Numbers, Polynomials, and Games: An excursion into algebraic worlds

John Perry

August 24, 2016

Wonder is the desire to understand an observation whose cause eludes us or exceeds our knowledge. So wonder can stimulate pleasure, insofar as it stimulates a hope of understanding what we observe. This is why wondrous things please us.

— Thomas Aquinas, *Summa Teologica*, Prima pars secundæ partis, q. 32 art. 8 co. (loose translation)

Typeset using Lyx and LaTeX, in the Gentium typeface, copyright SIL international. See www.lyx.org, www.tug.org, www.sil.org for details.

Some quotes were found using the Mathematical Quotation Server at Furman University.

Contents

Nomenclature

$[r]$	the element $r + n\mathbb{Z}$ of $\mathbb{Z}_n$
$\langle g \rangle$	the group (or ideal) generated by g
$я$	the identity element of a monoid or group
$\|P\|_{\text{sq}}$	the sqare distance of the point P to the origin
$a \equiv_d b$	a is equivalent to b (modulo d)
A_3	the alternating group on three elements
$A \lhd G$	for G a group, A is a normal subgroup of G
$A \lhd R$	for R a ring, A is an ideal of R
$\text{Aut}(S)$	the group of automorphisms on S
$[G, G]$	commutator subgroup of a group G
$[x, y]$	for x and y in a group G, the commutator of x and y
$D_n(\mathbb{R})$	the set of all diagonal matrices whose values along the diagonal is constant
$d\mathbb{Z}$	the set of integer multiples of d
$\mathbb{F}(\alpha)$	field extension of $\mathbb{F}$ by *alpha*
G/A	the set of left cosets of A
$G\backslash A$	the set of right cosets of A
gA	the left coset of A with g
$\text{GL}_m(\mathbb{R})$	the general linear group of invertible matrices
g^z	for G a group and $g, z \in G$, the conjugation of g by z, or zgz^{-1}
$H < G$	for G a group, H is a subgroup of G
$\text{lcm}(t, u)$	the least common multiple of the monomials t and u

Preface

> *A wise man speaks because he has something to say; a fool because he has to say something.*
>
> — Plato

Why this text?

This text has three goals.

The first goal is to introduce you to the algebraic view of the world. This view reveals strange mathematical creatures that connect seemingly unrelated mathematical ideas. I have tried to organize the excursion so that, by the time you're done reading at least the first chapter or two, you will understand that the world we inhabit is not merely different, but *wonderfully* different.

The second goal is to take you *immediately* into this wonderful world. While it is possible to teach algebra without ever mentioning polynomials, and that is in fact how I learned it, a student can find himself left with a gnawing question: What do groups, rings, etc. have to do with "algebra"? Surely they hold *some* relationship to polynomials and solving equations? The algebraic world strikes the newcomer as *exotic,* but there's no reason it has to be *esoteric.* You will encounter polynomials and their roots in the very first chapter — indeed, in the very first pages, though how they appear won't be clear until later.

The third goal is to lead you on an intuitive path into this world. Higher algebra is often called "abstract" algebra, with reason. Abstraction is difficult, and requires a certain amount of maturity, patience, and perseverance. Proofs are a big part of algebra, but many students arrive in the course with no more experience than a survey on proof techniques. One class on proofs does not a proof-writer make! Reflecting on my own experience as a student: I reached the requisite maturity later than many of my fellow students. This initially deterred me from pursuing doctoral studies, and even then it took me a while. I like to tell students that I don't have a PhD because I'm smart; I have a PhD because I was too dumb to quit. There's truth to that, but I was also lucky to have had two graduate professors who spent a lot of time elaborating on both how to find justification for an idea, and how to write the proof. I try to do that in class myself, and many exercises provide hints on how to begin and where to look.

What should you do?

Algebra is probably different from the math classes you've had before. Rather than *computation,* it expects *explanation.*

The word "proof" frightens students,[1] but it's really just another word for "explanation." The "Questions" in this text are not here to give you practice with a narrowly-tailored skill, but to develop your ability to speak the algebraic language. Sometimes you'll "see" an answer, but find it difficult to put into words. *That makes sense,* because you don't have much experience giving flesh to your ideas. It's one thing to repeat someone else's words; it's altogether something else to come up with your own. I would advise you to adopt a habit of *memorizing the definitions!* After all, you can't answer algebraic questions if you don't know what the words mean.

