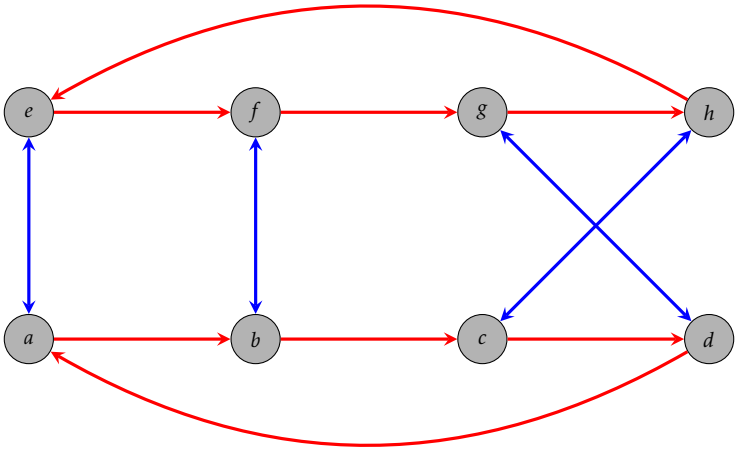


Figure 5.2. Diagram for Exercise 5.47.

As the previous two exercises indicate, the moral of the story is that our original intuitive definition of a group has a weakness. It does not agree with our formal definition of a group given in Definition 5.18. Let's see if we can figure out what goes wrong.

Consider the Cayley diagram for D_3 with generating set $\{r, s\}$ that is given in Figure 5.3. Notice that we labeled the lower right corner of the Cayley diagram with the word r^2s. This means that we first followed a blue arrow out of e and then two red arrows. However, we could also get to this vertex by first doing a red arrow out of e followed by a blue arrow. So, we could also have labeled this vertex with the word sr. The upshot is that $r^2s = sr$. These types of group equations are called **relations**.

We discovered this relation by starting at e and then traveling a sequence of arrows to get to the vertex in the lower right corner. However, notice that following a blue and then two red arrows is *always* the same as following a red arrow and then a blue arrow regardless of which vertex we start at. That is, the relation $r^2s = sr$ holds universally across the entire Cayley diagram.

Cayley diagrams for groups will always have this uniform symmetry. The fancy way of saying this is that Cayley diagrams are **regular**. In other words, a diagram is regular if every internal pattern repeats itself throughout the diagram.

Exercise 5.48. Identify two other relations that the Cayley diagram for D_3 given in Fig-

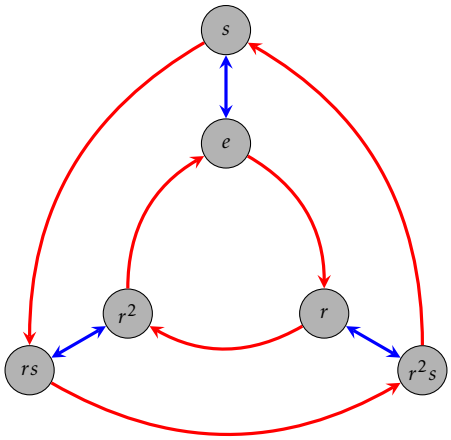


Figure 5.3. Cayley diagram for D_3 with generating set $\{r, s\}$.

ure 5.3 exhibits. Find one that involves both r^{-1} (i.e., following a red arrow backwards) and s. Convince yourself that your relations appear throughout the diagram.

Exercise 5.49. Verify that the diagrams in Exercise 5.46 and 5.47 are not regular.

Problem 5.50. Explain why the Cayley diagram for a group must be regular.

The discussion and exercises above lead us to conclude that one thing missing from our original intuitive definition of a group is regularity. It turns out that this is the only thing missing. That is, if we add the requirement of regularity to our intuitive definition, we could convert it into a rigorous definition that is equivalent to Definition 5.18. However, we won't bother doing this since our intuitive definition served its purpose.

We close this section with a problem that asks you to think about the structure of the Cayley diagrams for an abelian group.

Problem 5.51. Suppose $(G, *)$ is a group and suppose S is a generating set for G. Consider the Cayley diagram for G with generating set S.

(a) If G is abelian and $a, b \in S$, then what relationship must be true for the arrows in the Cayley diagram corresponding to the elements a and b?

(b) Is the converse of your claim true? That is, if every pair of edges in a Cayley diagram for G has the property you stated above, will the group be abelian?

5.5 Revisiting Subgroups

Back in Section 4.1, we introduced the notion of subgroup. In light of our official definition of a group, we more or less have the same definition as before, but let's restate it here using slightly more formal language.

Definition 5.52. Let $(G, *)$ be a group and let H be a subset of G. Then H is a **subgroup** of G, written $H \leq G$, provided that H is a group in its own right under the binary operation inherited from G.

The phrase "under the binary operation inherited from G" means that to combine two elements in H, we should treat the elements as if they were in G and perform G's binary operation.

Recall that Theorems 4.6 and 4.8 tell us that $\langle e \rangle = \{e\}$ and G are always subgroups of G. This is still true even using our official definition of a group. The subgroup $\{e\}$ is referred to as the **trivial subgroup**. All other subgroups are called **nontrivial**. If H is a subgroup of a group G with $H \neq G$, then we may write $H < G$ and refer to H as a **proper subgroup** of G.

The next theorem tells us that we don't need to verify all the axioms of a group to determine whether a nonempty subset is a subgroup.

Theorem 5.53. Suppose $(G, *)$ is a group and H is a nonempty subset of G. Then $H \leq G$ if and only if (i) for all $h \in H$, $h^{-1} \in H$, as well, and (ii) H is closed under the binary operation of G.

Remark 5.54. Notice that one of the hypotheses of Theorem 5.53 is that H be nonempty. This means that if we want to prove that a certain subset H is a subgroup of a group G, then one of the things we must do is verify that H is in fact nonempty.

Exercise 5.55. Consider $(\mathbb{R}^3, +)$, where $\mathbb{R}^3$ is the set of all 3-entry row vectors with real number entries (e.g., (a, b, c) where $a, b, c \in \mathbb{R}$) and $+$ is ordinary vector addition. It turns out that $(\mathbb{R}^3, +)$ is an abelian group with identity $(0, 0, 0)$.

(a) Let H be the subset of $\mathbb{R}^3$ consisting of vectors with first coordinate 0. Is H a subgroup of $\mathbb{R}^3$? Prove your answer.

(b) Let K be the subset of $\mathbb{R}^3$ consisting of vectors whose entries sum to 0. Is K a subgroup of $\mathbb{R}^3$? Prove your answer.

(c) Construct a subset of $\mathbb{R}^3$ (different from H and K) that is *not* a subgroup of $\mathbb{R}^3$.

Exercise 5.56. Consider the group $(\mathbb{Z}, +)$ (under ordinary addition).

(a) Show that the even integers, written $2\mathbb{Z} := \{2k \mid k \in \mathbb{Z}\}$, form a subgroup of $\mathbb{Z}$.

(b) Show that the odd integers are not a subgroup of $\mathbb{Z}$.

(c) Show that all subsets of the form $n\mathbb{Z} := \{nk \mid k \in \mathbb{Z}\}$ for $n \in \mathbb{Z}$ are subgroups of $\mathbb{Z}$.

(d) Are there any other subgroups besides the ones listed in part (c)? Explain your answer.

Exercise 5.57. Consider the group of symmetries of a regular octagon. This group is denoted by D_8, where the operation is composition of actions. The group D_8 consists of 16 elements (8 rotations and 8 reflections). Let H be the subset consisting of the following clockwise rotations: 0°, 90°, 180°, and 270°. Determine whether H is a subgroup of D_8 and justify your answer.

Exercise 5.58. Consider the groups $(\mathbb{R},+)$ and $(\mathbb{R}\setminus\{0\},\cdot)$. Explain why $\mathbb{R}\setminus\{0\}$ is not a subgroup of $\mathbb{R}$ despite the fact that $\mathbb{R}\setminus\{0\}\subseteq\mathbb{R}$ and both are groups (under the respective binary operations).

Theorem 5.59. Suppose $(G,*)$ is a group and let $H,K\leq G$. Then $H\cap K\leq G$.

Problem 5.60. Can we replace intersection with union in the theorem above? If so, prove the corresponding theorem. If not, then provide a specific counterexample.

Theorem 5.61. Suppose $(G,*)$ is an abelian group and let $H\leq G$. Then H is an abelian subgroup.

Problem 5.62. Is the converse of the previous theorem true? If so, prove it. Otherwise, provide a counterexample.

Theorem 5.63. Suppose $(G,*)$ is a group. Define

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

(called the **center** of G). Then $Z(G)$ is an abelian subgroup of G.

Exercise 5.64. Find the center of each of the following groups.

(a) S_2

(b) V_4

(c) S_3

(d) D_3

(e) D_4

(f) R_4

(g) R_6

(h) $\mathrm{Spin}_{1\times 2}$

(i) Q_8

(j) $(\mathbb{Z},+)$

(k) $(\mathbb{R}\setminus\{0\},\cdot)$

The following definition formalizes Definition 4.9.

Definition 5.65. Let $(G,*)$ be a group and let S be a nonempty subset of G. Then we define $\langle S\rangle$ to be the set consisting of all possible (finite) products of elements from S and their inverses. The set $\langle S\rangle$ is called the **subgroup generated by** S. The elements of S are called **generators** of $\langle S\rangle$.

Note that S may be finite or infinite. Moreover, even if S is finite, $\langle S \rangle$ may be infinite. Also, it is important to point out that we are not putting any restrictions about efficiency on S in the definition above. That is, it is possible that some elements are included in S that are not necessary to generate all of the elements of $\langle S \rangle$.

In cases when we know what the elements of S actually are, then we will list them inside the angle brackets without the set braces. For example, if $S = \{a, b, c\}$, then we will write $\langle a, b, c \rangle$ instead of $\langle \{a, b, c\} \rangle$. In the special case when S equals a single element, say $S = \{a\}$, then

$$\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\},$$

which is called the **subgroup generated by** a.

The set $\langle S \rangle$ is called the "subgroup generated by S", so it better be a subgroup!

Theorem 5.66. Let $(G, *)$ be a group and let $S \subseteq G$, where $S \neq \emptyset$. Then $\langle S \rangle \leq G$. In particular, $\langle S \rangle$ is the smallest subgroup of G containing S.

Exercise 5.67. In Exercise 5.56 we introduced the notation $n\mathbb{Z}$. Write these subgroups in the "generated by" notation. That is, find a set S such that $\langle S \rangle = n\mathbb{Z}$. Can you find more than one way to do it?

Every subgroup can be written in the "generated by" form. That is, if H is a subgroup of a group G, then there always exists a subset S of G such that $\langle S \rangle = H$. In particular, $\langle H \rangle = H$.

Let's explore a couple of examples. First, consider the group R_4 (where the operation is composition of actions). What are the subgroups of R_4? Theorems 4.6 and 4.8 tell us that $\{e\}$ and R_4 itself are subgroups of R_4. Are there any others? Theorem 5.53 tells us that if we want to find other subgroups of R_4, we need to find nonempty subsets of R_4 that are closed and contain all the necessary inverses. However, the previous paragraph indicates that we can find all of the subgroups of R_4 by forming the subgroups generated by various combinations of elements from R_4. We can certainly be more efficient, but below we list all of the possible subgroups we can generate using subsets of R_4. We are assuming that r is rotation by 90° clockwise. As you scan the list, you should take a moment to convince yourself that the list is accurate.

$\langle e \rangle = \{e\}$

$\langle r \rangle = \{e, r, r^2, r^3\}$

$\langle r^2 \rangle = \{e, r^2\}$

$\langle r^3 \rangle = \{e, r^3, r^2, r\}$

$\langle e, r \rangle = \{e, r, r^2, r^3\}$

$\langle e, r^2 \rangle = \{e, r^2\}$

$\langle e, r^3 \rangle = \{e, r^3, r^2, r\}$

$\langle r, r^2 \rangle = \{e, r, r^2, r^3\}$

$\langle r, r^3 \rangle = \{e, r, r^2, r^3\}$

$\langle r^2, r^3 \rangle = \{e, r, r^2, r^3\}$

$\langle e, r, r^2 \rangle = \{e, r, r^2, r^3\}$

$\langle e, r, r^3 \rangle = \{e, r, r^2, r^3\}$

$\langle e, r^2, r^3 \rangle = \{e, r, r^2, r^3\}$

$\langle r, r^2, r^3 \rangle = \{e, r, r^2, r^3\}$

$\langle e, r, r^2, r^3 \rangle = \{e, r, r^2, r^3\}$

Let's make a few observations. Scanning the list, we see only three distinct subgroups: $\{e\}, \{e, r^2\}, \{e, r, r^2, r^3\}$. Our exhaustive search guarantees that these are the only subgroups of R_4. It is also worth pointing out that if a subset contains either r or r^3, then that subset generates all of R_4. The reason for this is that r and r^3 are each generators for R_4, respectively. Also, observe that if we increase the size of the subset using an element that was already contained in the subgroup generated by the smaller set, then we don't get anything new. For example, consider $\langle r^2 \rangle = \{e, r^2\}$. Since $e \in \langle r^2 \rangle$, we don't get anything new by including e in our generating set. We can state this as a general fact.

Theorem 5.68. Let $(G, *)$ be a group and let $g_1, g_2, \ldots, g_n \in G$. If $x \in \langle g_1, g_2, \ldots, g_n \rangle$, then $\langle g_1, g_2, \ldots, g_n \rangle = \langle g_1, g_2, \ldots, g_n, x \rangle$.

It is important to point out that in the theorem above, we are not saying that $\{g_1, g_2, \ldots, g_n\}$ is a generating set for G—although this may be the case. Instead, are simply making a statement about the subgroup $\langle g_1, g_2, \ldots, g_n \rangle$, whatever it may be.

Let's return to our example involving R_4. We have three subgroups, namely the two trivial subgroups $\{e\}$ and R_4 itself, together with one nontrivial subgroup $\{e, r^2\}$. Notice that $\{e\}$ is also a subgroup of $\{e, r^2\}$. We can capture the overall relationship between the subgroups using a **subgroup lattice**, which we depict in Figure 5.4 case of R_4.

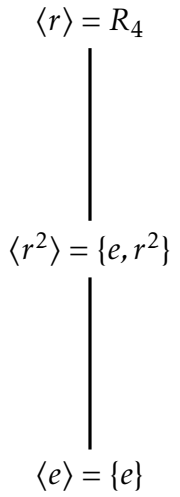


Figure 5.4. Subgroup lattice for R_4.

In general, subgroups of smaller order are towards the bottom of the lattice while larger subgroups are towards the top. Moreover, an edge between two subgroups means that the smaller set is a subgroup of the larger set.

Let's see what we can do with $V_4 = \{e, v, h, vh\}$. Using an exhaustive search, we find that there are five subgroups:

$\langle e \rangle = \{e\}$

$\langle h \rangle = \{e, h\}$

$\langle v \rangle = \{e, v\}$

$\langle vh \rangle = \{e, vh\}$

$\langle v, h \rangle = \langle v, vh \rangle = \langle h, vh \rangle = \{e, v, h, vh\} = V_4$

For each subgroup above, we've used minimal generating sets to determine the group. (Note that minimal generating sets are generating sets where we cannot remove any elements and still obtain the same group. Two minimal generating sets for the same group do not have to have the same number of generators.) In this case, we get the subgroup lattice in Figure 5.5.

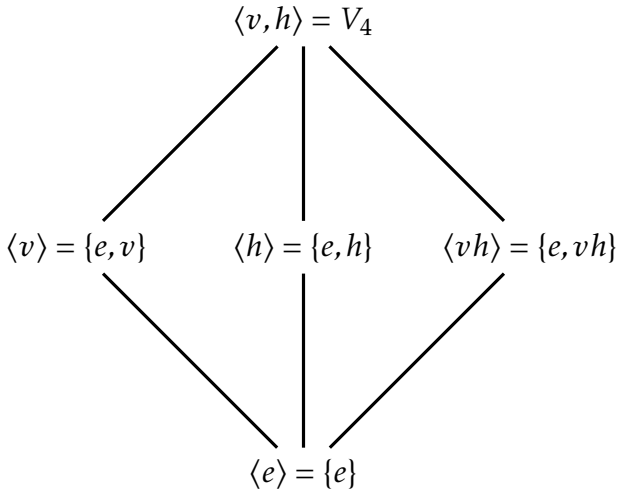


Figure 5.5. Subgroup lattice for V_4.

Notice that there are no edges among $\langle v \rangle, \langle h \rangle$, and $\langle vh \rangle$. The reason for this is that none of these groups are subgroups of each other. We already know that R_4 and V_4 are not isomorphic, but this becomes even more apparent if you compare their subgroup lattices.

In the next few exercises, you are asked to create subgroup lattices. As you do this, try to minimize the amount of work it takes to come up with all the subgroups. In particular, I do *not* recommend taking a full brute-force approach like we did for R_4.

Exercise 5.69. Find all the subgroups of $R_5 = \{e, r, r^2, r^3, r^4\}$ (where r is rotation clockwise of a regular pentagon by 72°) and then draw the subgroup lattice for R_5.

Exercise 5.70. Find all the subgroups of $R_6 = \{e, r, r^2, r^3, r^4, r^5\}$ (where r is rotation clockwise of a regular hexagon by 60°) and then draw the subgroup lattice for R_6.

Exercise 5.71. Find all the subgroups of $D_3 = \{e, r, r^2, s, sr, sr^2\}$ (where r and s are the usual actions) and then draw the subgroup lattice for D_3.

Exercise 5.72. Find all the subgroups of $S_3 = \langle s_1, s_2 \rangle$ (where s_1 is the action is that swaps the positions of the first and second coins and s_2 is the action that swaps the second and third coins; see Exercise 4.27) and then draw the subgroup lattice for S_3. How does your lattice compare to the one in Exercise 5.71? You should look back at Exercise 4.28 and ponder what just happened.

Exercise 5.73. Find all the subgroups of $D_4 = \{e, r, r^2, r^3, s, sr, sr^2, sr^3\}$ (where r and s are the usual actions) and then draw the subgroup lattice for D_4.

Exercise 5.74. Find all the subgroups of $Q_8 = \{1, -1, i, -i, j, -j, k, -k\}$ and then draw the subgroup lattice for Q_8.

Problem 5.75. What claims can be made about the subgroup lattices of two groups that are isomorphic? What claims can be made about the subgroup lattices of two groups that are not isomorphic? What claims can be made about two groups if their subgroup lattices look nothing alike? *Hint:* The answers to two of these questions should be obvious, but the answer to the remaining question should be something like, "we don't have enough information to make any claims."

Here are two final problems to conclude this section.

Problem 5.76. Several times we've referred to the fact that some subgroups are visible in a Cayley diagram for the parent group and some subgroups are not. Suppose $(G, *)$ is a group and let $H \leq G$. Can you describe a process for creating a Cayley diagram for G that "reveals" the subgroup H inside of this Cayley diagram?

Problem 5.77. Suppose $(G, *)$ is a finite group and let $H \leq G$. Can you describe a process that "reveals" the subgroup H inside the group table for G? Where will the clones for H end up?

5.6 Revisiting Isomorphisms

Suppose $(G_1, *)$ and $(G_2, \circ)$ are two groups. Recall that G_1 and G_2 are isomorphic, written $G_1 \cong G_2$, provided that we can choose generating sets for G_1 and G_2, respectively, so that the Cayley diagrams for both groups are identical (ignoring the labels on the vertices). When two groups are isomorphic, it means that they have identical structure up to relabeling the names of the elements of the group.

One consequence of two groups being isomorphic is that there is a one-to-one correspondence between the elements of the group. This correspondence is referred to as an isomorphism. In other words, an isomorphism is a one-to-one and onto function that preserves the structure of the two groups.

Having an isomorphism between two groups immediately implies that they have the same order, i.e., $|G_1| = |G_2|$ (see Theorem 4.20). However, it is extremely important to remember that two groups having the same order does *not* imply that the two groups are isomorphic. Said another way, having a one-to-one correspondence between two groups does not imply that the two groups are isomorphic. They must also have the same structure!

Exercise 5.78. Provide an example of two groups that have the same order but are not isomorphic.

After we introduced groups tables, we also discussed the fact that $G_1 \cong G_2$ exactly when we can arrange the rows and columns and color elements in such a way that the

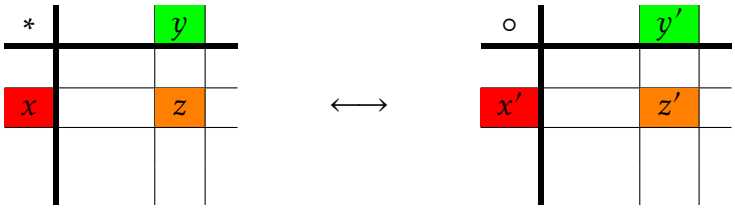


Figure 5.6

colorings for the two group tables agree (see Problem 5.39). The upshot of this is that if $G_1 \cong G_2$, then

the product of corresponding elements yields the corresponding result.

This is the essence of what it means for two groups to have the same structure.

Let's try to make a little more sense of this. Suppose that $G_1 \cong G_2$ and imagine we have arranged the rows and columns of their respective group tables and colored the elements in such a way that the colorings for the two group tables agree. Now, let $x, y \in G_1$. Then these two elements have corresponding elements in the group table for G_2, say x' and y', respectively. In other words, x and x' have the same color while y and y' have the same color. Since G_1 is closed under its binary operation $*$, there exists $z \in G_1$ such that $z = x * y$. There must exist a $z' \in G_2$ such that z' has the same color as z. What must be true of $x' \circ y'$? Since the two tables exhibit the same color pattern, it must be the case that $z' = x' \circ y'$. This is what is means for the product of corresponding elements to yield the corresponding result. Figure 5.6 depicts this phenomenon for group tables.

We can describe the isomorphism between G_1 and G_2 using a function. Let $\phi : G_1 \to G_2$ be the one-to-one and onto function that maps elements of G_1 to their corresponding elements in G_2. Then $\phi(x) = x'$, $\phi(y) = y'$, and $\phi(z) = z'$. Since $z' = x' \circ y'$, we can obtain

$$\phi(x * y) = \phi(z) = z' = x' \circ y' = \phi(x) \circ \phi(y).$$

In summary, it must be the case that

$$\phi(x * y) = \phi(x) \circ \phi(y).$$

We are now prepared to state a formal definition of what it means for two groups to be isomorphic.