Many students recoil from this suggestion, in part because lower-level mathematics classes tend to emphasize computation over definition.[2] Here, if you don't know the definitions, you won't understand the question, let alone find the answer, so start by reviewing definitions. When a student comes to me for help on a problem, I typically start with the question, "What does [very important term in the problem] mean?" More often than not, the student will shrug. Well, *of course* you can't solve it: you don't know what the words mean! Yet the definition is in the text; why didn't you start there?

Don't get the wrong idea: Knowing the definitions may be necessary, but it is rarely sufficient.[3] It is no less discouraging when the excitement of a seemingly great idea gives way to the crushing realization that it won't work out. *That's okay.* You will likely see your instructor goof up from time to time, unless he's the sort of stick-in-the-mud who comes to class perfectly prepared with detailed, impeccable notes. My students don't see that; there are days where I ask them to believe ten impossible things before breakfast.[4] I usually figure out they're impossible and set things straight, but *that's part of the point!* Students without much experience figuring things out need to see their professors do it.

You may be tempted to look up solutions elsewhere. *Don't do that.* To start with, it's often futile; some of the problems are uniquely mine. If you do find one somewhere, you cheat yourself out of both the pleasure of discovery, and the benefit of training your mind at problem-solving skills. A better idea is to talk about the problem with other students, or to question the instructor. Sometimes instructors are actually helpful.

Some of you won't like to read this, but you also need to put aside your expectation of the grades you've typically received heretofore. I'm not saying you won't earn the grade you're accustomed to earn; you may well do so, and I'd be glad for it. Statistically speaking, though, you won't — *and that's okay.* It makes you no less a person, no less a mathematician. I received an F on one of my *graduate-level* algebra tests, yet here I am, teaching it & publishing the occasional research article. Worry instead about this: did you learn something new every time you sat to work on it? That includes mistakes — if you learned only that such-and-such approach doesn't work with this-or-that problem, *you learned something!* Even that is closer to the end than it was before you started.

[1] When I started my PhD studies, I was astonished to learn that many of my classmates had earned their undergraduate degrees without ever writing a proof.

[2] If you doubt me, ask the average A student in Calculus I for the definition of a derivative — *not* how to compute it, so not the *formula,* but the honest-to-goodness definition.

[3] Well-begun is half done… but *only* half-done.

[4] With apologies to Humpty-Dumpty. For that matter, there are days when I discover these notes assert the truth of "facts" that are, in "fact," impossible.

Chapter 1

Noetherian behavior

This is a class on *algebra,* not on *games*, but we will allow ourselves a few moments now and again with two games that distill some important ideas of algebra into a convenient, easily-accessible package. The games are simple enough that children can play them, but some rather deep questions lie behind them.

The unifying theme of this chapter is "Noetherian behavior," named in honor of Emmy Noether, a brilliant mathematician of the early 20th century. Noetherian behavior occurs whenever an ordered chain of events must stabilize. For instance, consider the statement

$$a_1 \geqslant a_2 \geqslant a_3 \geqslant \cdots$$

In certain contexts, this sequence *must* eventually "stabilizes," by which we mean that

$$a_i = a_{i+1} = a_{i+2} = \cdots$$

This is an example of Noetherian behavior. It should be obvious that Noetherian behavior is not a universal principle; after all, the sequence

$$0 > -1 > -2 > -3 > \cdots$$

continues without stabilizing. Yet this behavior, when it does occur, is one of the most important tools of modern algebra.

1·1 Two games

> *Mathematics is a game played according to certain simple rules with meaningless marks on paper.*
>
> — David Hilbert, quoted by N. Rose, *Mathematical Maxims and Minims*

You may have played Nim before, perhaps as part of a computer game; it's rather famous in the theory of games. You have almost certainly not played Ideal Nim before. Both are fairly easy to play, and Nim turns out to be a special case of Ideal Nim. Yet while Nim is fairly easy to analyze, Ideal Nim is not, even though you can play it according to the same basic principles.

Both Nim and Ideal Nim involve fundamental ideas of algebra, so we use them as tools to introduce and illustrate these ideas.

Nim

The basic game of Nim has three rows of pebbles. The first row has seven pebbles; the second row, five; the third row, three.

Players take turns doing the following:

1. Choose a row that still has pebbles in it, and choose a pebble.
2. Remove from that row all pebbles beneath and to the right of your finger.

The player who takes the last pebble wins.