Definition 5.79. Let $(G_1, *)$ and $(G_2, \circ)$ be two groups. Then G_1 is **isomorphic** to G_2, written $G_1 \cong G_2$, if and only if there exists a one-to-one and onto function $\phi : G_1 \to G_2$ such that

$$\phi(x * y) = \phi(x) \circ \phi(y). \tag{5.1}$$

The function ϕ is referred to as an **isomorphism**. Equation 5.1 is often referred to as the **homomorphic property**.

You should definitely take a few minutes to convince yourself that the above definition agrees with our previous informal approach to isomorphisms. For those of you that have

had linear algebra, notice that our homomorphic property looks a lot like the requirement for a function on vector spaces to be a linear transformation. Linear transformations preserve the algebraic structure of vector spaces while the homomorphic property is preserving the algebraic structure of groups.

We've seen several instances of two groups being isomorphic, but now that we have a formal definition, we can open the door to more possibilities.

Problem 5.80. Consider the groups $(\mathbb{R}, +)$ and $(\mathbb{R}^+, \cdot)$, where $\mathbb{R}^+$ is the set of positive real numbers. It turns out that these two groups are isomorphic, but this would be difficult to discover using our previous techniques because the groups are infinite. Define $\phi : \mathbb{R} \to \mathbb{R}^+$ via $\phi(r) = e^r$ (where e is the natural base, not the identity). Prove that ϕ is an isomorphism.

Exercise 5.81. For each of the following pairs of groups, determine whether the given function is an isomorphism from the first group to the second group.

(a) $(\mathbb{Z}, +)$ and $(\mathbb{Z}, +)$, $\phi(n) = n + 1$.

(b) $(\mathbb{Z}, +)$ and $(\mathbb{Z}, +)$, $\phi(n) = -n$.

(c) $(\mathbb{Q}, +)$ and $(\mathbb{Q}, +)$, $\phi(x) = x/2$.

Problem 5.82. Show that the groups $(\mathbb{Z}, +)$ and $(2\mathbb{Z}, +)$ are isomorphic.

Perhaps one surprising consequence of the previous problem is that when dealing with infinite groups, a group can have a proper subgroup that it is isomorphic to. Of course, this never happens with finite groups.

Once we know that two groups are isomorphic, there are lots of interesting things we can say. The next theorem tells us that isomorphisms map the identity element of one group to the identity of the second group. It was already clear that this was the case using our informal definition of isomorphic. Prove the next theorem using Definition 5.79

Theorem 5.83. Suppose $\phi : G_1 \to G_2$ is an isomorphism from the group $(G_1, *)$ to the group $(G_2, \circ)$. If e and e' are the identity elements of G_1 and G_2, respectively, then $\phi(e) = e'$.

Theorem 5.84. Suppose $\phi : G_1 \to G_2$ is an isomorphism from the group $(G_1, *)$ to the group $(G_2, \circ)$. Then $\phi(g^{-1}) = [\phi(g)]^{-1}$.

Theorem 5.85. Suppose $\phi : G_1 \to G_2$ is an isomorphism from the group $(G_1, *)$ to the group $(G_2, \circ)$. If G_1 is abelian, then G_2 is abelian.

Theorem 5.86. Suppose $\phi : G_1 \to G_2$ is an isomorphism from the group $(G_1, *)$ to the group $(G_2, \circ)$. Then the function $\phi^{-1} : G_2 \to G_1$ is an isomorphism.

Theorem 5.87. Suppose $\phi : G_1 \to G_2$ and $\psi : G_2 \to G_3$ are isomorphisms from the groups $(G_1, *)$ to $(G_2, \odot)$ and $(G_2, \odot)$ to $(G_3, \star)$, respectively. Then the composite function $\psi \circ \phi$ is an isomorphism of G_1 and G_3.

Theorem 5.88. Let $\mathcal{G}$ be any nonempty collection of groups. Then the relation $\cong$ of being isomorphic is an equivalence relation.

Theorem 5.89. Suppose $\phi : G_1 \to G_2$ is an isomorphism from the group $(G_1, *)$ to the group $(G_2, \circ)$. If $H \leq G_1$, then $\phi(H) \leq G_2$, where

$$\phi(H) := \{y \in G_2 \mid \text{there exists } h \in H \text{ such that } \phi(h) = y\}.$$

Note that $\phi(H)$ is called the **image** of H.

Theorem 5.90. Suppose $(G, *)$ is a group and let $g \in G$. Define $\phi_g : G \to G$ via $\phi(x) = g^{-1}xg$. Then ϕ_g is an isomorphism from G to G. Note that the map ϕ_g is called **conjugation** by g.

Now that you've proved the above theorems, it's a good idea to review the key themes. If you were really paying attention, you may have noticed that in a few of the proofs, we did not use the fact that the function was one-to-one and onto despite assuming that the function was an isomorphism.

Problem 5.91. For which of the recent theorems could we remove the assumption that the function is one-to-one and onto and only assume that it satisfies the homomorphic property? Such functions are called **homomorphisms** and will be the subject of a future chapter.

Chapter 6

Families of Groups

In this chapter we will explore a few families of groups.

6.1 Cyclic Groups

Recall that if $(G,*)$ is a group and $a \in G$, then the subgroup generated by a is given by

$$\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}.$$

According to Theorem 5.66, $\langle a \rangle$ is the smallest subgroup containing a. We call $\langle a \rangle$ the **cyclic group generated by** a. It is important to point out that $\langle a \rangle$ may be finite or infinite. In the finite case, the Cayley diagram with generator a gives us a good indication where the word cyclic comes from.

Exercise 6.1. Suppose $\langle a \rangle$ is a finite group. Since $\langle a \rangle$ is a group in its own right, we can draw a Cayley diagram for this group. Using the generator a, what does the Cayley diagram for $\langle a \rangle$ look like? To rigorously prove that your intuitive thinking is correct, we'll need some results that appear later in this section.

Definition 6.2. Suppose $(G,*)$ is a group and let $a \in G$. We define the **order** of a, written $|a|$, to be the order of $\langle a \rangle$. That is,

$$|a| = |\langle a \rangle|.$$

Exercise 6.3. What is the order of the identity in any group?

Theorem 6.4. Suppose $(G,*)$ is a group and let $a \in G$. Then $\langle a \rangle = \langle a^{-1} \rangle$. In particular, $|a| = |a^{-1}|$.

Exercise 6.5. Find the orders of each of the elements in each of the following groups.

(a) S_2

(b) R_3

(c) R_4

(d) V_4

(e) R_5

(f) R_6

(g) D_3

(h) R_7

(i) R_8

(j) D_4

(k) Q_8

Exercise 6.6. Consider the group $(\mathbb{Z}, +)$. What is the order of 1? Are there any elements in $\mathbb{Z}$ with finite order?

Exercise 6.7. Consider the group of invertible 2×2 matrices with real number entries under the operation of matrix multiplication. This group is denoted $\mathrm{GL}_2(\mathbb{R})$. Find the order of each of the following elements in this group.

(a) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

Theorem 6.8. Suppose $(G, *)$ is a finite group and let $a \in G$. Then there exists a positive integer m such that $a^m = e$, where e is the identity in G.

In fact, we can say something even stronger. You likely noticed the following fact while exploring Exercise 6.5.

Theorem 6.9. Suppose $(G, *)$ is a group and let $a \in G$.

(a) If $|a| = n < \infty$, then $a^n = e$ and $e, a, a^2, \ldots, a^{n-1}$ are all distinct elements of $\langle a \rangle$.

(b) If $|a| = \infty$, then $a^n \neq e$ for all $n \neq 0$ and $a^n \neq a^m$ whenever $n \neq m$ in $\mathbb{Z}$.

Corollary 6.10. Suppose $(G, *)$ is a finite group and let $a \in G$. Then the order of a is the smallest positive integer n such that $a^n = e$.

Exercise 6.11. Notice that in the definition for $\langle a \rangle$, we allow the exponents on a to be negative. Explain why we only need to use positive exponents when $\langle a \rangle$ is a finite group. What about when $\langle a \rangle$ is infinite?

Problem 6.12. Suppose $(G, *)$ is a group $a \in G$ with $|a| = n$. For what other exponents k will it be true that $a^k = e$? You'll have an opportunity to prove your claim later.

We are finally ready to introduce our family of interest for this section.

Definition 6.13. Suppose $(G, *)$ is a group. Then we say that G is a **cyclic group** if and only if there exists $a \in G$ such that $\langle a \rangle = G$.

It is clear that if G is cyclic with generator a, then $|G| = |a|$. In fact, if $a \in G$, the converse is true, as well.

Exercise 6.14. Determine which of the groups from Exercise 6.5 are cyclic. If the group is cyclic, find at least one generator.

Exercise 6.15. Determine whether each of the following groups are cyclic. If the group is cyclic, find at least one generator.

(a) $(\mathbb{Z}, +)$

(b) $(\mathbb{R}, +)$

(c) $(\mathbb{R}^+, \cdot)$

(d) $(\{6^n \mid n \in \mathbb{Z}\}, \cdot)$

(e) $\mathrm{GL}_2(\mathbb{R})$ under matrix multiplication

(f) $\{(\cos(\pi/4) + i\sin(\pi/4))^n \mid n \in \mathbb{Z}\}$ under multiplication of complex numbers

Theorem 6.16. If $(G, *)$ is a cyclic group, then G is abelian.

Exercise 6.17. Provide an example of a finite group that is abelian but not cyclic.

Exercise 6.18. Provide an example of an infinite group that is abelian but not cyclic.

Theorem 6.19. Suppose $(G, *)$ is a cyclic group such that G has exactly one element that generates all of G. Then the order of G is at most order 2.

Theorem 6.20. Suppose $(G, *)$ is a group such that G has no proper nontrivial subgroups. Then G is cyclic.

Theorem 6.21. Suppose $(G, *)$ is an infinite cyclic group. Then G is isomorphic to $\mathbb{Z}$ (under the operation of addition).

Recall that for $n \geq 3$, R_n is the group of rotational symmetries of a regular n-gon, where the operation is composition of actions.

Theorem 6.22. For all $n \geq 3$, R_n is cyclic.

Theorem 6.23. Suppose $(G, *)$ is a finite cyclic group of order $n \geq 1$. Then G is isomorphic to R_n if $n \geq 3$, S_2 if $n = 2$, and the trivial group if $n = 1$.

The upshot of Theorems 6.21 and 6.23 is that up to isomorphism, we know exactly what all of the cyclic groups are.

Exercise 6.24. Suppose $(G, *)$ is a finite cyclic group of order n with generator a. If we write down the group table for G using $e, a, a^2, \ldots, a^{n-1}$ as the labels for the rows and columns, are there any interesting patterns in the table?

Recall that two integers are **relatively prime** if they have no factors other than 1 in common. That is, integers n and k are relatively prime iff $\gcd(n, k) = 1$.

Definition 6.25. Let $n \in \mathbb{N}$ and define the following sets.

(a) $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$

(b) $U(n) := \{k \in \mathbb{Z}_n \mid \gcd(n, k) = 1\}$

For each set above, the immediate goal is to find a binary operation that will yield a group. The key is to use modular arithmetic. To calculate the sum (respectively, product) of two integers mod n, add (respectively, multiply) the two numbers and then find the remainder after dividing the sum (respectively, product) by n. For example, $4 + 9$ is 3 mod 5 since 13 has remainder 3 when being divided by 5. Similarly, $4 \cdot 9$ is 1 mod 5 since 36 has remainder 1 when being divided by 5.

Theorem 6.26. The set $(\mathbb{Z}_n, + \bmod n)$ is a group.

Theorem 6.27. The set $(U(n), \cdot \bmod n)$ is a group.

Exercise 6.28. Consider $\mathbb{Z}_4$.

(a) Find the group table for $\mathbb{Z}_4$.

(b) Is $\mathbb{Z}_4$ cyclic? If so, list elements of $\mathbb{Z}_4$ that individually generate $\mathbb{Z}_4$. If $\mathbb{Z}_4$ is not cyclic, explain why.

(c) Is $\mathbb{Z}_4$ isomorphic to either of R_4 or V_4? Justify your answer.

(d) Draw the subgroup lattice for $\mathbb{Z}_4$.

Exercise 6.29. Consider $U(10) = \{1, 3, 7, 9\}$.

(a) Find the group table for $U(10)$.

(b) Is $U(10)$ cyclic? If so, list elements of $U(10)$ that individually generate $U(10)$. If $U(10)$ is not cyclic, explain why.

(c) Is $U(10)$ isomorphic to either of R_4 or V_4? Justify your answer.

(d) Is $U(10)$ isomorphic to $\mathbb{Z}_4$? Justify your answer.

(e) Draw the subgroup lattice for $U(10)$.

Exercise 6.30. Consider $U(12) = \{1, 5, 7, 11\}$.

(a) Find the group table for $U(12)$.

(b) Is $U(12)$ cyclic? If so, list elements of $U(12)$ that individually generate $U(12)$. If $U(12)$ is not cyclic, explain why.

(c) Is $U(12)$ isomorphic to either of R_4 or V_4? Justify your answer.

(d) Draw the subgroup lattice for $U(12)$.

In light of Exercise 6.29 and 6.30, $U(n)$ may or may not be cyclic. Nonetheless, as the next theorem illustrates, $U(n)$ is always abelian.

Theorem 6.31. For all n, $U(n)$ is abelian.

The upshot of the next theorem is that for $n \geq 3$, $\mathbb{Z}_n$ is just the set of (smallest nonnegative) exponents on r in R_n.

Theorem 6.32. For $n \geq 3$, $\mathbb{Z}_n \cong R_n$. Moreover, $\mathbb{Z}_2 \cong S_2$ and $\mathbb{Z}_1$ is isomorphic to the trivial group.

One consequence of the previous theorem is that $\mathbb{Z}_n$ is always cyclic. Combining the results of Theorems 6.23 and 6.21 together with Theorem 6.32, we immediately obtain the following.

Theorem 6.33. Let $(G,*)$ be a cyclic group. If the order of G is infinite, then $(G,*)$ is isomorphic to $(\mathbb{Z},+)$. If G has finite order n, then $(G,*)$ is isomorphic to $(\mathbb{Z}_n,+ \bmod n)$.

Now that we have a complete description of the cyclic groups, let's focus our attention on subgroups of cyclic groups. The next result should look familiar and will come in handy. In particular, it will be useful when proving Theorems 6.35 and6.37. We'll take the result for granted and not worry about proving it right now.

Theorem 6.34 (Division Algorithm for $\mathbb{Z}$)**.** If m is a positive integer and n is any integer, then there exist unique integers q (called the **quotient**) and r (called the **remainder**) such that $n = mq + r$, where $0 \leq r < m$.

Theorem 6.35. Suppose $(G,*)$ is a group and let $a \in G$ such that $|a| = n$. Then $a^i = a^j$ iff n divides $i - j$.

Compare the next result to Problem 6.12.

Corollary 6.36. Suppose $(G,*)$ is a group and let $a \in G$ such that $|a| = n$. If $a^k = e$, then $|a|$ divides k.

Theorem 6.37. Suppose $(G,*)$ is a cyclic group. If $H \leq G$, then H is also cyclic.

It turns out that for proper subgroups, the converse of Theorem 6.37 is not true.

Exercise 6.38. Provide an example of a group $(G,*)$ such that G is not cyclic, but all proper subgroups of G are cyclic.

The next result officially settles Exercise 5.56(d) and also provides a complete description of the subgroups of infinite cyclic groups up to isomorphism.

Corollary 6.39. The subgroups of $\mathbb{Z}$ are precisely the groups $n\mathbb{Z}$ under addition for $n \in \mathbb{Z}$.

What about finite cyclic groups?

Theorem 6.40. Suppose $(G,*)$ is a finite cyclic group with generator a such that $|G| = n$.

(a) Then $|a^s| = \dfrac{n}{\gcd(n,s)}$.

(b) Moreover, $\langle a^s \rangle = \langle a^t \rangle$ iff $\gcd(s,n) = \gcd(t,n)$.

Exercise 6.41. Suppose $(G,*)$ is a cyclic group of order 12 with generator a.

(a) Find the orders of each of the following elements: a^2, a^7, a^8.

(b) Which elements of G individually generate G?

Corollary 6.42. Suppose $(G,*)$ is a finite cyclic group with generator a such that $|G| = n$. Then $\langle a\rangle = \langle a^r\rangle$ iff n and r are relatively prime. That is, a^r generates G iff n and r are relatively prime.

Exercise 6.43. Consider $(\mathbb{Z}_{18}, + \bmod 18)$.

(a) Find all of the elements of $\mathbb{Z}_{18}$ that individually generate all of $\mathbb{Z}_{18}$.

(b) Draw the subgroup lattice for $\mathbb{Z}_{18}$. For each subgroup, list the elements of the corresponding set. Moreover, circle the the elements in each subgroup that individually generate that subgroup. For example, $\langle 2\rangle = \{0,2,4,6,8,10,12,14,16\}$. In this case, we should circle 2, 4, 8, 10, 14, and 16 since each of these elements individually generate $\langle 2\rangle$ and none of the remaining elements do. I'll leave it to you to figure out why this is true.

Exercise 6.44. Repeat the above exercise, but this time use $\mathbb{Z}_{12}$ instead of $\mathbb{Z}_{18}$.

Corollary 6.45. Suppose $(G,*)$ is a finite cyclic group such that $|G| = p$, where p is prime. Then G has no proper nontrivial subgroups.

Problem 6.46. Let p and q be distinct primes. Find the number of generators of $\mathbb{Z}_{pq}$.

Problem 6.47. Let p be a prime. Find the number of generators of $\mathbb{Z}_{p^r}$, where r is an integer greater than or equal to 1.

Problem 6.48. If there is exactly one group up to isomorphism of order n, then to what group are all the groups of order n isomorphic?

6.2 Dihedral Groups

We can think of cyclic groups as groups that describe rotational symmetry. In particular, R_n is the group of rotational symmetries of a regular n-gon. Dihedral groups are those groups that describe both rotational and reflection symmetry of regular n-gons.

Definition 6.49. For $n \geq 3$, the **dihedral group** D_n is defined to be the group consisting of the symmetry actions of a regular n-gon, where the operation is composition of actions.

For example, as we've seen, D_3 and D_4 are the symmetry groups of equilateral triangles and squares, respectively. The symmetry group of a regular pentagon is denoted by D_5. It is a well-known fact from geometry that the composition of two reflections in the plane is a rotation by twice the angle between the reflecting lines.

Theorem 6.50. The group D_n is a non-abelian group of order $2n$.

Theorem 6.51. For $n \geq 3$, $R_n \leq D_n$.

Theorem 6.52. Fix $n \geq 3$ and consider D_n. Let r be rotation clockwise by $360°/n$ and let s and s' be any two adjacent reflections of a regular n-gon. Then

(a) $D_n = \langle r, s \rangle = \{\underbrace{e, r, r^2, \ldots, r^{n-1}}_{\text{rotations}}, \underbrace{s, sr, sr^2, \ldots, sr^{n-1}}_{\text{reflections}}\}$ and

(b) $D_n = \langle s, s' \rangle =$ all possible products of s and s'.

Theorem 6.53. Fix $n \geq 3$ and consider D_n. Let r be rotation clockwise by $360°/n$ and let s and s' be any two adjacent reflections of a regular n-gon. Then the following relations hold.

(a) $r^n = s^2 = (s')^2 = e$,

(b) $r^{-k} = r^{n-k}$ (special case: $r^{-1} = r^{n-1}$),

(c) $sr^k = r^{n-k}s$ (special case: $sr = r^{n-1}s$),

(d) $\underbrace{ss's\cdots}_{n \text{ factors}} = \underbrace{s'ss'\cdots}_{n \text{ factors}}$.

Exercise 6.54. From Theorem 6.52, we know $D_n = \langle r, s \rangle = \{\underbrace{e, r, r^2, \ldots, r^{n-1}}_{\text{rotations}}, \underbrace{s, sr, sr^2, \ldots, sr^{n-1}}_{\text{reflections}}\}$. If you were to create the group table for D_n so that the rows and columns of the table were labeled by $e, r, r^2, \ldots, r^{n-1}, s, sr, sr^2, \ldots, sr^{n-1}$ (in exactly that order), do any patterns arise? *Hint:* Where are the rotations? Where are the reflections?

6.3 Symmetric Groups

Recall the group S_3 from Exercise 4.27. This group acts on three coins that are in a row by rearranging their positions (but not flipping them over). This group is an example of a **symmetric group**. In general, the symmetric group on n objects is the set of permutations that rearranges the n objects. The group operation is composition of permutations. Let's be a little more formal.

Definition 6.55. A **permutation of a set** A is a function $\sigma : A \to A$ that is both one-to-one and onto.

You should take a moment to convince yourself that the formal definition of a permutation agrees with the notion of rearranging the set of objects. The do-nothing action is the identity permutation, i.e., $\sigma(a) = a$ for all $a \in A$. There are many ways to represent a permutation. One visual way is using **permutation diagrams**, which we will introduce via examples.

Consider the following diagrams:

Each of these diagrams represents a permutation on five objects. I've given the permutations the names α, β, σ, and γ. The intention is to read the diagrams from the top down. The numbers labeling the nodes along the top are identifying position. Following an edge from the top row of nodes to the bottom row of nodes tells us what position an object moves to. It is important to remember that the numbers are referring to the position of an object, not the object itself. For example, β is the permutation that sends the object in the second position to the fourth position, the object in the third position to the second position, and the object in the fourth position to the third position. Moreover, the permutation β doesn't do anything to the objects in positions 1 and 5.

Exercise 6.56. Describe in words what the permutations σ and γ do.

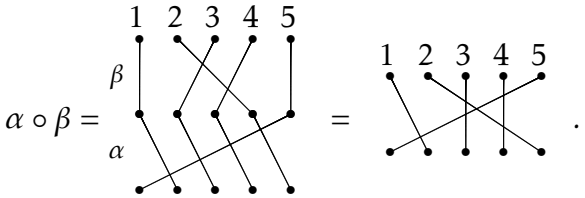
Exercise 6.57. Draw the permutation diagram for the do-nothing permutation on 5 objects. This is called the **identity permutation**. What does the identity permutation diagram look like in general for arbitrary n?

Definition 6.58. The set of all permutations on n objects is denoted by S_n.

Exercise 6.59. Draw all the permutation diagrams for the permutations in S_3.