In our examples, we always refer to the first player as Emmy, and the second player as David. The "first" pebble lies leftmost, and we count pebbles from there.

Suppose Emmy chooses the last pebble in the first row, leaving six in that row.

David chooses the first pebble in the second row, leaving none in that row.

This was a terrible move,[1] as Emmy now chooses the fourth pebble in the first row,

[1] To be fair, David has no good moves, so he might as well make that one and minimize the pain.

and David's done for.

Question 1·1 .

Explain why we say, "David's done for." One way to explain this is to show that no matter what move David makes from here on, Emmy always has at least one move left – not just on the first turn, but on every turn from here on.

Question 1·2 .

Try playing several games of Nim against a friend (preferably one who has never played the game before). See if you can work out any strategies for winning. Write them out in words. (Surely you can find *something,* at least something similar to Question 1.1.)

A common mathematical technique is *generalization:* take a scenario with specific numbers, replace them with symbols that stand for general numbers, and see how the scenario changes.

We can generalize Nim in the following way. Choose a number of rows, call it m, then choose m numbers of pebbles, call them $n_1, n_2, \ldots, n_m$.[2] Aside from that, the rules stay the same.

For example, Emmy and David might choose $m = 4$ and $n_1 = 3, n_2 = 4, n_3 = 5$, and $n_4 = 6$. That gives them the following game:

As you can imagine, there is no end to the number of ways you can play Nim.

Question 1·3 .

What values of m and $n_1, \ldots, n_m$ give us the game of Nim with that started the section?

Question 1·4 .

Games of Nim where $m = 1$ are boring. Why?

Question 1·5 .

Suppose $m = 2$.

(a) A game with $n_1 \neq n_2$ is easy for one of the players. Which player, and why?

(b) A game with $n_1 = n_2$ is also easy for one of the players. Which player, and why?

(c) So, really, games of Nim where $m = 2$ are also boring. Why?

[2] Don't let the subscripts frighten you; they're just labels. The symbol n_1 means "the first n," the symbol n_2 means "the second n," and so forth. Mathematicians often use subscripts to list and distinguish related values.

Question 1·6 .

Suppose $m > 2$ and $n_i = n_j$ for some i,j where $i \neq j$.

(a) What do we mean by the phrase "$m > 2$ and $n_i = n_j$ for some i,j where $i \neq j$?" Try to explain without using the symbols n, i, or j.

(b) With the given assumption, a player trying to decide on a good move might as well consider rows i and j to be have been completely played out already. Why?

Question 1·7 .

Suppose you generalize Nim further by letting a row of pebbles extend without end to the right. When this happens, we'll say[3] the number of pebbles in the row is ω. For example, the Nim game with $m = 3$ and $n_1 = 3$, $n_2 = 5$, and $n_3 = \omega$ would look like this:

Is it possible to play such a game indefinitely, making turns so that it never ends? If so, describe a sequence of moves that would never end. If, however, this is impossible, explain why.

Ideal Nim

Ideal Nim is another generalization of Nim. The playing board consists of points with integer values in the first quadrant of the x-y axis. Choose a few[4] points for a set F. Any point not northeast of at least one point in F lies within a *F*orbidden *F*rontier. Shade those points red. More precisely, (c,d) is red if for each $(a,b) \in F$, we have $0 \leqslant c < a$ or $0 \leqslant d < b$. There is also a gray region G, which is "*G*one from *G*ameplay," but it begins empty.

Players take turns doing the following:

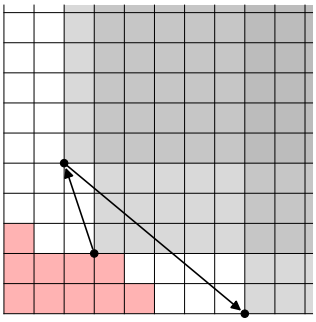
1. Choose a point (a,b) that is in neither the Forbidden Frontier, nor Gone from Gameplay.
2. Add to G the region of points (c,d) that are northeast of (a,b). More precisely, add to G all points (c,d) that satisfy $c \geqslant a$ and $d \geqslant b$.

The player who makes the last move wins.

[3] The last symbol in this sentence is a letter in the Greek alphabet called "omega," not the letter w in the "Latin" alphabet. The letter ω will show up repeatedly in these notes, often with very different meanings, and the letter w will show up, also with different meanings, so be careful.