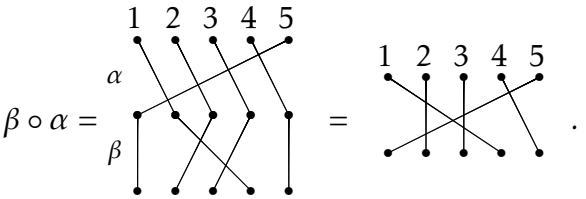
Exercise 6.60. How many distinct permutations are there in S_4? How about S_n for any $n \in \mathbb{N}$?

If S_n is going to be a group, we need to know how to compose permutations. This is easy to do using the permutation diagrams. Consider the permutations α and β from earlier. We can represent the composition $\alpha \circ \beta$ via



As you can see by looking at the figure, to compose two permutations, you stack the one that goes first in the composition (e.g., β in the example above) on top of the other and just follow the edges from the top through the middle to the bottom. If you think about how function composition works, this is very natural. The resulting permutation is determined by where we begin and where we end in the composition.

We already know that the order of composition matters for functions, and so it should matter for the composition of permutations. To make this crystal clear, let's compose α and β in the opposite order. We see that

$\beta \circ \alpha =$ 

.

The moral of the story is that composition of permutations does not necessarily commute.

Exercise 6.61. Consider α, β, σ, and γ from earlier. Can you find a pair of permutations that do commute? Can you identify any features about your diagrams that indicate why they commuted?

Exercise 6.62. Fix $n \in \mathbb{N}$. Convince yourself that any $\rho \in S_n$ composed with the identity permutation (in either order) equals ρ.

If S_n is going to be a group, we need to know what the inverse of a permutation is.

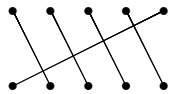
Problem 6.63. Given a permutation $\rho \in S_n$, describe a method for constructing ρ^{-1}. Briefly justify that $\rho \circ \rho^{-1}$ will yield the identity permutation.

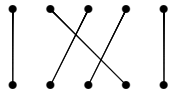
At this point, we have all the ingredients we need to prove that S_n forms a group under composition of permutations.

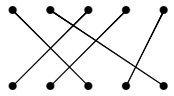
Theorem 6.64. The set of permutations on n objects forms a group under the operation of composition. That is, $(S_n, \circ)$ is a group. Moreover, $|S_n| = n!$.

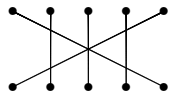
Note that it is standard convention to omit the composition symbol when writing down compositions in S_n. For example, we will simply write $\alpha\beta$ to denote $\alpha \circ \beta$.

Permutation diagrams are fun to play with, but we need a more efficient way of encoding information. One way to do this is using **cycle notation**. Consider α, β, σ, and γ in S_5 from the previous examples. Below I have indicated what each permutation is equal to using cycle notation.

$\alpha =$  $= (1,2,3,4,5)$

$\beta =$  $= (2,4,3)$

$\sigma =$  $= (1,3)(2,5,4)$

$\gamma =$  $= (1,5)$

Each string of numbers enclosed by parentheses is called a **cycle** and if the string of numbers has length k, then we call it a k-cycle. For example, α consists of a single 5-cycle, whereas σ consists of one 2-cycle and one 3-cycle. In the case of σ, we say that σ is the product of two **disjoint cycles**.

One observation that you hopefully made is that if an object in position i remains unchanged, then we don't bother listing that number in the cycle notation. However, if we wanted to, we could use the 1-cycle (i) to denote this. For example, we could write $\beta = (1)(2,3,4)(5)$. In particular, we could denote the identity permutation in S_5 using $(1)(2)(3)(4)(5)$. Yet, it is common to simply use (1) to denote the identity in S_n for all n.

Notice that the first number we choose to write down for a given cycle is arbitrary. However, the numbers that follow are not negotiable. Typically, we would use the smallest possible number first, but this is not necessary. For example, the cycle $(2,4,7)$ could also be written as $(4,7,2)$ or $(7,2,4)$.

Exercise 6.65. Write down all 6 elements in S_3 using cycle notation.

Exercise 6.66. Write down all 24 elements in S_4 using cycle notation.

Suppose $\sigma \in S_n$. Since σ is one-to-one and onto, it is clear that it is possible to write σ as a product of disjoint cycles such that each $i \in \{1, 2, \ldots, n\}$ appears exactly once.

Let's see if we can figure out how to multiply elements of S_n using cycle notation. Consider the permutations $\alpha = (1,3,2)$ and $\beta = (3,4)$ in S_4. To compute the composition $\alpha\beta = (1,3,2)(3,4)$, let's explore what happens in each position. Since we are doing function composition, we should work our way from right to left. Since 1 does not appear in the cycle notation for β, we know that $\beta(1) = 1$ (i.e., β maps 1 to 1). Now, we see what $\alpha(1) = 3$. Thus, the composition $\alpha\beta$ maps 1 to 3 (since $\alpha\beta(1) = \alpha(\beta(1)) = \alpha(1) = 3$). Next, we should return to β and see what happens to 3—which is where we ended a moment ago. We see that β maps 3 to 4 and then α maps 4 to 4 (since 4 does not appear in the cycle notation for α). So, $\alpha\beta(3) = 4$. Continuing this way, we see that β maps 4 to 3 and α maps 3 to 2, and so $\alpha\beta$ maps 4 to 2. Lastly, since $\beta(2) = 2$ and $\alpha(2) = 1$, we have $\alpha\beta(2) = 1$. Putting this altogether, we see that $\alpha\beta = (1,3,4,2)$. Now, you should try a few. Things get a little trickier if the composition of two permutations results in a permutation consisting of more than a single cycle.

Exercise 6.67. Consider α, β, σ, and γ for which we drew the permutation diagrams. Using cycle notation, compute each of the following.

(a) $\alpha\gamma$

(b) α^2

(c) α^3

(d) α^4

(e) α^5

(f) $\sigma\alpha$

(g) $\alpha^{-1}\sigma^{-1}$

(h) β^2

(i) β^3

(j) $\beta\gamma\alpha$

(k) σ^3

(l) σ^6

Exercise 6.68. Write down the group table for S_3 using cycle notation.

In Exercise 6.66, one of the permutations you should have written down is $(1,2)(3,4)$. This is a product of two disjoint 2-cycles. It is worth pointing out that each cycle is a permutation in its own right. That is, $(1,2)$ and $(3,4)$ are each permutations. It just so happens that their composition does not "simplify" any further. Moreover, these two disjoint 2-cycles commute since $(1,2)(3,4) = (3,4)(1,2)$. In fact, this phenomenon is always true.

Theorem 6.69. Suppose α and β are two disjoint cycles. Then $\alpha\beta = \beta\alpha$. That is, products of disjoint cycles commute.

Computing the order of a permutation is fairly easy using cycle notation once we figure out how to do it for a single cycle. In fact, you've probably already guessed at the following theorem.

Theorem 6.70. Suppose $\alpha \in S_n$ such that α consists of a single k-cycle. Then $|\alpha| = k$.

Theorem 6.71. Suppose $\alpha \in S_n$ such that α consists of m disjoint cycles of lengths $k_1, \ldots, k_m$. Then $|\alpha| = \text{lcm}(k_1, \ldots, k_m)$.*

Problem 6.72. Is the previous theorem true if we do not require the cycles to be disjoint? Justify your answer.

Exercise 6.73. Compute the orders of all the elements in S_3. See Exercise 6.65.

Exercise 6.74. Compute the orders of all the elements in S_4. See Exercise 6.66.

Exercise 6.75. What is the order of $(1,4,7)(2,5)(3,6,8,9)$?

Exercise 6.76. Draw the subgroup lattice for S_3.

Exercise 6.77. Now, using $(1,2)$ and $(1,2,3)$ as generators, draw the Cayley diagram for S_3. Look familiar?

It turns out that the subgroups of symmetric groups play an important role in group theory.

Definition 6.78. Every subgroup of a symmetric group is called a **permutation group**.

The proof of the following theorem isn't too bad, but for now we'll take it for granted.

Theorem 6.79 (Cayley's Theorem)**.** Every finite group is isomorphic to some permutation group. In particular, if $(G, *)$ is a group of order n, then G is isomorphic to a subgroup of S_n.

Cayley's Theorem guarantees that every finite group is isomorphic to a permutation group and it turns out that there is a rather simple algorithm for constructing the corresponding permutation group. I'll briefly explain an example and then let you try a couple.

Consider the Klein four-group $V_4 = \{e, v, h, vh\}$. Recall that V_4 has the following group table.

*Recall that $\text{lcm}(k_1, \ldots, k_m)$ is the **least common multiple** of $\{k_1, \ldots, k_m\}$.

$*$	e	v	h	vh
e	e	v	h	vh
v	v	e	vh	h
h	h	vh	e	v
vh	vh	h	v	e

If we number the elements e, v, h, and vh as $1, 2, 3$, and 4, respectively, then we obtain the following table.

	1	2	3	4
1	1	2	3	4
2	2	1	4	3
3	3	4	1	2
4	4	3	2	1

Comparing each of the four columns to the leftmost column, we can obtain the corresponding permutations. In particular, we obtain

$$\begin{aligned} e &\leftrightarrow (1) \\ v &\leftrightarrow (1,2)(3,4) \\ h &\leftrightarrow (1,3)(2,4) \\ vh &\leftrightarrow (1,4)(2,3). \end{aligned}$$

Do you see where these permutations came from? The claim is that the set of permutations $\{(1), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$ is isomorphic to V_4. In this particular case, it's fairly clear that this is true. However, it takes some work to prove that this process will always result in an isomorphic permutation group. In fact, verifying the algorithm is essentially the proof of Cayley's Theorem.

Since there are potentially many ways to rearrange the rows and columns of a given table, it should be clear that there are potentially many isomorphisms that could result from the algorithm described above.

Here's another way to obtain a permutation group that is isomorphic to a given group. Let's consider V_4 again. Recall that V_4 is a subset of D_4, which is the symmetry group for a square. Alternatively, V_4 is the symmetry group for a non-square rectangle. Label the corners of the rectangle 1, 2, 3, and 4 by starting in the upper left corner and continuing clockwise. Recall that v is the action that reflects the rectangle over the vertical midline. The result of this action is that the corners labeled by 1 and 2 switch places and the corners labeled by 3 and 4 switch places. Thus, v corresponds to the permutation $(1,2)(3,4)$. Similarly, h swaps the corners labeled by 1 and 4 and the corners labeled by 2 and 3, and so h corresponds to the permutation $(1,4)(2,3)$. Notice that this is not the same answer we got earlier and that's okay as there may be many permutation representations for a given group. Lastly, vh rotates the rectangle $180°$ which sends ends up swapping corners labeled 1 and 3 and swapping corners labeled by 2 and 4. Therefore, vh corresponds to the permutation $(1,3)(2,4)$.

Exercise 6.80. Find a permutation group that is isomorphic to D_4.

Exercise 6.81. Find a permutation group that is isomorphic to $\mathbb{Z}_6$.

Exercise 6.82. Consider S_3.

(a) Using $(1,2)$, $(1,3)$, and $(2,3)$ as generators, draw the Cayley diagram for S_3.

(b) In the previous part, we used a generating set with three elements. Is there a smaller generating set? If so, what is it?

Exercise 6.83. Recall that there are $4! = 24$ permutations in S_4.

(a) Pick any 12 permutations from S_4 and verify that you can write them as words in the 2-cycles $(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)$. In most circumstances, your words will not consist of products of disjoint 2-cycles. For example, the permutation $(1,2,3)$ can be decomposed into $(1,2)(2,3)$, which is a word consisting of two 2-cycles that happen to not be disjoint.

(b) Using your same 12 permutations, verify that you can write them as words only in the 2-cycles $(1,2),(2,3),(3,4)$.

By the way, it might take some trial and error to come up with a way to do this. Moreover, there is more than one way to do it.

As the previous exercises hinted at, the 2-cycles play a special role in the symmetric groups. In fact, they have a special name. A **transposition** is a single cycle of length 2. In the special case that the transposition is of the form $(i, i+1)$, we call it an **adjacent transposition**. For example, $(3,7)$ is a (non-adjacent) transposition while $(6,7)$ is an adjacent transposition.

It turns out that the set of transpositions in S_n is a generating set for S_n. In fact, the adjacent transpositions form an even smaller generating set S_n. To get some intuition, let's play with a few examples.

Exercise 6.84. Try to write each of the following permutations as a product of transpositions. You do not necessarily need to use adjacent transpositions.

(a) $(3,1,5)$

(b) $(2,4,6,8)$

(c) $(3,1,5)(2,4,6,8)$

(d) $(1,6)(2,5,3)$

The products you found in the previous exercise are called **transposition representations** of the given permutation.

Problem 6.85. Consider the arbitrary k-cycle $(a_1, a_2, \ldots, a_k)$ from S_n (with $k \leq n$). Find a way to write this permutation as a product of 2-cycles.

Problem 6.86. Consider the arbitrary 2-cycle (a, b) from S_n. Find a way to write this permutation as a product of adjacent 2-cycles.

The previous two problems imply the following theorem.

Theorem 6.87. Consider S_n.

1. Every permutation in S_n can be written as a product of transpositions.
2. Every permutation in S_n can be written as a product of adjacent transpositions.

Corollary 6.88. The set of transpositions (respectively, the set of adjacent transpositions) from S_n forms a generating set for S_n.

It is important to point out that the transposition representation of a permutation is not unique. That is, there are many words in the transpositions that will equal the same permutation. However, as we shall see in the next section, given two transposition representations for the same permutation, the number of transpositions will have the same parity (i.e., even versus odd).

Remark 6.89. Here are two interesting facts that I will let you ponder on your own time.

(a) The group of rigid motion symmetries for a cube is isomorphic to S_4. To convince yourself of this fact, first prove that this group has 24 actions and then ponder the action of S_4 on the four long diagonals of a cube.

(b) It turns out that you can generate S_4 with $(1,2)$ and $(1,2,3,4)$. Moreover, you can arrange the Cayley diagram for S_4 with these generators on a truncated cube, which is depicted in Figure 6.1. Try it.

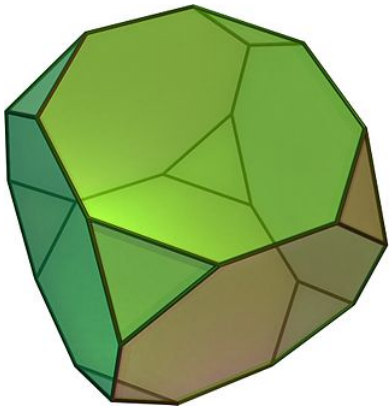


Figure 6.1. Truncated cube. [Image source: Wikipedia]

6.4 Alternating Groups

In this section, we describe a special class of permutation groups. To get started, let's play with a few exercises.

Exercise 6.90. Write down every permutation in S_3 as a product of 2-cycles in the most efficient way you can find (i.e., use the fewest possible transpositions). Now, write every permutation in S_3 as a product of adjacent 2-cycles, but don't worry about whether your decompositions are efficient. Any observations about the number of transpositions you used in each case? Think about even versus odd.

Lemma 6.91. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_k$ is a collection of 2-cycles in S_n such that $\alpha_1\alpha_2\cdots\alpha_k = (1)$. Then k must be even. *Hint:* Use strong induction on k. Start by showing that $k \neq 1$ but that the statement is true when $k = 2$. Then assume that $k > 2$ and proceed by induction.

Theorem 6.92. Let $\sigma \in S_n$. Then every transposition representation of σ has the same parity.

The previous theorem tells us that the following definition is well-defined.

Definition 6.93. A permutation is **even** (respectively, **odd**) if one of its transposition representations consists of an even (respectively, odd) number of transpositions.

Exercise 6.94. Classify all of the permutations in S_3 as even or odd.

Exercise 6.95. Classify all of the permutations in S_4 as even or odd.

Exercise 6.96. Determine whether $(1, 4, 2, 3, 5)$ is even or odd. How about $(1, 4, 2, 3, 5)(7, 9)$?

Problem 6.97. Consider the arbitrary k-cycle $(a_1, a_2, \ldots, a_k)$ from S_n (with $k \leq n$). When will this cycle be odd versus even? Briefly justify your answer.

Problem 6.98. Conjecture a statement about when a permutation will be even versus odd. Briefly justify your answer.

And finally, we are ready to introduce the alternating groups.

Definition 6.99. The set of all even permutations in S_n is denoted by A_n and is called the **alternating group**.

Since we referred to A_n as a group, it darn well better be a group!

Theorem 6.100. The set A_n forms a group under composition of permutations and has order $n!/2$.

Exercise 6.101. Find A_3. What group is A_3 isomorphic to?

Exercise 6.102. Find A_4 and then draw its subgroup lattice. Is A_4 abelian?

Exercise 6.103. What is the order of A_5? Is A_5 abelian?

Exercise 6.104. What are the possible orders for elements in S_6 and A_6? What about S_7 and A_7?

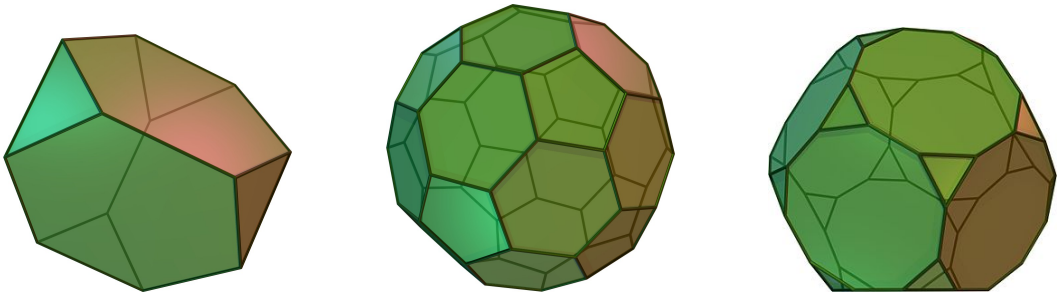
Exercise 6.105. Does A_8 contain an element of order 15? If so, find one. If not, explain why no such element exists.

Remark 6.106. Below are a few interesting facts about A_4 and A_5, which we will state without proof.

(a) The group of rigid motion symmetries for a regular tetrahedron is isomorphic to A_4.

(b) You can arrange the Cayley diagram for A_4 with generators $(1,2)(3,4)$ and $(2,3,4)$ on a truncated tetrahedron, which is depicted in Figure 6.2(a).

(c) You can arrange the Cayley diagram for A_5 with generators $(1,2)(3,4)$ and $(1,2,3,4,5)$ on a truncated icosahedron, which is given in Figure 6.2(b). You can also arrange the Cayley diagram for A_5 with generators $(1,2,3)$ and $(1,5)(2,4)$ on a truncated dodecahedron seen in Figure 6.2(c).



(a) (b) (c)

Figure 6.2. Truncated tetrahedron, truncated icosahedron, and truncated dodecahedron. [Image source: Wikipedia]

Chapter 7

Cosets, Lagrange's Theorem, and Normal Subgroups

7.1 Cosets

Undoubtably, you've noticed numerous times that if G is a group with $H \leq G$ and $g \in G$, then both $|H|$ and $|g|$ divide $|G|$. The theorem that says this is always the case is called Lagrange's theorem and we'll prove it towards the end of this chapter. We begin with a definition.

Definition 7.1. Let G be a group and let $H \leq G$ and $a \in G$. The subsets

$$aH := \{ah \mid h \in H\}$$

and

$$Ha := \{ha \mid h \in H\}$$

are called the **left** and **right cosets of** H **containing** a, respectively.

To gain some insight, let's tinker with an example. Consider the dihedral group $D_3 = \langle r, s \rangle$ and let $H = \langle s \rangle \leq D_3$. To compute the right cosets of H, we need to multiply all of the elements of H on the right by the elements of G. We see that

$$\begin{aligned}
He &= \{ee, se\} = \{e, s\} = H \\
Hr &= \{er, sr\} = \{r, sr\} \\
Hr^2 &= \{er^2, sr^2\} = \{r^2, rs\} \\
Hs &= \{es, ss\} = \{s, e\} = H \\
Hsr &= \{esr, ssr\} = \{sr, r\} \\
Hrs &= \{ers, srs\} = \{rs, ssr^2\} = \{rs, r^2\}.
\end{aligned}$$

Despite the fact that we made six calculations (one for each element in D_3), if we scan the list, we see that there are only 3 distinct cosets, namely

$$\begin{aligned}
H = He = Hs &= \{e, s\} \\
Hr = Hsr &= \{r, sr\} \\
Hr^2 = Hrs &= \{r^2, rs\}.
\end{aligned}$$

We can make a few more observations. First, the resulting cosets formed a partition of D_3. That is, every element of D_3 appears in exactly one coset. Moreover, all the cosets are the same size—two elements in each coset in this case. Lastly, each coset can be named in multiple ways. In particular, the elements of the coset are exactly the elements of D_3 we multiplied H by. For example, $Hr = Hsr$ and the elements of this coset are r and sr. Shortly, we will see that these observations hold, in general.

Here is another significant observation we can make. Consider the Cayley diagram for D_3 with generating set $\{r, s\}$ that is given in Figure 7.1. Given this Cayley diagram, we can visualize the subgroup H and it's clones. Moreover, H and it's clones are exactly the 3 right cosets of H. We'll see that, in general, the *right* cosets of a given subgroup are always the subgroup and its clones.

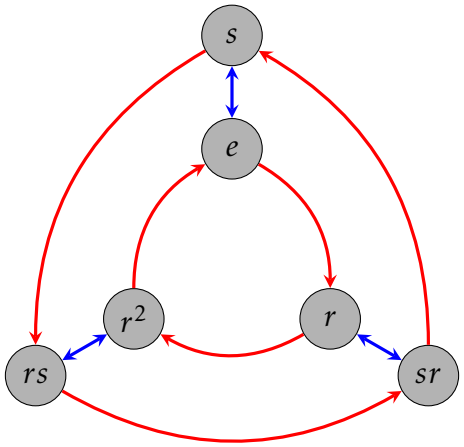


Figure 7.1. Cayley diagram for D_3 with generating set $\{r, s\}$.

Exercise 7.2. Consider the group D_3. Find all the left cosets for $H = \langle s \rangle$. Are they the same as the right cosets? Are they the same as the subgroup H and its clones that we can see in the Cayley graph for D_3 with generating set $\{r, s\}$?

As the previous exercise indicates, the collections of left and right cosets may not be the same and when they are not the same, the subgroup and its clones do not coincide with the left cosets.

You might be thinking that somehow right cosets are better than left cosets since we were able to visualize them in the Cayley graph. However, this is not the case. Our convention of composing actions from right to left is what is dictating the visualization. If we had adopted a left to right convention, then we would be able to visualize the left cosets.