[4] Not too many points, nor too large in value. Certainly not infinitely many. It doesn't change the properties of the game, but you'd waste an awwwwful lot of time trying to figure out the gameplay.

In the example below, Emmy and David have chosen the points $(0,3)$, $(1,2)$, $(4,1)$, and $(5,0)$ for F. Emmy chose the position $(3,2)$ on her first turn; David chose $(2,5)$ on his first turn; and Emmy chose $(8,0)$ on her second turn.



Don't overlook a difference in the definition of the regions. Players *may* choose a point on the border of the Forbidden Frontier; such points are northeast of a point in F. They *may not* choose a point on the border of the region Gone from Gameplay, as such points are considered northeast of G, and thus *in* G. So, Emmy was allowed to choose the point $(3,2)$, which borders the red region, but may not now choose the point $(3,3)$, because it borders the current gray region. She could, of course, choose the point $(2,2)$, or even the point $(1,2)$, as they border the red region, but not the gray.[5]

When playing this game, certain questions might arise. They may not seem mathematical, but *all of them are!* In fact, *all of them are related to algebraic ideas!*

- Must the game end? or is it possible to have a game that continues indefinitely? Why, or why not? Does the answer change if we play in three dimensions, or more?
- Is there a way to count the number of moves available, even when there are infinitely many?
- What strategy wins the game?

We consider these questions (and more!) throughout the text. We will *not* be able to answer all of them; at least one is an open research question.[6] Maybe you can solve them someday.

You should take from this introduction three main points.

- Mathematics can apply to problems that do not appear mathematical.
- Questions that seem unrelated to mathematics can be very important for mathematics.[7]
- It is a very, very good thing to ask questions!

[5] The game can be played so that players may not dance along the Forbidden Frontier, but then we'd have to interpret the word "northeast" differently for this region than for the other.

[6] An amazing aspect of mathematics is that simple questions can lead to profound results in research!

[7] This is *not* the same as the previous point. Make sure you understand why.

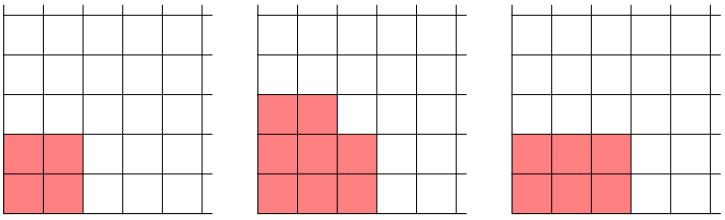
Meanwhile, *play the game!* A few example games appear below to help you along; some of them will be "partially" played.

But don't play thoughtlessly. As a student of mathematics, you should prepare yourself to think carefully and precisely. Intuition and insight are good and necessary, but deduction and dogged determination are no less required. When someone wins, talk about which moves seemed "obvious," and think about the strategy used. With enough effort, you should find a winning strategy for all the games given, but don't feel bad if you don't.

Your explanations to the questions need not look "mathematical", but *they should be yours,* and *they should be convincing,* or at least reasonable. If you can formulate reasonable answers, you will have succeeded at important tasks that helped solve important problems in mathematics. That's no small feat for someone just starting out in algebra!

Question 1·8 .

Play the following games with a friend. If you play carefully, you should find that *Emmy* (the starting player) is guaranteed a win for each game.

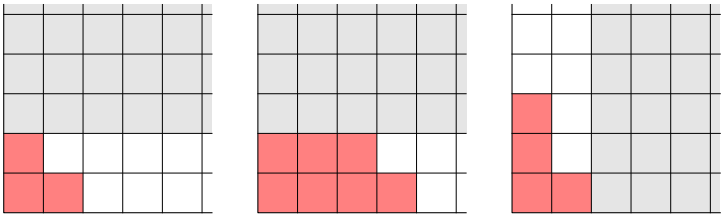


Question 1·9 .

What characteristic do all the games in Question 1.8 share? How does that characteristic guarantee Emmy a win? *Hint:* Think geometrically.

Question 1·10 .

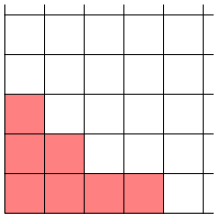
Play the following games with a friend. They have already been partially played. It is Emmy's turn, but this time *David* is guaranteed a win for each game. Try to find how.



Question 1·11 .

What characteristic do all the games in Question 1.10 share? How does that characteristic guarantee David a win? *Hint:* Think geometrically.