Computing left and right cosets using a group table is fairly easy. Hopefully, you figured out in Exercise 7.2 that the left cosets of $H = \langle s \rangle$ in D_3 are $H = \{e, s\}$, $srH = \{r^2, sr\}$, and $rsH = \{r, rs\}$. Now, consider the following group table for D_3 that has the rows and columns arranged according to the left cosets of H.

$*$	e	s	sr	r^2	rs	r
e	e	s	sr	r^2	rs	r
s	s	e	r	rs	r^2	sr
sr	sr	r^2	e	s	r	rs
r^2	r^2	sr	rs	r	s	e
rs	rs	r	r^2	sr	e	s
r	r	rs	s	e	sr	r^2

The left coset srH must appear in the row labeled by sr and in the columns labeled by the elements of $H = \{e, s\}$. We've depicted this below.

$*$	e	s	sr	r^2	rs	r
e	e	s	sr	r^2	rs	r
s	s	e	r	rs	r^2	sr
sr	sr	r^2	e	s	r	rs
r^2	r^2	sr	rs	r	s	e
rs	rs	r	r^2	sr	e	s
r	r	rs	s	e	sr	r^2

On the other hand, the right coset Hsr must appear in the column labeled by sr and the rows labeled by the elements of $H = \{e, s\}$:

$*$	e	s	sr	r^2	rs	r
e	e	s	sr	r^2	rs	r
s	s	e	r	rs	r^2	sr
sr	sr	r^2	e	s	r	rs
r^2	r^2	sr	rs	r	s	e
rs	rs	r	r^2	sr	e	s
r	r	rs	s	e	sr	r^2

As we can see from the tables, $srH \neq Hsr$ since $\{sr, r^2\} \neq \{sr, r\}$. If we color the entire group table for D_3 according to which *left* coset an element belongs to, we get the following.

$*$	e	s	sr	r^2	rs	r
e	e	s	sr	r^2	rs	r
s	s	e	r	rs	r^2	sr
sr	sr	r^2	e	s	r	rs
r^2	r^2	sr	rs	r	s	e
rs	rs	r	r^2	sr	e	s
r	r	rs	s	e	sr	r^2

We would get a similar table (but in this case, not identical) if we colored the elements according to the right cosets.

Let's tackle a few more examples.

Exercise 7.3. Consider D_3 and let $K = \langle r \rangle$.

(a) Find all of the left cosets of K and then find all of the right cosets of K in D_3. Any observations?

(b) Write down the group table for D_3, but this time arrange the rows and columns according to the left cosets for K. Color the entire table according to which *left* coset an element belongs to. Can you visualize the observations you made in part (a)?

Exercise 7.4. Consider Q_8. Let $H = \langle i \rangle$ and $K = \langle -1 \rangle$.

(a) Find all of the left cosets of H and all of the right cosets of H in Q_8.

(b) Write down the group table for Q_8 so that rows and columns are arranged according to the left cosets for H. Color the entire table according to which *left* coset an element belongs to.

(c) Find all of the left cosets of K and all of the right cosets of K in Q_8.

(d) Write down the group table for Q_8 so that rows and columns are arranged according to the left cosets for K. Color the entire table according to which *left* coset an element belongs to.

Exercise 7.5. Consider S_4. Find all of the left cosets and all of the right cosets of A_4 in S_4. Instead of doing brute-force, try to be clever. *Hint:* What happens when you compose two even permutations versus an even permutation and an odd permutation?

Exercise 7.6. Consider $\mathbb{Z}_8$. Find all of the left cosets and all of the right cosets of $\langle 4 \rangle$ in $\mathbb{Z}_8$. Why do you know the left and right cosets are the same without actually verifying?

Exercise 7.7. Consider $(\mathbb{Z}, +)$. Find all of the left cosets and all of the right cosets of $3\mathbb{Z}$ in $\mathbb{Z}$. Why do you know the left and right cosets are the same without actually verifying?

Now that we've played with a few examples, let's make a few general observations.

Theorem 7.8. Let G be a group and let $H \leq G$.

1. If $a \in G$, then $a \in aH$ (respectively, Ha).
2. If $b \in aH$ (respectively, Ha), then $aH = bH$ (respectively, $Ha = Hb$).
3. If $a \in H$, then $aH = H = Ha$.
4. If $a \notin H$, then for all $h \in H$, $ah \notin H$ (respectively, $ha \notin H$).

The upshot of part 2 of Theorem 7.8 is that cosets can have different names. In particular, if b is an element of the left coset aH, then we could have just as easily called the coset by the name bH. In this case, both a and b are called **coset representatives**.

In all of the examples we've seen so far, the left and right cosets partitioned G into equal-sized chunks. We need to prove that this is true in general. To prove that the cosets form a partition, we'll define an appropriate equivalence relation.

Theorem 7.9. Let G be a group and let $H \leq G$. Define $\sim_L$ and $\sim_R$ via

$$a \sim_L b \text{ iff } a^{-1}b \in H$$

and

$$a \sim_R b \text{ iff } ab^{-1} \in H.$$

Then both $\sim_L$ and $\sim_R$ are equivalence relations.*

Problem 7.10. If $[a]_{\sim_L}$ (respectively, $[a]_{\sim_R}$) denotes the equivalence class of a under $\sim_L$ (respectively, $\sim_R$), what is $[a]_{\sim_L}$ (respectively, $[a]_{\sim_R}$)? *Hint:* It's got something to do with cosets.

Corollary 7.11. If G is a group and $H \leq G$, then the left (respectively, right) cosets of H form a partition of G.

Next, we argue that all of the cosets have the same size.

Theorem 7.12. Let G be a group, $H \leq G$, and $a \in G$. Define $\phi : H \to aH$ via $\phi(h) = ah$. Then ϕ is one-to-one and onto.

Corollary 7.13. Let G be a group and let $H \leq G$. Then all of the left and right cosets of H are the same size as H. In other words $\#(aH) = |H| = \#(Ha)$ for all $a \in G$.†

7.2 Lagrange's Theorem

We're finally ready to state Lagrange's Theorem, which is named after the Italian born mathematician Joseph Louis Lagrange. It turns out that Lagrange did not actually prove the theorem that is named after him. The theorem was actually proved by Carl Friedrich Gauss in 1801.

Theorem 7.14 (Lagrange's Theorem)**.** Let G be a finite group and let $H \leq G$. Then $|H|$ divides $|G|$.

This simple sounding theorem is extremely powerful. One consequence is that groups and subgroups have a fairly rigid structure. Suppose G is a finite group and let $H \leq G$. Since G is finite, there must be a finite number of distinct left cosets, say $H, a_2H, \ldots, a_nH$. Corollary 7.13 tells us that each of these cosets is the same size. In particular, Lagrange's Theorem implies that for each $i \in \{1, \ldots, n\}$, $|a_iH| = |G|/n$, or equivalently $n = |G|/|a_iH|$. This is depicted in Figure 7.2, where each rectangle represents a coset and we've labeled a single coset representative in each case.

One important consequence of Lagrange's Theorem is that it narrows down the possible sizes for subgroups.

Exercise 7.15. Suppose G is a group of order 48. What are the possible orders for subgroups of G?

*You only need to prove that either $\sim_L$ or $\sim_R$ is an equivalence relation as the proof for the other is similar.

†As you probably expect, $\#(aH)$ denotes the size of aH. Note that everything works out just fine even if H has infinite order.

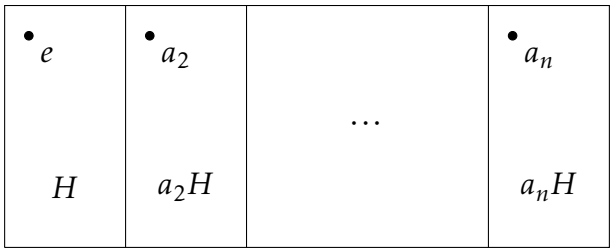


Figure 7.2

Lagrange's Theorem tells us what the possible orders of a subgroup are, but if k is a divisor of the order of a group, it does not guarantee that there is a subgroup of order k. It's not too hard to show that the converse of Lagrange's Theorem is true for cyclic groups. However, it's not true, in general. The next problem will show that A_4 is an example of a group where the converse of Lagrange's Theorem fails. Can you think of others?

Problem 7.16. Consider the alternating group A_4. Lagrange's Theorem tells us that the possible orders of subgroups for A_4 are 1, 2, 3, 4, 6, and 12.

(a) Find examples of subgroups of A_4 of orders 1, 2, 3, 4, and 12.

(b) Write down all of the elements of order 2 in A_4.

(c) Argue that any subgroup of A_4 that contains any two elements of order 2 must contain a subgroup isomorphic to V_4.

(d) Argue that if A_4 has a subgroup of order 6, that it cannot be isomorphic to R_6.

(e) It turns out that up to isomorphism, there are only two groups of order 6, namely S_3 and R_6. Suppose that H is a subgroup of A_4 of order 6. Part (d) guarantees that $H \cong S_3$. Argue that H must contain all of the elements of order 2 from A_4.

(f) Explain why A_4 cannot have a subgroup of order 6.

Using Lagrange's Theorem, we can quickly prove both of the following theorems.

Theorem 7.17. Let G be a finite group and let $a \in G$. Then $|a|$ divides $|G|$.

Theorem 7.18. Every group of prime order is cyclic.

Since the converse of Lagrange's Theorem is not true, the converse of Theorem 7.17 is not true either. However, it is much easier to find a counterexample.

Problem 7.19. Argue that S_4 does not have any elements of order 8.

Lagrange's Theorem motivates the following definition.

Definition 7.20. Let G be a group and let $H \leq G$. The **index** of H in G is the number of cosets (left or right) of H in G. Equivalently, if G is finite, then the index of H in G is equal to $|G|/|H|$. We denote the index via $[G : H]$.

Exercise 7.21. Let $H = \langle (1,2)(3,4), (1,3)(2,4) \rangle$.

(a) Find $[A_4 : H]$.

(b) Find $[S_4 : H]$.

Exercise 7.22. Find $[\mathbb{Z} : 4\mathbb{Z}]$.

7.3 Normal Subgroups

We've seen an example where the left and right cosets of a subgroup were different and a few examples where they coincided. In the latter case, the subgroup has a special name.

Definition 7.23. Let G be a group and let $H \leq G$. If $aH = Ha$ for all $a \in G$, then we say that H is a **normal subgroup**. If H is a normal subgroup of G, then we write $H \trianglelefteq G$.

Exercise 7.24. Provide an example of group that has a subgroup that is not normal.

Problem 7.25. Suppose G is a finite group and let $H \leq G$. If $H \trianglelefteq G$ and we arrange the rows and columns of the group table for G according to the left cosets of H and then color the corresponding cosets, what property will the table have? Is the converse true? That is, if the table has the property you discovered, will H be normal in G?

There are a few instances where we can guarantee that a subgroup will be normal.

Theorem 7.26. Suppose G is a group. Then $\{e\} \trianglelefteq G$ and $G \trianglelefteq G$.

Theorem 7.27. If G is an abelian group, then all subgroups of G are normal.

A group does not have to be abelian in order for all the proper subgroups to be normal.

Problem 7.28. Argue that all of the proper subgroups of Q_8 are normal in Q_8.

Theorem 7.29. Suppose G is a group and let $H \leq G$ such that $[G : H] = 2$. Then $H \trianglelefteq G$.

It turns out that normality is not transitive.

Problem 7.30. Consider $\langle s \rangle = \{e, s\}$ and $\langle r^2, sr^2 \rangle = \{e, r^2, sr^2, s\}$. It is clear that

$$\langle s \rangle \leq \langle r^2, sr^2 \rangle \leq D_4.$$

Show that $\langle s \rangle \trianglelefteq \langle r^2, sr^2 \rangle$ and $\langle r^2, sr^2 \rangle \trianglelefteq D_4$, but $\langle s \rangle \not\trianglelefteq D_4$.

The previous problem illustrates that $H \trianglelefteq K \trianglelefteq G$ does not imply $H \trianglelefteq G$.

Theorem 7.31. Suppose G is a group and let $H \leq G$. Then $H \trianglelefteq G$ if and only if $aHa^{-1} = H$ for all $a \in G$, where

$$aHa^{-1} = \{aha^{-1} \mid h \in H\}.$$

Note that the expression gHg^{-1} is called the **conjugate** of H by g. Another way of thinking about normal subgroups is that they are "closed under conjugation." The previous theorem is often used as the definition of normal. It also motivates the following definition.

Definition 7.32. Let G be a group and let $H \leq G$. The **normalizer of** H **in** G is defined via

$$N_G(H) := \{g \in G \mid gHg^{-1} = H\}.$$

Theorem 7.33. If G is a group and $H \leq G$, then $N_G(H)$ is a subgroup of G.

Theorem 7.34. If G is a group and $H \leq G$, then $H \trianglelefteq N_G(H)$. Moreover, $N_G(H)$ is the largest subgroup of G in which H is normal.

It is worth pointing out that the "smallest" $N_G(H)$ can be is H itself—certainly a subgroup is a normal subgroup of itself. Also, the "largest" that $N_G(H)$ can be is G, which happens precisely when H is normal in G.

Exercise 7.35. Find $N_{D_4}(V_4)$.

Exercise 7.36. Find $N_{D_3}(\langle s \rangle)$.

We conclude this chapter with a few remarks. We've seen examples of groups that have subgroups that are normal and subgroups that are not normal. In an abelian group, all the subgroups are normal. It turns out that there are examples of groups that have no normal subgroups. These groups are called **simple groups**. The smallest simple group is A_5, which has 120 elements and lots of subgroups, none of which are normal.

The classification of the finite simple groups is a theorem stating that every finite simple group belongs to one of four categories:

1. A cyclic group with prime order;
2. An alternating group of degree at least 5;
3. A simple group of Lie type, including both
   (a) the classical Lie groups, namely the simple groups related to the projective special linear, unitary, symplectic, or orthogonal transformations over a finite field;
   (b) the exceptional and twisted groups of Lie type (including the Tits group);
4. The 26 sporadic simple groups.

These groups can be seen as the basic building blocks of all finite groups, in a way reminiscent of the way the prime numbers are the basic building blocks of the natural numbers.

The classification theorem has applications in many branches of mathematics, as questions about the structure of finite groups (and their action on other mathematical objects)

can sometimes be reduced to questions about finite simple groups. Thanks to the classification theorem, such questions can sometimes be answered by checking each family of simple groups and each sporadic group. The proof of the theorem consists of tens of thousands of pages in several hundred journal articles written by about 100 authors, published mostly between 1955 and 2004.

The classification of the finite simple groups is a modern achievement in abstract algebra and I highly encourage you to go learn more about it. You might be especially interested in learning about one of the sporadic groups called the **Monster Group**.

Chapter 8

Products and Quotients of Groups

8.1 Products of Groups

In this section, we will discuss a method for using existing groups as building blocks to form new groups.

Suppose $(G,*)$ and $(H,\circ)$ are two groups. Recall that the Cartesian product of G and H is defined to be

$$G \times H = \{(g,h) : g \in G, h \in H\}$$

For more information on Cartesian products, see Definition A.26. Using the binary operations for the groups G and H, we can define a binary operation on the set $G \times H$. Define $\star$ on $G \times H$ via

$$(g_1,h_1) \star (g_2,h_2) = (g_1 * g_2, h_1 \circ h_2).$$

This looks fancier than it is. We're just doing the operation of each group in the appropriate component. It turns out that $(G \times H, \star)$ is a group.

Theorem 8.1. Suppose $(G,*)$ and $(H,\circ)$ are two groups, where e and e' are the identity elements of G and H, respectively. Then $(G \times H, \star)$ is a group, where $\star$ is defined as above. Moreover, (e,e') is the identity of $G \times H$ and the inverse of $(g,h) \in G \times H$ is given by $(g,h)^{-1} = (g^{-1}, h^{-1})$.

We refer to $G \times H$ as the **direct product** of the groups G and H. Note that we abbreviate $(g_1,h_1) \star (g_2,h_2) = (g_1 * g_2, h_1 \circ h_2)$ by $(g_1,h_1)(g_2,h_2) = (g_1 g_2, h_1 h_2)$.

There's no reason we can't do this for more than two groups. If $A_1, A_2, \ldots, A_n$ is a collection of sets, we define

$$\prod_{i=1}^{n} A_i := A_1 \times A_2 \times \cdots \times A_n.$$

Each element of $\prod_{i=1}^n A_i$ is of the form $(a_1, a_2, \ldots, a_n)$, where $a_i \in A_i$.

Theorem 8.2. Let $G_1, G_2, \ldots, G_n$ be groups. For $(a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \prod_{i=1}^n G_i$, define

$$(a_1, a_2, \ldots, a_n)(b_1, b_2, \ldots, b_n) = (a_1 b_1, a_2 b_2, \ldots, a_n b_n).$$

Then $\prod_{i=1}^n G_i$, the **direct product** of G_i, is a group under this binary operation.

Note that each G_i above is called a **factor** of the direct product. One way to think about direct products is that we can navigate the product by navigating each factor simultaneously but independently.

Theorem 8.3. Let $G_1, G_2, \ldots, G_n$ be finite groups. Then

$$|G_1 \times G_2 \times \cdots \times G_n| = |G_1| \cdot |G_2| \cdots |G_n|.$$

Theorem 8.4. Let $G_1, G_2, \ldots, G_n$ be groups. Then $|G_1 \times G_2 \times \cdots \times G_n|$ is infinite if and only if at least one $|G_i|$ is infinite.

The following theorem should be clear.

Theorem 8.5. Let $G_1, G_2, \ldots, G_n$ be groups. Then $\prod_{i=1}^{n} G_i$ is abelian if and only if each G_i is abelian.

If each G_i is abelian, then we may use additive notation. For example, consider $\mathbb{Z}_2 \times \mathbb{Z}_3$ under the operation of addition mod 2 in the first component and addition mod 3 in the second component. Then

$$\mathbb{Z}_2 \times \mathbb{Z}_3 = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2)\}.$$

Since $\mathbb{Z}_2$ and $\mathbb{Z}_3$ are cyclic, both groups are abelian, and hence $\mathbb{Z}_2 \times \mathbb{Z}_3$ is abelian. In this case, we will use addition notation in $\mathbb{Z}_2 \times \mathbb{Z}_3$. For example,

$$(0,1) + (1,2) = (1,0)$$

and

$$(1,2) + (0,2) = (1,1).$$

There is a very natural generating set for $\mathbb{Z}_2 \times \mathbb{Z}_3$, namely, $\{(1,0), (0,1)\}$ since $1 \in \mathbb{Z}_2$ and $1 \in \mathbb{Z}_3$ generate $\mathbb{Z}_2$ and $\mathbb{Z}_3$, respectively.

Exercise 8.6. Draw the Cayley diagram for $\mathbb{Z}_2 \times \mathbb{Z}_3$ using $\{(1,0), (0,1)\}$ as the generating set. Do you see a subgroup of $\mathbb{Z}_2 \times \mathbb{Z}_3$ isomorphic to $\mathbb{Z}_2$ in the Cayley diagram? What is this subgroup? How about a subgroup isomorphic to $\mathbb{Z}_3$?

Exercise 8.7. Prove that $\mathbb{Z}_2 \times \mathbb{Z}_3$ is a cyclic group of order 6 and hence isomorphic to R_6.

Let's play with a few more examples.

Exercise 8.8. Consider $\mathbb{Z}_2 \times \mathbb{Z}_2$ under the operation of addition mod 2 in each component. Find a generating set for $\mathbb{Z}_2 \times \mathbb{Z}_2$ and then create a Cayley diagram for this group. What well-known group is $\mathbb{Z}_2 \times \mathbb{Z}_2$ isomorphic to?

Consider the similarities and differences between $\mathbb{Z}_2 \times \mathbb{Z}_3$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$. Both groups are abelian by Theorem 8.5, but only the former is cyclic. Here's another exercise.

Problem 8.9. Consider $\mathbb{Z}_2 \times \mathbb{Z}_4$ under the operation of addition mod 2 in the first component and addition mod 4 in the second component.

(a) Using $\{(1,0),(0,1)\}$ as the generating set, draw the Cayley diagram for $\mathbb{Z}_2 \times \mathbb{Z}_4$.

(b) Draw the subgroup lattice for $\mathbb{Z}_2 \times \mathbb{Z}_4$.

(c) Show that $\mathbb{Z}_2 \times \mathbb{Z}_4$ is abelian but not cyclic.

(d) Argue that $\mathbb{Z}_2 \times \mathbb{Z}_4$ cannot be isomorphic to any of D_4, R_8, and Q_8.

The upshot of the previous problem is that there are at least 4 groups of order 8 up to isomorphism. We'll show later that there are actually (at least) 5. The previous exercises have hinted at the following theorem.

Theorem 8.10. The group $\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic if and only if m and n are relatively prime.

Corollary 8.11. The group $\mathbb{Z}_m \times \mathbb{Z}_n$ is isomorphic to $\mathbb{Z}_{mn}$ if and only if m and n are relatively prime.

The previous results can be extended to more than two factors.

Theorem 8.12. The group $\prod_{i=1}^{n} \mathbb{Z}_{m_i}$ is cyclic and isomorphic to $\mathbb{Z}_{m_1 m_2 \cdots m_n}$ if and only if every pair from the collection $\{m_1, m_2, \ldots, m_n\}$ is relatively prime.

Exercise 8.13. Determine whether each of the following groups is cyclic.

(a) $\mathbb{Z}_7 \times \mathbb{Z}_8$

(b) $\mathbb{Z}_7 \times \mathbb{Z}_7$

(c) $\mathbb{Z}_2 \times \mathbb{Z}_7 \times \mathbb{Z}_8$

(d) $\mathbb{Z}_5 \times \mathbb{Z}_7 \times \mathbb{Z}_8$

Theorem 8.14. Suppose $n = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$, where each p_i is a distinct prime number. Then

$$\mathbb{Z}_n \cong \mathbb{Z}_{p_1^{n_1}} \times \mathbb{Z}_{p_2^{n_2}} \times \cdots \times \mathbb{Z}_{p_r^{n_r}}.$$

Theorem 8.15. Suppose G and H are two groups. Then $G \times H \cong H \times G$.

The next theorem tells us how to compute the order of an element in a direct product of groups.

Theorem 8.16. Suppose $G_1, G_2, \ldots, G_n$ are groups and let $(g_1, g_2, \ldots, g_n) \in \prod_{i=1}^{n} G_i$. If $|g_i| = r_i < \infty$, then $|(g_1, g_2, \ldots, g_n)| = \operatorname{lcm}(r_1, r_2, \ldots, r_n)$.

Exercise 8.17. Find the order of each of the following elements.

(a) $(6,5) \in \mathbb{Z}_{12} \times \mathbb{Z}_7$.

(b) $(r, i) \in D_3 \times Q_8$.

(c) $((1,2)(3,4),3) \in S_4 \times \mathbb{Z}_{15}$.

Exercise 8.18. Find the largest possible order in each of the following groups.

(a) $\mathbb{Z}_6 \times \mathbb{Z}_8$

(b) $\mathbb{Z}_9 \times \mathbb{Z}_{12}$

(c) $\mathbb{Z}_4 \times \mathbb{Z}_{18} \times \mathbb{Z}_{15}$

Theorem 8.19. Suppose G_1 and G_2 are groups such that $H_1 \leq G_1$ and $H_2 \leq G_2$. Then $H_1 \times H_2 \leq G_1 \times G_2$.

However, not every subgroup of a direct product has the form above.

Problem 8.20. Find an example that illustrates that not every subgroup of a direct product is the direct product of subgroups of the factors.

Theorem 8.21. Suppose G_1 and G_2 are groups with identities e_1 and e_2, respectively. Then $\{e_1\} \times G_2 \trianglelefteq G_1 \times G_2$ and $G_1 \times \{e_2\} \trianglelefteq G_1 \times G_2$.

Theorem 8.22. Suppose G_1 and G_2 are groups with identities e_1 and e_2, respectively. Then $\{e_1\} \times G_2 \cong G_2$ and $G_1 \times \{e_2\} \cong G_1$.

The next theorem describes precisely the structure of finite abelian groups. We will omit its proof, but allow ourselves to utilize it as needed.

Theorem 8.23 (Fundamental Theorem of Finitely Generated Abelian Groups)**.** Every finitely generated abelian group G is isomorphic to a direct product of cyclic groups of the form

$$\mathbb{Z}_{p_1^{n_1}} \times \mathbb{Z}_{p_2^{n_2}} \times \cdots \times \mathbb{Z}_{p_r^{n_r}} \times \mathbb{Z}^k,$$

where each p_i is a prime number (not necessarily distinct). The product is unique up to rearrangement of the factors.

Note that the number k is called the **Betti number**. A finitely generated abelian group is finite if and only if the Betti number is 0.

Exercise 8.24. Find all abelian groups up to isomorphism of order 8. How many different groups up to isomorphism (both abelian and non-abelian) have we seen and what are they?

Exercise 8.25. Find all abelian groups up to isomorphism for each of the following orders.

(a) 16

(b) 12

(c) 25

(d) 30

(e) 60

8.2 Quotients of Groups

In the previous section, we discussed a method for constructing "larger" groups from "smaller" groups using a direct product construction. In this section, we will in some sense do the opposite.

Problem 7.25 hinted that if $H \leq G$ and we arrange the group table according to the left cosets of H, then the group table will have checkerboard pattern if and only if H is normal in G (i.e., the left and right cosets of H are the same). For example, see the colored table prior to Exercise 7.3 versus the ones you created in Exercises 7.3, 7.4. If we have the checkerboard pattern in the group table that arises from a normal subgroup, then by "gluing together" the colored blocks, we obtain a group table for a smaller group that has the cosets as the elements.

For example, let's consider $K = \langle -1 \rangle \leq Q_8$. Exercise 7.4 showed us that K is normal Q_8. The left (and right) cosets of K in Q_8 are

$$K = \{1, -1\}, iK = \{i, -i\}, jK = \{j, -j\}, \text{ and } kK = \{k, -k\}.$$

As you found in Exercise 7.4, if we arrange the rows and columns of Q_8 according to these cosets, we obtain the following group table.

$*$	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	k	$-k$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

If we consider the 2×2 blocks as elements, it appears that we have a group table for a group with 4 elements. Closer inspection reveals that this looks like the table for V_4. If the table of 2×2 blocks is going to represent a group, we need to understand the binary operation. How do we "multiply" cosets? For example, the table suggest that the coset jK (colored in red) times the coset iK (colored in blue) is equal to kK (colored in purple) despite the fact that $ji = -k \neq k$. Yet, it is true that the product $ji = -k$ is an element in the coset kK. In fact, if we look closely at the table, we see that if we pick any two cosets, the product of any element of the first coset times any element of the second coset will always result in an element in the same coset regardless of which representatives we chose.

In other words, it looks like we can multiply cosets by choosing any representative from each coset and then seeing what coset the product of the representatives lies in. However, it is important to point out that this will only work if we have a checkerboard pattern of cosets, which we have seen evidence of only happening when the corresponding subgroup is normal.

Before continuing, let's continue tinkering with the same example. Consider the Cayley diagram for Q_8 with generators $\{i, j, -1\}$ that is given in Figure 8.1.

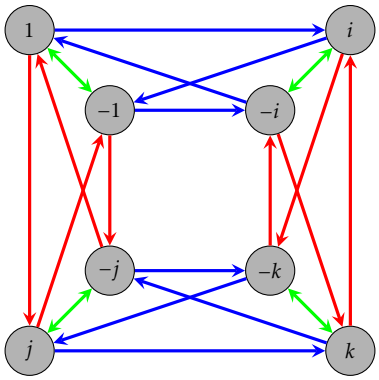


Figure 8.1. Cayley diagram for Q_8 with generating set $\{i, j, -1\}$.

We can visualize the left cosets of K as the clumps of vertices connected together with the two-way green arrows. In this case, we are also seeing the right cosets since K is normal in Q_8. If we collapse the cosets onto each other and collapse the corresponding arrows, we obtain the diagram given in Figure 8.2. It is clear that this diagram is the Cayley diagram for a group that is isomorphic to V_4. For reasons we will understand shortly, this processing of collapsing a Cayley diagram according to the cosets of a normal subgroup is called the "quotient process."

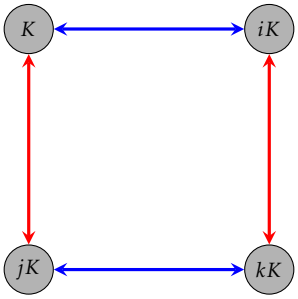


Figure 8.2. The collapsed Cayley diagram for Q_8 according to the cosets of $K = \langle -1 \rangle$.

Exercise 8.26. Let's see what happens if we attempt the quotient process for a subgroup that is not normal. Consider $H = \langle s \rangle \leq D_3$. In Exercise 7.2, we discovered that the left cosets of H are not the same as the right cosets of H. This implies that H is not normal in D_3. Consider the standard Cayley diagram for D_3 that uses the generators r and s. Draw the diagram that results from attempting the quotient process on D_3 using the subgroup H. Explain why this diagram cannot be the diagram for a group.

The problem that arises in Exercise 8.26 is that if the same arrow types (i.e., those representing the same generator) leaving a coset do not point at elements in the same coset, attempting the quotient process will result in a diagram that violates Rule 3 (every action is deterministic) of Definition 2.14. In Figure 8.3, we illustrate what goes wrong if all the arrows out of a coset do not unanimously point to the same coset. In the second subfigure, all the arrows point to the same coset, and in this case, it appears that everything works out just fine.

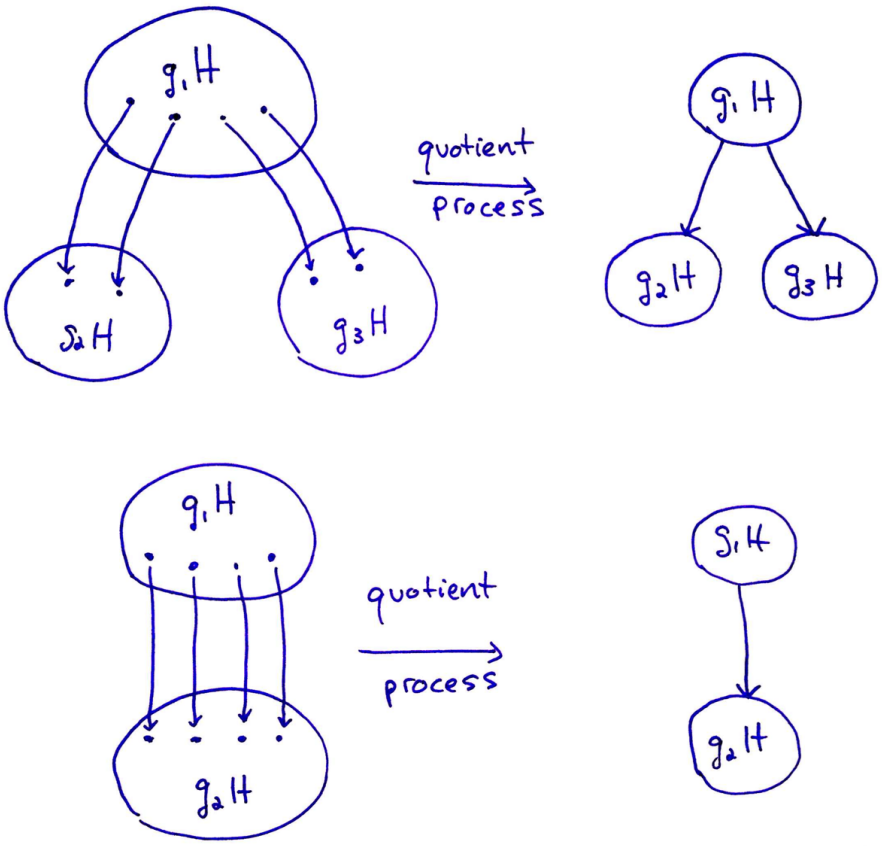


Figure 8.3. The quotient process.

Exercise 8.27. In Exercise 7.3, we learned that the subgroup $K = \langle r \rangle$ is normal in D_3 since the left cosets are equal to the right cosets. Note that this follows immediately from Theorem 7.29 since $[D_3 : K] = 2$. Draw the diagram that results from performing the quotient process to D_3 using the subgroup K. Does the resulting diagram represent a group? If so, what group is it isomorphic to?

Now, suppose G is an arbitrary group and let $H \leq G$. Consider the set of left cosets of H. We define

$$(aH)(bH) := (ab)H.$$

The natural question to ask is whether this operation is well-defined. That is, does the result of multiplying two left cosets depend on our choice of representatives? More specifically, suppose $c \in aH$ and $d \in bH$. Then $cH = aH$ and $dH = bH$. According to the operation defined above, $(cH)(dH) = cdH$. It better be the case that $cdH = abH$, otherwise the operation is not well-defined.

Exercise 8.28. Let $H = \langle s \rangle \leq D_3$. Find specific examples of $a, b, c, d \in D_3$ such that

$$(aH)(bH) \neq (cH)(dH)$$

even though $aH = cH$ and $bH = dH$.

Theorem 8.29. Let G be a group and let $H \leq G$. Then left coset multiplication (as defined above) is well-defined if and only if $H \trianglelefteq G$.

Theorem 8.30. Let G be a group and let $H \trianglelefteq G$. Then the set of left cosets of H in G forms a group under left coset multiplication.

The group from Theorem 8.30 is denoted by G/H, read "G mod H", and is referred to as the **quotient group** (or **factor group**) of G by H. If G is a finite group, then G/H is exactly the group that arises from "gluing together" the colored blocks in a checkerboard-patterned group table. It's also the group that we get after applying the quotient process to the Cayley diagram. It's important to point out once more that this only works properly if H is a normal subgroup.

Theorem 8.31. Let G be a group and let $H \trianglelefteq G$. Then $|G/H| = [G : H]$. In particular, if G is finite, then $|G/H| = |G|/|H|$.

Exercise 8.32. Find the order of the given element in the quotient group. You may assume that we are taking the quotient by a normal subgroup.

(a) $s\langle r\rangle \in D_4/\langle r\rangle$

(b) $j\langle -1\rangle \in Q_8/\langle -1\rangle$

(c) $5 + \langle 4\rangle \in \mathbb{Z}_{12}/\langle 4\rangle$

(d) $(2,1) + \langle (1,1)\rangle \in (\mathbb{Z}_3 \times \mathbb{Z}_6)/\langle (1,1)\rangle$

Exercise 8.33. For each quotient group below, describe the group. If possible, state what group each is isomorphic to. You may assume that we are taking the quotient by a normal subgroup.

(a) $Q_8/\langle -1\rangle$

(b) $Q_8/\langle i\rangle$

(c) $\mathbb{Z}_4/\langle 2\rangle$

(d) $V_4/\langle h\rangle$

(e) $A_4/\langle (1,2)(3,4), (1,3)(2,4)\rangle$

(f) $(\mathbb{Z}_2 \times \mathbb{Z}_2)/\langle (1,1)\rangle$

(g) $\mathbb{Z}/4\mathbb{Z}$

(h) S_4/A_4

(i) $(\mathbb{Z}_4 \times \mathbb{Z}_2)/(\{0\} \times \mathbb{Z}_2)$

Theorem 8.34. Let G be a group. Then

1. $G/\{e\} \cong G$

2. $G/G \cong \{e\}$

Theorem 8.35. For all $n \in \mathbb{N}$, we have the following.

1. $S_n/A_n \cong \mathbb{Z}_2$ (for $n \geq 3$)
2. $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$
3. $\mathbb{R}/n\mathbb{R} \cong \{e\}$

Theorem 8.36. Let G be a group and let $H \trianglelefteq G$. If G is abelian, then so is G/H.

Problem 8.37. Show that the converse of the previous theorem is not true by providing a specific counterexample.

Exercise 8.38. Consider the quotient group $(\mathbb{Z}_4 \times \mathbb{Z}_6)/\langle(0,1)\rangle$.

(a) What is the order of $(\mathbb{Z}_4 \times \mathbb{Z}_6)/\langle(0,1)\rangle$?

(b) Is the group abelian? Why?

(c) Write down all the elements of $(\mathbb{Z}_4 \times \mathbb{Z}_6)/\langle(0,1)\rangle$.

(d) Does one of the elements generate the group?

(e) What well-known group is $(\mathbb{Z}_4 \times \mathbb{Z}_6)/\langle(0,1)\rangle$ isomorphic to?

Theorem 8.39. Let G be a group and let $H \trianglelefteq G$. If G is cyclic, then so is G/H.

Problem 8.40. Show that the converse of the previous theorem is not true by providing a specific counterexample.

Here are few additional exercises. These ones are a bit tougher.

Exercise 8.41. For each quotient group below, describe the group. If possible, state what group each is isomorphic to. You may assume that we are taking the quotient by a normal subgroup.

(a) $(\mathbb{Z}_4 \times \mathbb{Z}_6)/\langle(0,2)\rangle$

(b) $(\mathbb{Z} \times \mathbb{Z})/\langle(1,1)\rangle$

(c) $\mathbb{Q}/\langle 1\rangle$ (the operation on $\mathbb{Q}$ is addition)

Chapter 9

Homomorphisms and the Isomorphism Theorems

9.1 Homomorphisms

Let G_1 and G_2 be groups. Recall that $\phi : G_1 \to G_2$ is an isomorphism iff ϕ

(a) is one-to-one,

(b) is onto, and

(c) satisfies the homomorphic property.

We say that G_1 is isomorphic to G_2 and write $G_1 \cong G_2$ if such a ϕ exists. Loosely speaking, two groups are isomorphic if they have the "same structure." What if we drop the one-to-one and onto requirement?

Definition 9.1. Let $(G_1, *)$ and $(G_2, \circ)$ be groups. A function $\phi : G_1 \to G_2$ is a **homomorphism** iff ϕ satisfies the homomorphic property:

$$\phi(x * y) = \phi(x) \circ \phi(y)$$

for all $x, y \in G_1$. At the risk of introducing ambiguity, we will usually omit making explicit reference to the binary operations and write the homomorphic property as

$$\phi(xy) = \phi(x)\phi(y).$$

Group homomorphisms are analogous to linear transformations on vector spaces that one encounters in linear algebra.

Figure 9.1 captures a visual representation of the homomorphic property. We encountered this same representation in Figure 5.6. If $\phi(x) = x'$, $\phi(y) = y'$, and $\phi(z) = z'$ while $z' = x' \circ y'$, then the only way G_2 may respect the structure of G_1 is for

$$\phi(x * y) = \phi(z) = z' = x' \circ y' = \phi(x) \circ \phi(y).$$

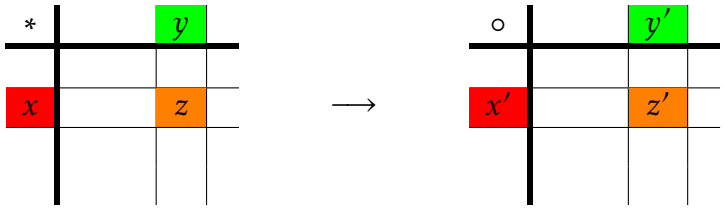


Figure 9.1

Exercise 9.2. Define $\phi : \mathbb{Z}_3 \to D_3$ via $\phi(k) = r^k$. Prove that ϕ is a homomorphism and then determine whether ϕ is one-to-one or onto. Also, try to draw a picture of the homomorphism in terms of Cayley diagrams.

Exercise 9.3. Let G and H be groups. Prove that the function $\phi : G \times H \to G$ given by $\phi(g,h) = g$ is a homomorphism. This function is an example of a **projection map**.

There is always at least one homomorphism between two groups.

Theorem 9.4. Let G_1 and G_2 be groups. Define $\phi : G_1 \to G_2$ via $\phi(g) = e_2$ (where e_2 is the identity of G_2). Then ϕ is a homomorphism. This function is often referred to as the **trivial homomorphism** or the 0-**map**.

Back in Section 5.5, we encountered several theorems about isomorphisms. However, at the end of that section we remarked that some of those theorems did not require that the function be one-to-one and onto. We collect those results here for convenience.

Theorem 9.5. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism.

1. If e_1 and e_2 are the identity elements of G_1 and G_2, respectively, then $\phi(e_1) = e_2$.
2. For all $g \in G_1$, we have $\phi(g^{-1}) = [\phi(g)]^{-1}$.
3. If $H \leq G_1$, then $\phi(H) \leq G_2$, where

$$\phi(H) := \{y \in G_2 \mid \text{there exists } h \in H \text{ such that } \phi(h) = y\}.$$

Note that $\phi(H)$ is called the **image** of H. A special case is when $H = G_1$. Notice that ϕ is onto exactly when $\phi(G_1) = G_2$.

The next two theorems tell us that under a homomorphism, the order of the image must divide the order of the preimage.

Theorem 9.6. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. If G_1 is finite, then $|\phi(G_1)|$ divides $|G_1|$.

Theorem 9.7. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. If $g \in G_1$ such that $|g|$ is finite, then $|\phi(g)|$ divides $|g|$.

Every homomorphism has an important subset of the domain associated with it.

Definition 9.8. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. The **kernel** of ϕ is defined via

$$\ker(\phi) := \{g \in G_1 \mid \phi(g) = e_2\}.$$

The kernel of a homomorphism is analogous to the null space of a linear transformation of vector spaces.

Exercise 9.9. Identify the kernel and image for the homomorphism given in Exercise 9.2.

Exercise 9.10. What is the kernel of a trivial homomorphism (see Theorem 9.4).

Theorem 9.11. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. Then $\ker(\phi) \trianglelefteq G_1$.

It turns out that the kernel can tell us something about whether ϕ is one-to-one.

Theorem 9.12. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. Then ϕ is one-to-one iff $\ker(\phi) = \{e_1\}$.

Remark 9.13. Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. Given a generating set for G_1, the homomorphism ϕ is uniquely determined by its action on the generating set for G_1. In particular, if you have a word for a group element written in terms of the generators, just apply the homomorphic property to the word to find the image of the corresponding group element.

Exercise 9.14. Suppose $\phi : Q_8 \to V_4$ is a group homomorphism satisfying $\phi(i) = h$ and $\phi(j) = v$.

(a) Find $\phi(1)$, $\phi(-1)$, $\phi(k)$, $\phi(-i)$, $\phi(-j)$, and $\phi(-k)$.

(b) Find $\ker(\phi)$.

(c) What well-known group is $Q_8/\ker(\phi)$ isomorphic to?

Exercise 9.15. Find a non-trivial homomorphism from $\mathbb{Z}_{10}$ to $\mathbb{Z}_6$.

Exercise 9.16. Find all non-trivial homomorphisms from $\mathbb{Z}_3$ to $\mathbb{Z}_6$.

Problem 9.17. Prove that the only homomorphism from D_3 to $\mathbb{Z}_3$ is the trivial homomorphism.

Exercise 9.18. Let F be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$ and let D be the subset of differentiable functions on $\mathbb{R}$. It turns out that F is a group under addition of functions and D is a subgroup of F (you do not need to prove this). Define $\phi : D \to F$ via $\phi(f) = f'$ (where f' is the derivative of f). Prove that ϕ is a homomorphism. You may recall facts from calculus without proving them. Is ϕ one-to-one? Onto?

9.2 The Isomorphism Theorems

We begin with a theorem.

Theorem 9.19. Let G be a group and let $H \trianglelefteq G$. Then the map $\gamma : G \to G/H$ given by $\gamma(g) = gH$ is a homomorphism with $\ker(\gamma) = H$. This map is called the **canonical projection map**.

The upshot of Theorems 9.11 and 9.19 is that kernels of homomorphisms are always normal and every normal subgroup is the kernel of some homomorphism.

The next theorem is arguably the crowning achievement of the course.

Theorem 9.20 (The First Isomorphism Theorem)**.** Let G_1 and G_2 be groups and suppose $\phi : G_1 \to G_2$ is a homomorphism. Then

$$G_1/\ker(\phi) \cong \phi(G_1).$$

If ϕ is onto, then

$$G_1/\ker(\phi) \cong G_2.$$

Exercise 9.21. Let $\phi : Q_8 \to V_4$ be the homomorphism described in Exercise 9.14. Use the First Isomorphism Theorem to prove that $Q_8/\langle -1 \rangle \cong V_4$.

Exercise 9.22. Define $\phi : S_n \to \mathbb{Z}_2$ via

$$\phi(\sigma) = \begin{cases} 0, & \sigma \text{ even} \\ 1, & \sigma \text{ odd.} \end{cases}$$

Use the First Isomorphism Theorem to prove that $S_n/A_n \cong \mathbb{Z}_2$.

Exercise 9.23. Use the First Isomorphism Theorem to prove that $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}_6$. Attempt to draw a picture of this using Cayley diagrams.