Question 1·12 .
What move guarantees Emmy a win in the following game? *Why* does that move guarantee her a win? *Hint:* Try to use the previous two problems.



Question 1·13 .
Suppose two players with infinite lifespan and patience are presented with an arbitrary game of Ideal Nim (the heavenly emanations of David Hilbert and Emmy Noether, perhaps). Does their game *have* to end, or could it go on for ever? Why or why not?

1·2 Sets

> *The fear of infinity is a form of myopia that destroys the possibility of seeing the actual infinite, even though its highest form has created and sustains us, and its secondary, transfinite forms occur all around us and even inhabit our minds.*
> — Georg Cantor

One of the fundamental objects of mathematical study, if not *the* fundamental object of mathematical study, is the set. We assume you've seen sets before, so we won't go into much detail, and in some cases will content ourselves with intuitive discussion rather than precise rigor.

Definition 1·14. A **set** consists of all objects that share a certain property. This property may simply be membership.

- Any object with that property is an **element** of the set.
- A set S is a **subset** of a set T if every element of S is also an element of T.
- Two sets are **equal** if and only if each is a subset of the other.

We typically write a set explicitly by enclosing *or describing* its elements within braces. I emphasize "describing" because it is typically burdensome, even impossible, to list all elements of a set explicitly. For instance, we can list explicitly the set of names for the fingers on one's hand as $F = \{\text{thumb, index, middle, ring, pinky}\}$, but any set with infinitely many elements requires description. Sometimes, that simply means listing a few elements, then concluding with an ellipsis to show that the pattern should continue. Other times, it requires a description in words. It may amaze you that words can encapsulate ideas about infinity within a few marks on paper, but it's true.

Fundamental sets

The fact that they are "fundamental" is a pretty big hint that you'll need to remember the following sets.

- The set of **natural numbers** is[8]

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}.$$

  The funny-looking $\mathbb{N}$ is a standard symbol to represent the natural numbers; the style is called "blackboard bold".[9]

- Since even a small plus sign can make a big difference, we adopt a similar symbol for the set of **positive numbers**

$$\mathbb{N}^+ = \{1, 2, 3, \dots\}.$$

  The set of **integers** is

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$

  We can also define it in **set-builder** notation,

$$\mathbb{Z} = \mathbb{N} \cup \{-x : x \in \mathbb{N}\}.$$

Don't pass over that set-builder notation too quickly. Take a moment to decipher it, as this notation pops up from time to time. Don't let it intimidate you! The world is a complex place, and it's amazing how a good choice of words can simplify complexity.[10] Transliterated, the set-builder definition says,

> The set of integers ($\mathbb{Z}$) is ($=$) the union ($\cup$) of the naturals ($\mathbb{N}$) and
> the set of elements ($\{\dots\}$) that are the opposite ($-x$) of any natural number ($: x \in \mathbb{N}$).

Translated, the integers are the union of the naturals with their opposites.

Some readers might think it clearer to write, "$\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$", and I suppose we could have, but then we'd have to explain what $-\mathbb{N}$ means, because that construction won't always make obvious sense. (Think about $-F$, where F is the set of fingers.) In fact, some authors use $-S$ to mean the complement of S, which you may have seen as $\sim S$ or S^c, something completely different from "the set of negatives." Not everyone writes mathematics the same way.

Elements of a set can appear in other sets, as well; when *all* elements of one set appear in another, the first is a **subset** of the second. When S is a subset of T, we write $S \subseteq T$; the bottom bar emphasizes that a subset can equal its containers, in the same way that $\leqslant$ applies to two equal numbers. You can chain these, so our fundamental sets so far satisfy

$$\mathbb{N}^+ \subseteq \mathbb{N} \subseteq \mathbb{Z}.$$

[8] Not everyone starts $\mathbb{N}$ with 0, and some authors refer to $\{0, 1, 2, 3, \dots\}$ as the "whole numbers". While this can be confusing, it's not uncommon, and highlights how you have to pay careful attention to definitions.

[9] I've read somewhere (can't remember where) that textbooks originally indicated these sets with bold characters. Professors can't write bold at the blackboard, or at least not easily, so they resorted to doubling the letters. Textbooks nowadays have adopted the professors' notation.