Exercise 9.24. Use the First Isomorphism Theorem to prove that $(\mathbb{Z}_4 \times \mathbb{Z}_2)/(\{0\} \times \mathbb{Z}_2) \cong \mathbb{Z}_4$.

We finish the chapter by listing a few of the remaining isomorphism theorems, but we won't prove these in this course.

Theorem 9.25 (The Second Isomorphism Theorem)**.** Let G be a group with $H \leq G$ and $N \trianglelefteq G$. Then

1. $HN := \{hn \mid h \in H, n \in N\} \leq G$;
2. $H \cap N \trianglelefteq H$;
3. $H/H \cap N \cong HN/N$.

Theorem 9.26 (The Third Isomorphism Theorem)**.** Let G be a group with $H, K \trianglelefteq G$ and $K \leq H$. Then

$$G/H \cong (G/K)/(H/K).$$

Chapter 10

An Introduction to Rings

10.1 Definitions and Examples

Recall that a group is a set together with a single binary operation, which together satisfy a few modest properties. Loosely speaking, a ring is a set together with two binary operations (called addition and multiplication) that are related via a distributive property.

Definition 10.1. A **ring** R is a set together with two binary operations $+$ and $\cdot$ (called **addition** and **multiplication**, respectively) satisfying the following:

(i) $(R, +)$ is an abelian group.

(ii) $\cdot$ is associative: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$.

(iii) The **distributive property** holds: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$ for all $a, b, c \in R$.

Remark 10.2. We make a couple comments about notation.

(a) We often write ab in place $a \cdot b$.

(b) The additive inverse of the ring element $a \in R$ is denoted $-a$.

Theorem 10.3. Let R be a ring. Then for all $a, b \in R$:

1. $0a = a0 = 0$
2. $(-a)b = a(-b) = -(ab)$
3. $(-a)(-b) = ab$

Definition 10.4. A ring R is called **commutative** if multiplication is commutative.

Definition 10.5. A ring R is said to have an **identity** (or called a **ring with 1**) if there is an element $1 \in R$ such that $1a = a1 = a$ for all $a \in R$.

Exercise 10.6. Justify that $\mathbb{Z}$ is a commutative ring with 1 under the usual operations of addition and multiplication. Which elements have multiplicative inverses in $\mathbb{Z}$?

Exercise 10.7. Justify that $\mathbb{Z}_n$ is a commutative ring with 1 under addition and multiplication mod n.

Exercise 10.8. Consider the set $\mathbb{Z}_{10} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Which elements have multiplicative inverses in $\mathbb{Z}_{10}$?

Exercise 10.9. For each of the following, find a positive integer n such that the ring $\mathbb{Z}_n$ does not have the stated property.

(a) $a^2 = a$ implies $a = 0$ or $a = 1$.

(b) $ab = 0$ implies $a = 0$ or $b = 0$.

(c) $ab = ac$ and $a \neq 0$ imply $b = c$.

Theorem 10.10. If R is a ring with 1, then the multiplicative identity is unique and $-a = (-1)a$.

Problem 10.11. Requiring $(R, +)$ to be a group is fairly natural, but why require $(R, +)$ to be abelian? Suppose R has a 1. Compute $(1 + 1)(a + b)$ in two different ways.

Definition 10.12. A ring R with 1 (with $1 \neq 0$) is called a **division ring** if every nonzero element in R has a multiplicative inverse: if $a \in R \setminus \{0\}$, then there exists $b \in R$ such that $ab = ba = 1$.

Definition 10.13. A commutative division ring is called a **field**.

Definition 10.14. A nonzero element a in a ring R is called a **zero divisor** if there is a nonzero element $b \in R$ such that either $ab = 0$ or $ba = 0$.

Exercise 10.15. Are there any zero divisors in $\mathbb{Z}_{10}$? If so, find all of them.

Exercise 10.16. Are there any zero divisors in $\mathbb{Z}_5$? If so, find all of them.

Exercise 10.17. Provide an example of a ring R and elements $a, b \in R$ such that $ax = b$ has more than one solution. How does this compare with groups?

Theorem 10.18 (Cancellation Law)**.** Assume $a, b, c \in R$ such that a is not a zero divisor. If $ab = ac$, then either $a = 0$ or $b = c$.

Definition 10.19. Assume R is a ring with 1 with $1 \neq 0$. An element $u \in R$ is called a **unit** in R if u has a multiplicative inverse (i.e., there exists $v \in R$ such that $uv = vu = 1$. The set of units in R is denoted $U(R)$.

Exercise 10.20. Consider the ring $\mathbb{Z}_{20}$.

(a) Find $U(\mathbb{Z}_{20})$.

(b) Find the zero divisors of $\mathbb{Z}_{20}$.

(c) Any observations?

Theorem 10.21. If $U(R) \neq \emptyset$, then $U(R)$ forms a group under multiplication.

Remark 10.22. We make a few observations.

(a) A field is a commutative ring F with identity $1 \neq 0$ in which every nonzero element is a unit, i.e., $U(F) = F \setminus \{0\}$.

(b) Zero divisors can never be units.

(c) Fields never have zero divisors.

Definition 10.23. A commutative ring with identity $1 \neq 0$ is called an **integral domain** if it has no zero divisors.

Remark 10.24. The Cancellation Law (Theorem 10.18) holds in integral domains for any three elements.

Theorem 10.25. Any finite integral domain is a field.

Example 10.26. Here are a few examples. Details left as an exercise.

(a) **Zero Ring**: If $R = \{0\}$, we can turn R into a ring in the obvious way. The zero ring is a finite commutative ring with 1. It is the only ring where the additive and multiplicative identities are equal. The zero ring is not a division ring, not a field, and not an integral domain.

(b) **Trivial Ring:** Given any abelian group R, we can turn R into a ring by defining multiplication via $ab = 0$ for all $a, b \in R$. Trivial rings are commutative rings in which every nonzero element is a zero divisor. Hence a trivial ring is not a division ring, not a field, and not a integral domain.

(c) The integers form an integral domain, but $\mathbb{Z}$ is not a division ring, and hence not a field.

(d) The rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$ are fields under the usual operations of addition and multiplication.

(e) The group of units $U(\mathbb{Z}_n)$ is the set of elements in $\mathbb{Z}_n$ that are relatively prime to n. All other nonzero elements are zero divisors. It turns out that $\mathbb{Z}_n$ forms a finite field iff n is prime.

(f) The set of even integers $2\mathbb{Z}$ forms a commutative ring under the usual operations of addition and multiplication. However, $2\mathbb{Z}$ does not have a 1, and hence cannot be a division ring nor a field nor an integral domain.

(g) **Polynomial Ring:** Fix a commutative ring R. Let $R[x]$ denote the set of polynomials in the variable x with coefficients in R. Then $R[x]$ is a commutative ring with 1. The units of $R[x]$ are exactly the units of R (if there are any). So, $R[x]$ is never a division ring nor a field. However, if R is an integral domain, then so is $R[x]$.

(h) **Matrix Ring:** Fix a ring R and let n be a positive integer. Let $M_n(R)$ be the set of $n \times n$ matrices with entries from R. Then $M_n(R)$ forms a ring under ordinary matrix addition and multiplication. If R is nontrivial and $n \geq 2$, then $M_n(R)$ always has zero divisors and $M_n(R)$ is not commutative even if R is. If R has a 1, then the matrix with 1's down the diagonal and 0's elsewhere is the multiplicative identity in $M_n(R)$. In this case, the group of units is the set of invertible $n \times n$ matrices, denoted $GL_n(R)$ and called the **general linear group of degree** n **over** R.

(i) **Quadratic Field:** Define $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. It turns out that $\mathbb{Q}(\sqrt{2})$ is a field. In fact, we can replace 2 with any rational number that is not a perfect square in $\mathbb{Q}$.

(j) **Hamilton Quaternions:** Define $\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}, i, j, k \in Q_8\}$ Then $\mathbb{H}$ forms a ring, where addition is definite componentwise in i, j, and k and multiplication is defined by expanding products and the simplifying using the relations of Q_8. It turns out that $\mathbb{H}$ is a non-commutative ring with 1.

Exercise 10.27. Find an example of a ring R and an element $a \in R \setminus \{0\}$ such that a is neither a zero divisor nor a unit.

Definition 10.28. A **subring** of a ring R is a subgroup of R that is closed under multiplication.

Remark 10.29. The property "is a subring" is clearly transitive. To show that a subset S of a ring R is a subring, it suffices to show that $S \neq \emptyset$, S is closed under subtraction, and S is closed under multiplication.

Example 10.30. Here are a few quick examples.

(a) $\mathbb{Z}$ is a subring of $\mathbb{Q}$, which is a subring of $\mathbb{R}$, which in turn is a subring of $\mathbb{C}$.

(b) $2\mathbb{Z}$ is a subring of $\mathbb{Z}$.

(c) The set $\mathbb{Z}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{Q}(\sqrt{2})$.

(d) The ring R is a subring of $R[x]$ if we identify R with set of constant functions.

(e) The set of polynomials with zero constant term in $R[x]$ is a subring of $R[x]$.

(f) $\mathbb{Z}[x]$ is a subring of $\mathbb{Q}[x]$.

(g) $\mathbb{Z}_n$ is *not* a subring of $\mathbb{Z}$ as the operations are different.

Problem 10.31. Consider the ring $\mathbb{Z}_{10}$ from Exercise 10.8. Let $S = \{0, 2, 4, 6, 8\}$.

(a) Argue that S is a subring of R.

(b) Is S a ring with 1? If so, find the multiplicative identity. If not, explain why.

(c) Is S a field? Justify your answer.

Problem 10.32. Suppose R is a ring and let $a \in R$. Define $S = \{x \in R \mid ax = 0\}$. Prove that S is a subring of R.

Problem 10.33. Consider the ring $\mathbb{Z}$. It turns out that $2\mathbb{Z}$ and $3\mathbb{Z}$ are subrings (but you don't need to prove this). Determine whether $2\mathbb{Z} \cup 3\mathbb{Z}$ is a subring of $\mathbb{Z}$. Justify your answer.

10.2 Ring Homomorphisms

Definition 10.34. Let R and S be rings. A **ring homomorphism** is a map $\phi : R \to S$ satisfying

(i) $\phi(a+b) = \phi(a) + \phi(b)$

(ii) $\phi(ab) = \phi(a)\phi(b)$

for all $a, b \in R$. The **kernel** of ϕ is defined via $\ker(\phi) = \{a \in R \mid \phi(a) = 0\}$. If ϕ is a bijection, then ϕ is called an **isomorphism**, in which case, we say that R and S are **isomorphic rings** and write $R \cong S$.

Example 10.35.

(a) For $n \in \mathbb{Z}$, define $\phi_n : \mathbb{Z} \to \mathbb{Z}$ via $\phi_n(x) = nx$. We see that $\phi_n(x+y) = n(x+y) = nx+ny = \phi_n(x)+\phi_n(y)$. However, $\phi_n(xy) = n(xy)$ while $\phi_n(x)\phi_n(y) = (nx)(ny) = n^2xy$. It follows that ϕ_n is a ring homomorphism exactly when $n \in \{0, 1\}$.

(b) Define $\phi : \mathbb{Q}[x] \to \mathbb{Q}$ via $\phi(p(x)) = p(0)$ (called **evaluation at 0**). It turns out that ϕ is a ring homomorphism, where $\ker(\phi)$ is the set of polynomials with 0 constant term.

Exercise 10.36. For each of the following, determine whether the given function is a ring homomorphism. Justify your answers.

(a) Define $\phi : \mathbb{Z}_4 \to \mathbb{Z}_{12}$ via $\phi(x) = 3x$.

(b) Define $\phi : \mathbb{Z}_{10} \to \mathbb{Z}_{10}$ via $\phi(x) = 5x$.

(c) Let $S = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$. Define $\phi : \mathbb{C} \to S$ via $\phi(a+ib) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

(d) Let $T = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid a, b \in \mathbb{Z} \right\}$. Define $\phi : T \to \mathbb{Z}$ via $\phi\left(\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\right) = a$.

Theorem 10.37. Let $\phi : R \to S$ be a ring homomorphism.

1. $\phi(R)$ is a subring of S.
2. $\ker(\phi)$ is a subring of R.

Problem 10.38. Suppose $\phi : R \to S$ is a ring homomorphism such that R is a ring with 1, call it 1_R. Prove that $\phi(1_R)$ is the multiplicative identity in $\phi(R)$ (which is a subring of S). Can you think of an example of a ring homomorphism where S has a multiplicative identity that is not equal to $\phi(1_R)$?

Theorem 10.37(2) states that the kernel of a ring homomorphism is a subring. This is analogous to the kernel of a group homomorphism being a subgroup. However, recall that the kernel of a group homomorphism is also a normal subgroup. Like the situation with groups, we can say something even stronger about the kernel of a ring homomorphism. This will lead us to the notion of an **ideal**.

Theorem 10.39. Let $\phi : R \to S$ be a ring homomorphism. If $a \in \ker(\phi)$ and $r \in R$, then $ar, ra \in \ker(\phi)$. That is, $\ker(\phi)$ is closed under multiplication by elements of R.

10.3 Ideals and Quotient Rings

Recall that in the case of a homomorphism ϕ of groups, the cosets of $\ker(\phi)$ have the structure of a group (that happens to be isomorphic to the image of ϕ by the First Isomorphism Theorem). In this case, $\ker(\phi)$ is the identity of the associated quotient group. Moreover, recall that every kernel is a normal subgroup of the domain and every normal subgroup can be realized as the kernel of some group homomorphism. Can we do the same sort of thing for rings?

Let $\phi : R \to S$ be a ring homomorphism with $\ker(\phi) = I$. Note that ϕ is also a group homomorphism of abelian groups and the cosets of $\ker(\phi)$ are of the form $r + I$. More specifically, if $\phi(r) = a$, then $\phi^{-1}(a) = r + I$.

These cosets naturally have the structure of a ring isomorphic to the image of ϕ:

$$(r + I) + (s + I) = (r + s) + I \tag{10.1}$$

$$(r + I)(s + I) = (rs) + I \tag{10.2}$$

The reason for this is that if $\phi^{-1}(a) = X$ and $\phi^{-1}(b) = Y$, then the inverse image of $a + b$ and ab are $X + Y$ and XY, respectively.

The corresponding ring of cosets is called the **quotient ring** of R by $I = \ker(\phi)$ and is denoted by R/I. The additive structure of the quotient ring R/I is exactly the additive quotient group of the additive abelian group R by the normal subgroup R (all subgroups are normal in abelian groups). When I is the kernel of some ring homomorphism ϕ, the additive abelian quotient group R/I also has a multiplicative structure defined in (2) above, making R/I into a ring.

Can we make R/I into a ring for any subring I?

The answer is "no" in general, just like in the situation with groups. But perhaps this isn't obvious because if I is an arbitrary subring of R, then I is necessarily an additive subgroup of the abelian group R, which implies that I is an additive normal subgroup of the group R. It turns out that the multiplicative structure of R/I may not be well-defined if I is an arbitrary subring.

Let I be an arbitrary *subgroup* of the additive group R. Let $r+I$ and $s+I$ be two arbitrary cosets. In order for multiplication of the cosets to be well-defined, the product of the two cosets must be independent of choice of representatives. Let $r+\alpha$ and $s+\beta$ be arbitrary representatives of $r+I$ and $s+I$, respectively ($\alpha, \beta \in I$), so that $r+I = (r+\alpha)+I$ and $s+I = (s+\beta)+I$. We must have

$$(r+\alpha)(s+\beta)+I = rs+I. \tag{10.3}$$

This needs to be true for all possible choices of $r, s \in R$ and $\alpha, \beta \in I$. In particular, it must be true when $r = s = 0$. In this case, we must have

$$\alpha\beta + I = I. \tag{10.4}$$

But this only happens when $\alpha\beta \in I$. That is, one requirement for multiplication of cosets to be well-defined is that I must be closed under multiplication, making I a *subring*.

Next, if we let $s = 0$ and let r be arbitrary, we see that we must have $r\beta \in I$ for every $r \in R$ and every $\beta \in I$. That is, it must be the case that I is closed under multiplication on the left by elements from R. Similarly, letting $r = 0$, we can conclude that we must have I closed under multiplication on the right by elements from R.

On the other hand, if I is closed under multiplication on the left and on the right by elements from R, then it is clear that relation (4) above is satisfied.

It is easy to verify that if the multiplication of cosets defined in (2) above is well-defined, then this multiplication makes the additive quotient group R/I into a ring (just check the axioms for being a ring).

We have shown that the quotient R/I of the ring R by a subgroup I has a natural ring structure iff I is closed under multiplication on the left and right by elements of R (which also forces I to be a subring). Such subrings are called **ideals**.

Definition 10.40. Let R be a ring and let I be a subset of R.

(a) I is a **left ideal** (respectively, **right ideal**) of R iff I is a subring and $rI \subseteq I$ (respectively, $Ir \subseteq I$) for all $r \in R$.

(b) I is an **ideal** (or **two-sided ideal**) iff I is both a left and a right ideal.

Here's a summary of everything that just happened.

Theorem 10.41. Let R be a ring and let I be an ideal of R. Then the additive quotient group R/I is a ring under the binary operations:

$$(r+I)+(s+I) = (r+s)+I \tag{10.5}$$

$$(r+I)(s+I) = (rs)+I \tag{10.6}$$

for all $r, s \in R$. Conversely, if I is any subgroup such that the above operations are well-defined, then I is an ideal of R.

Theorem 10.42. Suppose I and J are ideals of the ring R. Then $I \cap J$ is an ideal of R.

As you might expect, we have some isomorphism theorems.

Theorem 10.43 (First Isomorphism Theorem for Rings)**.** If $\phi : R \to S$ is a ring homomorphism, then $\ker(\phi)$ is an ideal of R and $R/\ker(\phi) \cong \phi(R)$.

We also have the expected Second and Third Isomorphism Theorems for rings.

The next theorem tells us that a subring is an ideal iff it is a kernel of a ring homomorphism.

Theorem 10.44. If I is any ideal of R, then the **natural projection** $\pi : R \to R/I$ defined via $\pi(r) = r + I$ is a surjective ring homomorphism with $\ker(\pi) = I$.

For the remainder of this section, assume that R is a ring with identity $1 \neq 0$.

Definition 10.45. Let A be any subset of R. Let (A) denote the smallest ideal of R containing A, called the **ideal generated by** A. If A consists of a single element, say $A = \{a\}$, then $(a) := (\{a\})$ is called a **principal ideal**.

Remark 10.46. The following facts are easily verified.

(a) (A) is the intersection of all ideals containing A.

(b) If R is commutative, then $(a) = aR := \{ar \mid r \in R\}$.

Example 10.47. In $\mathbb{Z}$, $n\mathbb{Z} = (n) = (-n)$. In fact, these are the only ideals in $\mathbb{Z}$ (since these are the only subgroups). So, all the ideals in $\mathbb{Z}$ are principal. If m and n are positive integers, then $n\mathbb{Z} \subseteq m\mathbb{Z}$ iff m divides n. Moreover, we have $(m, n) = (d)$, where d is the greatest common divisor of m and n.

Problem 10.48. Consider the ideal $(2, x)$ in $\mathbb{Z}[x]$. Note that $(2, x) = \{2p(x)+xq(x) \mid p(x), q(x) \in \mathbb{Z}[x]\}$. Argue that $(2, x)$ is not a principal ideal, i.e., there is no single polynomial in $\mathbb{Z}[x]$ that we can use to generate $(2, x)$.

Theorem 10.49. Assume R is a commutative ring with $1 \neq 0$. Let I be an ideal of R. Then $I = R$ iff I contains a unit.

Theorem 10.50. Assume R is a commutative ring with $1 \neq 0$. Then R is a field iff its only ideals are 0 and R.

Loosely speaking, the previous results say that fields are "like simple groups" (i.e, groups with no non-trivial normal subgroups).

Corollary 10.51. If R is a field, then every nonzero ring homomorphism from R into another ring is an injection.

10.4 Maximal and Prime Ideals

Throughout this entire section, we assume that all rings have a multiplicative identity $1 \neq 0$.

In this section of notes, we will study two important classes of ideals, namely **maximal** and **prime** ideals, and study the relationship between them.

Definition 10.52. Assume R is a commutative ring with 1. An ideal M in a ring R is called a **maximal ideal** if $M \neq R$ and the only ideals containing M are M and R.

Example 10.53. Here are a few examples. Checking the details is left as an exercise.

(1) In $\mathbb{Z}$, all the ideals are of the form $n\mathbb{Z}$ for $n \in \mathbb{Z}^+$. The maximal ideals correspond to the ideals $p\mathbb{Z}$, where p is prime.

(2) Consider the integral domain $\mathbb{Z}[x]$. The ideals (x) (i.e., the subring containing polynomials with 0 constant term) and (2) (i.e, the set of polynomials with even coefficients) are not maximal since both are contained in the proper ideal $(2, x)$. However, as we shall see soon, $(2, x)$ is maximal in $\mathbb{Z}[x]$.

(3) The zero ring has no maximal ideals.

(4) Consider the abelian group $\mathbb{Q}$ under addition. We can turn $\mathbb{Q}$ into a trivial ring by defining $ab = 0$ for all $a, b \in \mathbb{Q}$. In this case, the ideals are exactly the additive subgroups of $\mathbb{Q}$. However, $\mathbb{Q}$ has no maximal subgroups, and so $\mathbb{Q}$ has no maximal ideals.

The next result states that rings with an identity $1 \neq 0$ always have maximal ideals. It turns out that we won't need this result going forward, so we'll skip its proof. However, it is worth noting that all known proofs make use of Zorn's Lemma (equivalent to the Axiom of Choice), which is also true for the proofs that a finitely generated group has maximal subgroups or that every vector spaces has a basis.

Theorem 10.54. In a ring with 1, every proper ideal is contained in a maximal ideal.

For commutative rings, there is a very nice characterization about maximal ideals in terms of the structure of their quotient rings.

Theorem 10.55. Assume R is a commutative ring with 1. Then M is a maximal ideal iff the quotient ring R/M is a field.

Example 10.56. We can use the previous theorem to verify whether an ideal is maximal.

(1) Recall that $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$ and that $\mathbb{Z}_n$ is a field iff n is prime. We can conclude that $n\mathbb{Z}$ is a maximal ideal precisely when n is prime.

(2) Define $\phi : \mathbb{Z}[x] \to \mathbb{Z}$ via $\phi(p(x)) = p(0)$. Then ϕ is surjective and $\ker(\phi) = (x)$. By the First Isomorphism Theorem for Rings, we see that $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$. However, $\mathbb{Z}$ is not a field. Hence (x) is not maximal in $\mathbb{Z}[x]$. Now, define $\psi : \mathbb{Z} \to \mathbb{Z}_2$ via $\psi(x) = x \mod 2$ and consider the composite homomorphism $\psi \circ \phi : \mathbb{Z} \to \mathbb{Z}_2$. It is clear that $\psi \circ \phi$ is onto and the kernel of $\psi \circ \phi$ is given by $\{p(x) \in \mathbb{Z}[x] \mid p(0) \in 2\mathbb{Z}\} = (2, x)$. Again by the First Isomorphism Theorem for Rings, $\mathbb{Z}[x]/(2, x) \cong \mathbb{Z}_2$. Since $\mathbb{Z}_2$ is a field, $(2, x)$ is a maximal ideal.

Definition 10.57. Assume R is a commutative ring with 1. An ideal P is called a **prime ideal** if $P \neq R$ and whenever the product $ab \in P$ for $a, b \in R$, then at least one of a or b is in P.

Example 10.58. In any integral domain, the 0 ideal (0) is a prime ideal. What if the ring is not an integral domain?

Remark 10.59. The notion of a prime ideal is a generalization of "prime" in $\mathbb{Z}$. Suppose $n \in \mathbb{Z}^+ \setminus \{1\}$ such that n divides ab. In this case, n is guaranteed to divide either a or b exactly when n is prime. Now, let $n\mathbb{Z}$ be a proper ideal in $\mathbb{Z}$ with $n > 1$ and suppose $ab \in \mathbb{Z}$ for $a, b \in \mathbb{Z}$. In order for $n\mathbb{Z}$ to be a prime ideal, it must be true that n divides either a or b. However, this is only guaranteed to be true for all $a, b \in \mathbb{Z}$ when p is prime. That is, the nonzero prime ideals of $\mathbb{Z}$ are of the form $p\mathbb{Z}$, where p is prime. Note that in the case of the integers, the maximal and nonzero prime ideals are the same.

Theorem 10.60. Assume R is a commutative ring with 1. Then P is a prime ideal in R iff the quotient ring R/P is an integral domain.

Corollary 10.61. Assume R is a commutative ring with 1. Every maximal ideal of R is a prime ideal.

Example 10.62. Recall that $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$. Since $\mathbb{Z}$ is an integral domain, it must be the case that (x) is a prime ideal in $\mathbb{Z}[x]$. However, as we saw in an earlier example, (x) is not maximal in $\mathbb{Z}[x]$ since $\mathbb{Z}$ is not a field. This shows that the converse of the previous corollary is not true.

Appendix A

Prerequisites

I'll organize this section better later, but for now, here's a brain dump of some concepts you should be familiar with.

A.1 Basic Set Theory

Definition A.1. A **set** is a collection of objects called **elements**. If A is a set and x is an element of A, we write $x \in A$. Otherwise, we write $x \notin A$.

Definition A.2. The set containing no elements is called the **empty set**, and is denoted by the symbol $\emptyset$.

If we think of a set as a box containing some stuff, then the empty set is a box with nothing in it.

Definition A.3 (Interval Notation)**.** For $a, b \in \mathbb{R}$ with $a < b$, we define the following.

1. $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$
2. $(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$
3. $(-\infty, b) = \{x \in \mathbb{R} \mid x < b\}$
4. $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$

We analogously define $[a, b)$, $(a, b]$, $[a, \infty)$, and $(-\infty, b]$.

Remark A.4. There are a few sets with common names that we should be familiar with.

1. **Natural Numbers:** $\mathbb{N} = \{1, 2, 3, \ldots\}$
2. **Integers:** $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
3. **Real Numbers:** $\mathbb{R} = (-\infty, \infty)$*

*This is really a cop out. If you look at the definition of the interval $(-\infty, \infty)$, we are being circular.

4. **Complex Numbers:** $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$, where $i = \sqrt{-1}$ is the imaginary unit.

Definition A.5. The language associated to sets is specific. We will often define sets using the following notation, called **set builder notation**.

$$S = \{x \in A \mid x \text{ satisfies some condition}\}$$

The first part "$x \in A$" denotes what type of x is being considered. The statements to the right of the colon are the conditions that x must satisfy in order to be members of the set. This notation is read as "The set of all x in A such that x satisfies some condition," where "some condition" is something specific about the restrictions on x relative to A.

Definition A.6. If A and B are sets, then we say that A is a **subset** of B, written $A \subseteq B$, provided that every element of A is also an element of B.

Remark A.7. Observe that $A \subseteq B$ is equivalent to "For all x (in the universe of discourse), if $x \in A$, then $x \in B$." Since we know how to deal with "for all" statements and conditional propositions, we know how to go about proving $A \subseteq B$.

Theorem A.8 (Transitivity of subsets)**.** Suppose that A, B, and C are sets. If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

Definition A.9. If $A \subseteq B$, then A is called a **proper subset** provided that $A \neq B$. In this case, we may write $A \subset B$ or $A \subsetneq B$.[†]

Definition A.10. Let A and B be sets.

1. The **union** of the sets A and B is $A \cup B = \{x \in U \mid x \in A \text{ or } x \in B\}$.
2. The **intersection** of the sets A and B is $A \cap B = \{x \in U \mid x \in A \text{ and } x \in B\}$.
3. The **set difference** of the sets A and B is $A \setminus B = \{x \in U \mid x \in A \text{ and } x \notin B\}$.
4. The **complement of** A (relative to U) is the set $A^c = U \setminus A = \{x \in U \mid x \notin A\}$.

Definition A.11. If two sets A and B have the property that $A \cap B = \emptyset$, then we say that A and B are **disjoint** sets.

Theorem A.12. Let A and B be sets. If $A \subseteq B$, then $B^c \subseteq A^c$.

Definition A.13. Two sets A and B are **equal** if and only if $A \subseteq B$ and $B \subseteq A$. In this case we write $A = B$.

Remark A.14. Given two sets A and B, if we want to prove $A = B$, then we have to do two separate "mini" proofs: one for $A \subseteq B$ and one for $B \subseteq A$.

Theorem A.15. Let A and B be sets. Then $A \setminus B = A \cap B^c$.

Theorem A.16 (DeMorgan's Law)**.** Let A and B be sets. Then

[†] *Warning:* Some books use $\subset$ to mean $\subseteq$.

1. $(A \cup B)^c = A^c \cap B^c$,
2. $(A \cap B)^c = A^c \cup B^c$.

Theorem A.17 (Distribution of Union and Intersection)**.** Let A, B, and C be sets. Then

1. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
2. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Definition A.18. Suppose we have a collection $\{A_\alpha\}_{\alpha \in \Delta}$.

1. The **union of the entire collection** is defined via
$$\bigcup_{\alpha \in \Delta} A_\alpha = \{x \mid x \in A_\alpha \text{ for some } \alpha \in \Delta\}.$$
2. The **intersection of the entire collection** is defined via
$$\bigcap_{\alpha \in \Delta} A_\alpha = \{x \mid x \in A_\alpha \text{ for all } \alpha \in \Delta\}.$$

Example A.19. In the special case that $\Delta = \mathbb{N}$, we write

$$\bigcup_{n=1}^{\infty} A_n = \{x \mid x \in A_n \text{ for some } n \in \mathbb{N}\} = A_1 \cup A_2 \cup A_3 \cup \cdots$$

and

$$\bigcap_{n=1}^{\infty} A_n = \{x \mid x \in A_n \text{ for all } n \in \mathbb{N}\} = A_1 \cap A_2 \cap A_3 \cap \cdots$$

Similarly, if $\Delta = \{1, 2, 3, 4\}$, then

$$\bigcup_{n=1}^{4} A_n = A_1 \cup A_2 \cup A_3 \cup A_4$$

and

$$\bigcap_{n=1}^{4} A_n = A_1 \cap A_2 \cap A_3 \cap A_4.$$

Remark A.20. Notice the difference between "$\bigcup$" and "$\cup$" (respectively, "$\bigcap$" and "$\cap$"). The larger versions of the union and intersection symbols very much like the notation that you've likely seen for sums (e.g., $\sum_{i=1}^{\infty} i^2$).

Definition A.21. We say that a collection of sets $\{A_\alpha\}_{\alpha \in \Delta}$ is **pairwise disjoint** if $A_\alpha \cap A_\beta = \emptyset$ whenever $\alpha \neq \beta$.

Exercise A.22. Draw a Venn diagram of a collection of 3 sets that are pairwise disjoint.

Exercise A.23. Provide an example of a collection of three sets, say $\{A_1, A_2, A_3\}$, such that the collection is *not* pairwise disjoint, but

$$\bigcap_{n=1}^{3} A_n = \emptyset.$$

Definition A.24. An **ordered pair** is an object of the form (x, y). Two ordered pairs (x, y) and (a, b) are **equal** if $x = a$ and $y = b$.

Definition A.25. An n-**tuple** is an object of the form $(x_1, x_2, \ldots, x_n)$. Each x_i is referred to as the ith **component**.

Note that an ordered pair is just a 2-tuple.

Definition A.26. If X and Y are sets, the **Cartesian product** of X and Y is defined by

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

That is, $X \times Y$ is the set of all ordered pairs where the first element is from X and the second element is from Y. The set $X \times X$ is sometimes denoted by X^2. We similarly define the Cartesian product of n sets, say $X_1, \ldots, X_n$, by

$$\prod_{i=1}^{n} X_i = X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n) \mid \text{each } x_i \in X_i\}.$$

Exercise A.27. What general conclusion can you make about $X \times Y$ versus $Y \times X$? When will they be equal?

Exercise A.28. If X and Y are both finite sets, then how many elements will $X \times Y$ have? Be as specific as possible.

Exercise A.29. Let $X = [0, 1]$ and let $Y = \{1\}$. Describe geometrically what $X \times Y$, $Y \times X$, $X \times X$, and $Y \times Y$ look like.

A.2 Relations

Definition A.30. Let X and Y be sets. A **relation** from a set X to a set Y is a subset of $X \times Y$. A relation on X is a subset of $X \times X$.

Remark A.31. Different notations for relations are used in different contexts. When talking about relations in the abstract, we indicate that a pair (a, b) is in the relation by some notation like $a \sim b$, which is read "a is related to b."

Remark A.32. We can often represent relations using graphs or digraphs. Given a finite set X and a relation $\sim$ on X, a **digraph** (short for *directed graph*) is a discrete graph having the members of X as vertices and a directed edge from x to y iff $x \sim y$.

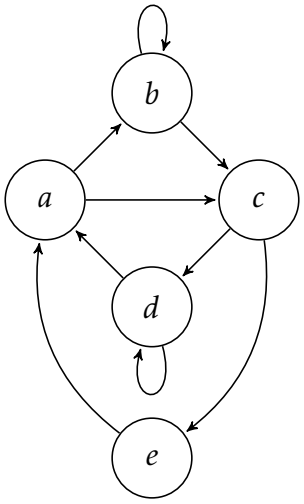


Figure A.1. An example of a digraph for a relation.

Example A.33. Figure A.1 depicts a digraph that represents a relation R given by

$$R = \{(a,b),(a,c),(b,b),(b,c),(c,d),(c,e),(d,d),(d,a),(e,a)\}.$$

Exercise A.34. Let $A = \{a,b,c\}$ and define $\sim = \{(a,a),(a,b),(b,c),(c,b),(c,a)\}$. Draw the digraph for $\sim$.

Definition A.35. Let $\sim$ be a relation on a set A.

1. $\sim$ is **reflexive** if for all $x \in A$, $x \sim x$ (every element is related to itself).
2. $\sim$ is **symmetric** if for all $x, y \in A$, if $x \sim y$, then $y \sim x$.
3. $\sim$ is **transitive** if for all $x, y, z \in A$, if $x \sim y$ and $y \sim z$, then $x \sim z$.

Exercise A.36. Given a finite set A and a relation $\sim$, describe what each of reflexive, symmetric, and transitive look like in terms of a digraph.

Exercise A.37. Let P be the set of people at a party and define N via $(x,y) \in N$ iff x knows the name of y. Describe what it would mean for N to be reflexive, symmetric, and transitive.

Definition A.38. Let $\sim$ be a relation on a set A. Then $\sim$ is called an **equivalence relation** if $\sim$ is reflexive, symmetric, and transitive.

Exercise A.39. Determine which of the following are equivalence relations.

1. Let P_f denote the set of all people with accounts on Facebook. Define F via xFy iff x is friends with y.
2. Let P be the set of all people and define H via xHy iff x and y have the same height.

3. Let P be the set of all people and define T via xTy iff x is taller than y.
4. Consider the relation "divides" on $\mathbb{N}$.
5. Let L be the set of lines and define $\|$ via $l_1 \| l_2$ iff l_1 is parallel to l_2.
6. Let $C[0,1]$ be the set of continuous functions on $[0,1]$. Define $f \sim g$ iff
$$\int_0^1 |f(x)|\,dx = \int_0^1 |g(x)|\,dx.$$
7. Define $\sim$ on $\mathbb{N}$ via $n \sim m$ iff $n+m$ is even.
8. Define D on $\mathbb{R}$ via $(x,y) \in D$ iff $x = 2y$.
9. Define $\sim$ on $\mathbb{Z}$ via $a \sim b$ iff $a-b$ is a multiple of 5.
10. Define $\sim$ on $\mathbb{R}^2$ via $(x_1,y_1) \sim (x_2,y_2)$ iff $x_1^2 + y_1^2 = x_2^2 + y_2^2$.
11. Define $\sim$ on $\mathbb{R}$ via $x \sim y$ iff $\lfloor x \rfloor = \lfloor y \rfloor$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x (e.g., $\lfloor \pi \rfloor = 3$, $\lfloor -1.5 \rfloor = -2$, and $\lfloor 4 \rfloor = 4$).
12. Define $\sim$ on $\mathbb{R}$ via $x \sim y$ iff $|x-y| < 1$.

Definition A.40. Let $\sim$ be a relation on a set A (not necessarily an equivalence relation) and let $x \in A$. Then we define the **set of relatives of** x via
$$[x] = \{y \in A \mid x \sim y\}.$$
Also, define
$$\Omega_\sim = \{[x] \mid x \in A\}.$$
Notice that $\Omega_\sim$ is a set of sets. In particular, an element in $\Omega_\sim$ is a subset of A (equivalently, an element of $\mathcal{P}(A)$). Other common notations for $[x]$ include $\overline{x}$ and R_x.

Exercise A.41. Find $[1]$ and $[2]$ for the relation given in part 9 of Exercise A.39. How many different sets of relatives are there? What are they?

Exercise A.42. If $\sim$ is an equivalence relation on a finite set A, then what is the connection between the equivalence classes and the corresponding digraph?

Theorem A.43. Suppose $\sim$ is an equivalence relation on a set A and let $a, b \in A$. Then $[a] = [b]$ iff $a \sim b$.

Theorem A.44. Suppose $\sim$ is an equivalence relation on a set A. Then

1. $\bigcup_{x \in A} [x] = A$, and
2. for all $x, y \in A$, either $[x] = [y]$ or $[x] \cap [y] = \emptyset$.

Definition A.45. In light of Theorem A.44, if $\sim$ is an equivalence relation on a set A, then we refer to each $[x]$ as the **equivalence class** of x. In this case, $\Omega_\sim$ is the set of equivalence classes determined by $\sim$.

Remark A.46. The upshot of Theorem A.44 is that given an equivalence relation, every element lives in exactly one equivalence class. We'll see in the next section of notes that we can run this in reverse. That is, if we separate out the elements of a set so that every element is an element of exactly one subset (like the bins of my kid's toys), then this determines an equivalence relation. More on this later.

A.3 Partitions

Definition A.47. A collection Ω of nonempty subsets of a set A is said to be a **partition** of A if the elements of Ω satisfy:

1. Given $X, Y \in \Omega$, either $X = Y$ or $X \cap Y = \emptyset$ (We can't have both at the same time. Do you see why?), and
2. $\bigcup_{X \in \Omega} X = A$.

That is, the elements of Ω are pairwise disjoint and their union is all of A.

The next theorem spells out half of the close connection between partitions and equivalence relations. Hopefully you were anticipating this.

Theorem A.48. Let $\sim$ be an equivalence relation on a set A. Then $\Omega_\sim$ forms a partition of A.

Exercise A.49. Consider the equivalence relation

$$\sim = \{(1,1),(1,2),(2,1),(2,2),(3,3),(4,4),(4,5),(5,4),(5,5),(6,6),(5,6),(6,5),(4,6),(6,4)\}$$

on the set $A = \{1,2,3,4,5,6\}$. Find the partition determined by $\Omega_\sim$.

It turns out that we can reverse the situation, as well. That is, given a partition, we can form an equivalence relation. Before proving this, we need a definition.

Definition A.50. Let A be a set and Ω any collection of subsets from $\mathcal{P}(A)$ (not necessarily a partition). If $a, b \in A$, we will define a to be Ω-related to b if there exists an $R \in \Omega$ that contains both a and b. This relation is denoted by $\sim_\Omega$ and is called the **relation on** A **associated to** Ω.

Remark A.51. This definition may look more awkward than the actual underlying concept. The idea is that if two elements are in the same subset, then they are related. For example, when my kids pick up all their toys and put them in the appropriate toy bins, we say that two toys are related if they are in the same bin.

Remark A.52. Notice that we have two notations that looks similar: $\Omega_\sim$ and $\sim_\Omega$.

1. $\Omega_\sim$ is the collection of subsets of A determined by the relation $\sim$.
2. $\sim_\Omega$ is the relation determined by the collection of subsets Ω.

Theorem A.53. Let A be a set and let Ω be a partition of A. Then $\sim_\Omega$ is an equivalence relation.

Remark A.54. The previous theorem says that every partition determines a natural equivalence relation. Namely, two elements are related if they are in the same equivalence class.

A.4 Functions

Definition A.55. Let X and Y be two nonempty sets. A **function** from set X to set Y, denoted $f : X \to Y$, is a relation (i.e., subset of $X \times Y$) such that:

1. For each $x \in X$, there exists $y \in Y$ such that $(x, y) \in f$, and
2. If $(x, y_1), (x, y_2) \in f$, then $y_1 = y_2$.

Note that if $(x, y) \in f$, we usually write $y = f(x)$ and say that "f maps x to y."

Remark A.56. Item 1 of Definition A.55 says that every element of X appears in the first coordinate of an ordered pair in the relation. Item 2 says that each element of X only appears once in the first coordinate of an ordered pair in the relation. It is important to note that there are no restrictions on whether an element of Y ever appears in the second coordinate. Furthermore, if an element of B appears in the second coordinate, it may appear again in a different ordered pair.

Definition A.57. The set X from Definition A.55 is called the **domain** of f and is denoted by $\mathrm{Dom}(f)$. The set Y is called the **codomain** of f and is denoted by $\mathrm{Codom}(f)$. The set

$$\mathrm{Rng}(f) = \{y \in Y \mid \text{there exists } x \text{ such that } y = f(x)\}$$

is called the **range** of f or the **image of** X under f.

Remark A.58. It follows immediately from the definition that $\mathrm{Rng}(f) \subseteq \mathrm{Codom}(f)$. However, it is possible that the range of f is strictly smaller.

Remark A.59. If f is a function and $(x, y) \in f$, then we may refer to x as the **input** of f and y as the **output** of f.

Exercise A.60. Let $X = \{\circ, \square, \triangle, \odot\}$ and $Y = \{a, b, c, d, e\}$. Determine whether each of the following represent functions. Explain. If the relation is a function, determine the domain, codomain, and range.

1. $f : X \to Y$ defined via $f = \{(\circ, a), (\square, b), (\triangle, c), (\odot, d)\}$.
2. $g : X \to Y$ defined via $g = \{(\circ, a), (\square, b), (\triangle, c), (\odot, c)\}$.

3. $h : X \to Y$ defined via $h = \{(\circ, a), (\square, b), (\triangle, c), (\circ, d)\}$.
4. $k : X \to Y$ defined via $k = \{(\circ, a), (\square, b), (\triangle, c), (\odot, d), (\square, e)\}$.
5. $l : X \to Y$ defined via $l = \{(\circ, e), (\square, e), (\triangle, e), (\odot, e)\}$.
6. $m : X \to Y$ defined via $m = \{(\circ, a), (\triangle, b), (\odot, c)\}$.
7. happy : $Y \to X$ defined via happy$(y) = \odot$ for all $y \in Y$.
8. id : $X \to X$ defined via $\mathrm{id}(x) = x$ for all $x \in X$.
9. nugget : $X \to X$ defined via

$$\mathrm{nugget}(x) = \begin{cases} x, & \text{if } x \text{ is a geometric shape,} \\ \square, & \text{otherwise.} \end{cases}$$

Exercise A.61. Let $f : X \to Y$ be a function and suppose that X and Y have n and m elements in them, respectively. Also, suppose that $n < m$. Is it possible for $\mathrm{Rng}(f) = \mathrm{Codom}(f)$? Explain.

Exercise A.62. In high school I am sure that you were told that a graph represents a function if it passes the **vertical line test**. Using our terminology of ordered pairs, explain why this works.

Definition A.63. Two functions are equal if they have the same domain, same codomain, and the same set of ordered pairs in the relation.

Remark A.64. If two functions are defined by the same algebraic formula, but have different domains, then they are *not* equal. For example, the function $f : \mathbb{R} \to \mathbb{R}$ defined via $f(x) = x^2$ is not equal to the function $g : \mathbb{N} \to \mathbb{N}$ defined via $g(x) = x^2$.

Theorem A.65. If $f : X \to Y$ and $g : X \to Y$ are functions, then $f = g$ iff $f(x) = g(x)$ for all $x \in X$.

Definition A.66. Let $f : X \to Y$ be a function.

1. The function f is said to be **one-to-one** (or **injective**) if for all $y \in \mathrm{Rng}(f)$, there is a unique $x \in X$ such that $y = f(x)$.
2. The function f is said to be **onto** (or **surjective**) if for all $y \in Y$, there exists $x \in X$ such that $y = f(x)$.
3. If f is both one-to-one and onto, we say that f is a **one-to-one correspondence** (or a **bijection**).

Exercise A.67. Provide an example of each of the following. You may draw a bubble diagram, write down a list of ordered pairs, or write a formula (as long as the domain and codomain are clear). Assume that X and Y are finite sets.

1. A function $f : X \to Y$ that is one-to-one but not onto.
2. A function $f : X \to Y$ that is onto but not one-to-one.
3. A function $f : X \to Y$ that is both one-to-one and onto.
4. A function $f : X \to Y$ that is neither one-to-one nor onto.

Theorem A.68. Let $f : X \to Y$ be a function. Then f is one-to-one iff for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$, then $x_1 = x_2$.

Remark A.69. The previous theorem gives a technique for proving that a given function is one-to-one. Start by assuming that $f(x_1) = f(x_2)$ and then work to show that $x_1 = x_2$.

Remark A.70. To show that a given function is onto, you should start with an arbitrary $y \in \text{Rng}(f)$ and then work to show that there exists $x \in X$ such that $y = f(x)$.

Definition A.71. If $f : X \to Y$ and $g : Y \to Z$ are functions, then a new function $g \circ f : X \to Z$ can be defined by $(g \circ f)(x) = g(f(x))$ for all $x \in \text{Dom}(f)$.

Remark A.72. It is important to notice that the function on the right is the one that "goes first."

Exercise A.73. In each case, give examples of finite sets X, Y, and Z, and functions $f : X \to Y$ and $g : Y \to Z$ that satisfy the given conditions. Drawing bubble diagrams is sufficient.

1. f is onto, but $g \circ f$ is not onto.
2. g is onto, but $g \circ f$ is not onto.
3. f is one-to-one, but $g \circ f$ is not one-to-one.
4. g is one-to-one, but $g \circ f$ is not.

Theorem A.74. If $f : X \to Y$ and $g : Y \to Z$ are both functions that are onto, then $g \circ f$ is also onto.

Theorem A.75. If $f : X \to Y$ and $g : Y \to Z$ are both functions that are one-to-one, then $g \circ f$ is also one-to-one.

Corollary A.76. If $f : X \to Y$ and $g : Y \to Z$ are both one-to-one correspondences, then $g \circ f$ is also a one-to-one correspondence.

Definition A.77. Let $f : X \to Y$ be a function. The relation f^{-1}, called f **inverse**, is defined via

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

Remark A.78. Notice that we called f^{-1} a relation and not a function. In some circumstances f^{-1} will be a function and sometimes it won't be.

Exercise A.79. Provide an example of a function $f : X \to Y$ such that f^{-1} is *not* a function. A bubble diagram is sufficient.

Theorem A.80. Let $f : X \to Y$ be a function. Then f^{-1} is a function iff f is 1-1.

Theorem A.81. Let $f : X \to Y$ be a function and suppose that f^{-1} is a function. Then

1. $(f \circ f^{-1})(x) = x$ for all $x \in Y$, and
2. $(f^{-1} \circ f)(x) = x$ for all $x \in X$.

(You only need to prove one of these statements; the other is similar.)

Theorem A.82. Let $f : X \to Y$ and $g : Y \to X$ be functions such that f is a one-to-one correspondence. If $(f \circ g)(x) = x$ for all $x \in Y$ and $(g \circ f)(x) = x$ for all $x \in X$, then $g = f^{-1}$.

Remark A.83. The upshot of the previous two theorems is that if f^{-1} is a function, then it is the only one satisfying the two-sided "undoing" property exhibited in Theorem A.81.

The next theorem can be considered to be a converse of Theorem A.82.

Theorem A.84. Let $f : X \to Y$ and $g : Y \to X$ be functions satisfying $(f \circ g)(x) = x$ for all $x \in Y$ and $(g \circ f)(x) = x$ for all $x \in X$. Then f is a one-to-one correspondence.

Theorem A.85. Let $f : X \to Y$ and $g : Y \to Z$ be functions. If f and g are both one-to-one correspondences, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

A.5 Induction

Induction is a technique for proving statements of the form "For all $n \in \mathbb{N}$, $P(n)$," where $P(n)$ is some predicate involving n. Notice that this is a statement about natural numbers and not some other set.

Axiom A.86 (Axiom of Induction)**.** Let $S \subseteq \mathbb{N}$ such that both

1. $1 \in S$, and
2. if $k \in S$, then $k + 1 \in S$.

Then $S = \mathbb{N}$.

Remark A.87. Recall that an axiom is a basic mathematical assumption. That is, we are assuming that the Axiom of Induction is true, which I'm hoping that you can agree is a pretty reasonable assumption. I like to think of the first hypothesis of the Axiom of Induction as saying that we have a first rung of a ladder. The second hypothesis says that if we have some random rung, we can always get to the next rung. Taken together, this says that we can get from the first rung to the second, from the second to the third, and so on. Again, we are assuming that the "and so on" works as expected here.

Theorem A.88 (Principle of Mathematical Induction)**.** Let $P(1), P(2), P(3), \ldots$ be a sequence of statements, one for each natural number.[‡] Assume

1. $P(1)$ is true, and
2. If $P(k)$ is true, then $P(k+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbb{N}$.[§]

Remark A.89. The Principal of Mathematical Induction (PMI) provides us with a process for proving statements of the form: "For all $n \in \mathbb{N}$, $P(n)$," where $P(n)$ is some predicate involving n. Hypothesis (1) above is called the **base step** while (2) is called the **inductive step**.

Skeleton Proof A.90 (Proof by induction for $(\forall n \in \mathbb{N})P(n)$)**.** Here is what the general structure for a proof by induction looks like. Remarks are in parentheses.

Proof. We proceed by induction.

(i) Base step: (Verify that $P(1)$ is true. This often, but not always, amounts to plugging $n = 1$ into two sides of some claimed equation and verifying that both sides are actually equal. Don't assume that they are equal!)

(ii) Inductive step: (Your goal is to prove that "For all $k \in \mathbb{N}$, if $P(k)$ is true, then $P(k+1)$ is true.") Let $k \in \mathbb{N}$ and assume that $P(k)$ is true. (Now, do some stuff to show that $P(k+1)$ is true.) Therefore, $P(k+1)$ is true.

Thus, by the PMI, $P(n)$ is true for all $n \in \mathbb{N}$. □

[‡]In this case, you should think of $P(n)$ as a predicate, where $P(1)$ is the statement that corresponds to substituting in the value 1 for n.

[§]*Hint:* Let $S = \{k \in \mathbb{N} \mid P_k \text{ is true}\}$ and use the Axiom of Induction. The set S is sometimes called the *truth set*. Your job is to show that the truth set is all of $\mathbb{N}$.

Appendix B

Elements of Style for Proofs

Years of elementary school math taught us incorrectly that the answer to a math problem is just a single number, "the right answer." It is time to unlearn those lessons; those days are over. From here on out, mathematics is about discovering proofs and writing them clearly and compellingly.

The following rules apply whenever you write a proof. I may refer to them, by number, in my comments on your homework and exams. Keep these rules handy so that you may refer to them as you write your proofs.

1. **The writing process.** Use the same writing process that you would for any writing project.

   (a) Prewriting. This is the most mathematical step of the process. Often this step takes place on scratch paper. Figure out the mathematics: test conjectures, work out examples, try various proof techniques, etc.

   (b) Writing. When you understand the mathematics it is time to write the first draft. The draft may have extraneous information, be missing information, be written in the wrong order, contain some minor mathematical errors, etc.

   (c) Revising. Once you have a first draft, go back and revise the writing. Focus on large changes such as adding, removing, rearranging, and replacing. Fix any mathematical errors.

   (d) Editing/proofreading. At this stage you must attend to the fine details. Fix any problems with spelling, grammar, word choice, punctuation, etc. Make sure all of the mathematics is typeset correctly.

   (e) Publishing. Make the final changes so that you can submit your work. You may need to fit it to a style guide (get the margins correct, add a title page, etc.), convert it to a certain file type, or print it.

2. **The burden of communication lies on you, not on your reader.** It is your job to explain your thoughts; it is not your reader's job to guess them from a few hints. You are trying to convince a skeptical reader who doesn't believe you, so you need to argue with airtight logic in crystal clear language; otherwise the reader will continue

to doubt. If you didn't write something on the paper, then (a) you didn't communicate it, (b) the reader didn't learn it, and (c) the grader has to assume you didn't know it in the first place.

3. **Tell the reader what you're proving.** The reader doesn't necessarily know or remember what "Theorem 2.13" is. Even a professor grading a stack of papers might lose track from time to time. Therefore, the statement you are proving should be on the same page as the beginning of your proof. For an exam this won't be a problem, of course, but on your homework, recopy the claim you are proving. This has the additional advantage that when you study for exams by reviewing your homework, you won't have to flip back in the notes/textbook to know what you were proving.

4. **Use English words.** Although there will usually be equations or mathematical statements in your proofs, use English sentences to connect them and display their logical relationships. If you look in your notes/textbook, you'll see that each proof consists mostly of English words.

5. **Use complete sentences.** If you wrote a history essay in sentence fragments, the reader would not understand what you meant; likewise in mathematics you must use complete sentences, with verbs, to convey your logical train of thought.

   Some complete sentences can be written purely in mathematical symbols, such as equations (e.g., $a^3 = b^{-1}$), inequalities (e.g., $x < 5$), and other relations (like $5 \mid 10$ or $7 \in \mathbb{Z}$). These statements usually express a relationship between two mathematical *objects*, like numbers or sets. However, it is considered bad style to begin a sentence with symbols. A common phrase to use to avoid starting a sentence with mathematical symbols is "We see that..."

6. **Show the logical connections among your sentences.** Use phrases like "Therefore" or "because" or "if. . . , then. . . " or "if and only if" to connect your sentences.

7. **Know the difference between statements and objects.** A mathematical object is a *thing*, a noun, such as a group, an element, a vector space, a number, an ordered pair, etc. Objects either exist or don't exist. Statements, on the other hand, are mathematical *sentences*: they can be true or false.

   When you see or write a cluster of math symbols, be sure you know whether it's an object (e.g., "$x^2 + 3$") or a statement (e.g., "$x^2 + 3 < 7$"). One way to tell is that every mathematical statement includes a verb, such as $=$, $\leq$, "divides", etc.

8. **"=" means equals.** Don't write $A = B$ unless you mean that A actually equals B. This rule seems obvious, but there is a great temptation to be sloppy. In calculus, for example, some people might write $f(x) = x^2 = 2x$ (which is false), when they really mean that "if $f(x) = x^2$, then $f'(x) = 2x$."

9. **Don't interchange = and $\implies$.** The equals sign connects two *objects*, as in "$x^2 = b$"; the symbol "$\implies$" is an abbreviation for "implies" and connects two *statements*, as in "$a + b = a \implies b = 0$." You should avoid using $\implies$ in your formal write-ups.

10. **Say exactly what you mean.** Just as the = is sometimes abused, so too people sometimes write $A \in B$ when they mean $A \subseteq B$, or write $a_{ij} \in A$ when they mean that a_{ij} is an entry in matrix A. Mathematics is a very precise language, and there is a way to say exactly what you mean; find it and use it.

11. **Don't write anything unproven.** Every statement on your paper should be something you *know* to be true. The reader expects your proof to be a series of statements, each proven by the statements that came before it. If you ever need to write something you don't yet know is true, you *must* preface it with words like "assume," "suppose," or "if" (if you are temporarily assuming it), or with words like "we need to show that" or "we claim that" (if it is your goal). Otherwise the reader will think they have missed part of your proof.

12. **Write strings of equalities (or inequalities) in the proper order.** When your reader sees something like
$$A = B \leq C = D,$$
he/she expects to understand easily why $A = B$, why $B \leq C$, and why $C = D$, and he/she expects the *point* of the entire line to be the more complicated fact that $A \leq D$. For example, if you were computing the distance d of the point $(12, 5)$ from the origin, you could write
$$d = \sqrt{12^2 + 5^2} = 13.$$
In this string of equalities, the first equals sign is true by the Pythagorean theorem, the second is just arithmetic, and the *point* is that the first item equals the last item: $d = 13$.

    A common error is to write strings of equations in the wrong order. For example, if you were to write "$\sqrt{12^2 + 5^2} = 13 = d$", your reader would understand the first equals sign, would be baffled as to how we know $d = 13$, and would be utterly perplexed as to why you wanted or needed to go through 13 to prove that $\sqrt{12^2 + 5^2} = d$.

13. **Avoid circularity.** Be sure that no step in your proof makes use of the conclusion!

14. **Don't write the proof backwards.** Beginning students often attempt to write "proofs" like the following, which attempts to prove that $\tan^2(x) = \sec^2(x) - 1$:
$$\begin{aligned}
\tan^2(x) &= \sec^2(x) - 1 \\
\left(\frac{\sin(x)}{\cos(x)}\right)^2 &= \frac{1}{\cos^2(x)} - 1 \\
\frac{\sin^2(x)}{\cos^2(x)} &= \frac{1 - \cos^2(x)}{\cos^2(x)} \\
\sin^2(x) &= 1 - \cos^2(x) \\
\sin^2(x) + \cos^2(x) &= 1 \\
1 &= 1
\end{aligned}$$

Notice what has happened here: the writer *started* with the conclusion, and deduced the true statement "1 = 1." In other words, he/she has proved "If $\tan^2(x) = \sec^2(x) - 1$, then $1 = 1$," which is true but highly uninteresting.

Now this isn't a bad way of *finding* a proof. Working backwards from your goal often is a good strategy *on your scratch paper*, but when it's time to *write* your proof, you have to start with the hypotheses and work to the conclusion.

15. **Be concise.** Most students err by writing their proofs too short, so that the reader can't understand their logic. It is nevertheless quite possible to be too wordy, and if you find yourself writing a full-page essay, it's probably because you don't really have a proof, but just an intuition. When you find a way to turn that intuition into a formal proof, it will be much shorter.

16. **Introduce every symbol you use.** If you use the letter "k," the reader should know exactly what k is. Good phrases for introducing symbols include "Let $n \in \mathbb{N}$," "Let k be the least integer such that…," "For every real number a…," and "Suppose that X is a counterexample."

17. **Use appropriate quantifiers (once).** When you introduce a variable $x \in S$, it must be clear to your reader whether you mean "for all $x \in S$" or just "for some $x \in S$." If you just say something like "$y = x^2$ where $x \in S$," the word "where" doesn't indicate whether you mean "for all" or "some".

    Phrases indicating the quantifier "for all" include "Let $x \in S$"; "for all $x \in S$"; "for every $x \in S$"; "for each $x \in S$"; etc. Phrases indicating the quantifier "some" (or "there exists") include "for some $x \in S$"; "there exists an $x \in S$"; "for a suitable choice of $x \in S$"; etc.

    On the other hand, don't introduce a variable more than once! Once you have said "Let $x \in S$," the letter x has its meaning defined. You don't *need* to say "for all $x \in S$" again, and you definitely should *not* say "let $x \in S$" again.

18. **Use a symbol to mean only one thing.** Once you use the letter x once, its meaning is fixed for the duration of your proof. You cannot use x to mean anything else.

19. **Don't "prove by example."** Most problems ask you to prove that something is true "for all"—You *cannot* prove this by giving a single example, or even a hundred. Your answer will need to be a logical argument that holds for *every example there possibly could be*.

20. **Write "Let $x = \ldots$," not "Let $\cdots = x$."** When you have an existing expression, say a^2, and you want to give it a new, simpler name like b, you should write "Let $b = a^2$," which means, "Let the new symbol b mean a^2."This convention makes it clear to the reader that b is the brand-new symbol and a^2 is the old expression he/she already understands.

    If you were to write it backwards, saying "Let $a^2 = b$," then your startled reader would ask, "What if $a^2 \neq b$?"

21. **Make your counterexamples concrete and specific.** Proofs need to be entirely general, but counterexamples should be absolutely concrete. When you provide an example or counterexample, make it as specific as possible. For a set, for example, you must name its elements, and for a function you must give its rule. Do not say things like "θ could be one-to-one but not onto"; instead, provide an actual function θ that *is* one-to-one but not onto.

22. **Don't include examples in proofs.** Including an example very rarely adds anything to your proof. If your logic is sound, then it doesn't need an example to back it up. If your logic is bad, a dozen examples won't help it (see rule 19). There are only two valid reasons to include an example in a proof: if it is a *counterexample* disproving something, or if you are performing complicated manipulations in a general setting and the example is just to help the reader understand what you are saying.

23. **Use scratch paper.** Finding your proof will be a long, potentially messy process, full of false starts and dead ends. Do all that on scratch paper until you find a real proof, and only then break out your clean paper to write your final proof carefully. *Do not hand in your scratch work!*

    Only sentences that actually contribute to your proof should be part of the proof. Do not just perform a "brain dump," throwing everything you know onto the paper before showing the logical steps that prove the conclusion. *That is what scratch paper is for.*

Appendix C

Fancy Mathematical Terms

Here are some important mathematical terms that you will encounter in this course and throughout your mathematical career.

1. **Definition**—a precise and unambiguous description of the meaning of a mathematical term. It characterizes the meaning of a word by giving all the properties and only those properties that must be true.

2. **Theorem**—a mathematical statement that is proved using rigorous mathematical reasoning. In a mathematical paper, the term theorem is often reserved for the most important results.

3. **Lemma**—a minor result whose sole purpose is to help in proving a theorem. It is a stepping stone on the path to proving a theorem. Very occasionally lemmas can take on a life of their own (Zorn's lemma, Urysohn's lemma, Burnside's lemma, Sperner's lemma).

4. **Corollary**—a result in which the (usually short) proof relies heavily on a given theorem (we often say that "this is a corollary of Theorem A").

5. **Proposition**—a proved and often interesting result, but generally less important than a theorem.

6. **Conjecture**—a statement that is unproved, but is believed to be true (Collatz conjecture, Goldbach conjecture, twin prime conjecture).

7. **Claim**—an assertion that is then proved. It is often used like an informal lemma.

8. **Axiom/Postulate**—a statement that is assumed to be true without proof. These are the basic building blocks from which all theorems are proved (Euclid's five postulates, Zermelo-Frankel axioms, Peano axioms).

9. **Identity**—a mathematical expression giving the equality of two (often variable) quantities (trigonometric identities, Euler's identity).

10. **Paradox**—a statement that can be shown, using a given set of axioms and definitions, to be both true and false. Paradoxes are often used to show the inconsistencies in a flawed theory (Russell's paradox). The term paradox is often used informally to describe a surprising or counterintuitive result that follows from a given set of rules (Banach-Tarski paradox, Alabama paradox, Gabriel's horn).

Appendix D

Definitions in Mathematics

It is difficult to overstate the importance of definitions in mathematics. Definitions play a different role in mathematics than they do in everyday life.

Suppose you give your friend a piece of paper containing the definition of the rarely-used word **rodomontade**. According to the Oxford English Dictionary* (OED) it is:

> A vainglorious brag or boast; an extravagantly boastful, arrogant, or bombastic speech or piece of writing; an arrogant act.

Give your friend some time to study the definition. Then take away the paper. Ten minutes later ask her to define rodomontade. Most likely she will be able to give a reasonably accurate definition. Maybe she'd say something like, "It is a speech or act or piece of writing created by a pompous or egotistical person who wants to show off how great they are." It is unlikely that she will have quoted the OED word-for-word. In everyday English that is fine—you would probably agree that your friend knows the meaning of the rodomontade. This is because most definitions are *descriptive*. They describe the common usage of a word.

Let us take a mathematical example. The OED† gives this definition of *continuous*.

> Characterized by continuity; extending in space without interruption of substance; having no interstices or breaks; having its parts in immediate connection; connected, unbroken.

Likewise, we often hear calculus students speak of a continuous function as one whose graph can be drawn "without picking up the pencil." This definition is descriptive. (As we learned in calculus the picking-up-the-pencil description is not a perfect description of continuous functions.) This is not a mathematical definition.

Mathematical definitions are *prescriptive*. The definition must prescribe the exact and correct meaning of a word. Contrast the OED's descriptive definition of continuous with the the definition of continuous found in a real analysis textbook.

> A function $f : A \to \mathbb{R}$ is **continuous at a point** $c \in A$ if, for all $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $|x-c| < \delta$ (and $x \in A$) it follows that $|f(x)-f(c)| < \varepsilon$. If f

*http://www.oed.com/view/Entry/166837
†http://www.oed.com/view/Entry/40280

> is continuous at every point in the domain A, then we say that f is **continuous on** A.[‡]

In mathematics there is very little freedom in definitions. Mathematics is a deductive theory; it is impossible to state and prove theorems without clear definitions of the mathematical terms. The definition of a term must completely, accurately, and unambiguously describe the term. Each word is chosen very carefully and the order of the words is critical. In the definition of continuity changing "there exists" to "for all," changing the orders of quantifiers, changing $<$ to $\leq$ or $>$, or changing $\mathbb{R}$ to $\mathbb{Z}$ would completely change the meaning of the definition.

What does this mean for you, the student? Our recommendation is that at this stage you memorize the definitions word-for-word. It is the safest way to guarantee that you have it correct. As you gain confidence and familiarity with the subject you may be ready to modify the wording. You may want to change "for all" to "given any" or you may want to change $|x-c|<\delta$ to $-\delta<x-c<\delta$ or to "the distance between x and c is less than δ."

Of course, memorization is not enough; you must have a conceptual understanding of the term, you must see how the formal definition matches up with your conceptual understanding, and you must know how to work with the definition. It is perhaps with the first of these that descriptive definitions are useful. They are useful for building intuition and for painting the "big picture." Only after days (weeks, months, years?) of experience does one get an intuitive feel for the ε,δ-definition of continuity; most mathematicians have the "picking-up-the-pencil" definitions in their head. This is fine as long as we know that it is imperfect, and that when we prove theorems about continuous functions mathematics we use the mathematical definition.

We end this discussion with an amusing real-life example in which a descriptive definition was not sufficient. In 2003 the German version of the game show *Who wants to be a millionaire?* contained the following question: "Every rectangle is: (a) a rhombus, (b) a trapezoid, (c) a square, (d) a parallelogram."

The confused contestant decided to skip the question and left with €4000. Afterward the show received letters from irate viewers. Why were the contestant and the viewers upset with this problem? Clearly a rectangle is a parallelogram, so (d) is the answer. But what about (b)? Is a rectangle a trapezoid? We would describe a trapezoid as a quadrilateral with a pair of parallel sides. But this leaves open the question: can a trapezoid have *two* pairs of parallel sides or must there only be *one* pair? The viewers said two pairs is allowed, the producers of the television show said it is not. This is a case in which a clear, precise, mathematical definition is required.

[‡]This definition is taken from page 109 of Stephen Abbott's *Understanding Analysis*, but the definition would be essentially the same in any modern real analysis textbook.