[10] "Brevity is the soul of wit." — Shakespeare, *Hamlet*

When we know a subset S is *not* equal to its container T, and we want to emphasize this, we cross out the bottom bar and write $S \subsetneq T$.[11] You can chain these, as well, so that

$$\mathbb{N}^+ \subsetneq \mathbb{N} \subsetneq \mathbb{Z}.$$

Subsets of this latter variety are called **proper subsets**. Don't confuse this with $S \not\subseteq T$, which means that S is *not* a subset of T. This happens when at least one element of S is not in T, whereas $S \subsetneq T$ means every element of S is in T, but at least one element of T is not in S.

Set arithmetic

We assume you've seen **union**s and **intersection**s. We can define them with set-builder notation:

$$S \cup T = \{x : x \in S \text{ or } x \in T\};$$
$$S \cap T = \{x : x \in S \text{ and } x \in T\}.$$

You may not have seen **set difference**; the difference of S and T is the set of elements in A that are not in B. That is,

$$S \backslash T = \{s \in S : s \notin T\}.$$

For example, we could describe the set of negative numbers as $\mathbb{Z}\backslash\mathbb{N}$.

A very useful construction is the **Cartesian product**, which creates *new objects* from two sets, in the form of a sequence of two elements:

$$S \times T = \{(s,t) : s \in S \text{ and } t \in T\}.$$

You've already see an example of this; the playing field of Ideal Nim is $\mathbb{N}\times\mathbb{N}$, since any position is a point "with integer values in the first quadrant of the x-y axis." Points are pairs (a, b), and the qualified, "the first quadrant," tells us that both $a, b \in \mathbb{N}$. The set $\mathbb{N} \times \mathbb{N}$ is important enough to remember by a name, and will appear again (at least when we play the game) so we will call it **the natural lattice**, or just **the lattice** when we're feeling a bit lazy, which we usually are, since in any case we don't typically deal with other lattices in this text.

Question 1·15 .
Suppose $S = \{1, 3, 5, 7\}$, $T = \{2, 4, 6, 8\}$, and $U = \{3, 4, 5, 6\}$. Construct (a) $S \cup T$, (b) $S \cap T$, (c) $(S \cup T) \backslash U$, and (d) $S \times T$.

A "real-life" example of a Cartesian product that the author is all too familiar with is the absent-minded tic of touching a hand's fingers to each other. (Guess what I was doing a few moments ago.) Each touch is a *pairing* of fingers, such as (thumb, middle) or (pinky, pinky). Inasmuch as pairings correlate to Cartesian products, we can describe the pairings of all fingers as $F \times F$, where F is again the set of all fingers.

Question 1·16 .
How large is $F \times F$? That is, how many elements does it have?

[11] Some authors use $\subset$, but other authors use $\subset$ when the two sets are equal, so we avoid $\subset$ altogether.

Question 1·17 .
If a set S has m elements and a set T has n elements, how many elements will $S \times T$ have? Explain why.

If $S = T$, we can write S^2 instead of $S \times T$. Hence we can abbreviate the lattice of Ideal Nim as $\mathbb{N}^2$.

When needed, we can chain sets in the Cartesian product to make sequences longer than mere pairs; we can even describe all infinite sequences of integers as

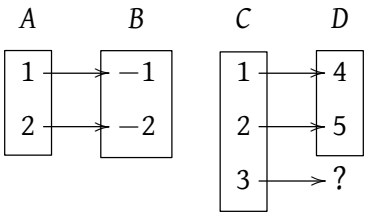
$$\mathbb{Z}^\infty = \prod_{i=1}^{\infty} \mathbb{Z} = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \times \cdots = \{(a, b, c, \dots) : a, b, c, \dots \in \mathbb{Z}\}.$$

That new symbol, $\prod$, means "product", much as Σ means "sum". Writing phrases like "the first element of P" or "the four hundred twenty-fifth element of P" all the time grows cumbersome, so we'll adopt the convention that if P is a sequence of numbers, then p_i will stand for the ith element of P. For example, if $P = (5, 8, 3, -2)$ then $p_1 = 5$ and $p_4 = -2$.

Definition 1·18. Two sets S and T have the same size (or **cardinality**) if you can match each element of S to a unique element of T, covering all the elements of T in the process. More precisely, S and T have the same cardinality if you can create a mapping from S to T where

- each element of S maps to a unique element of T (so the function is **one-to-one**), and
- for any element of T, you can find an element of S that maps there (so the function is **onto**).

For example, the sets $A = \{1, 2\}$ and $B = \{-1, -2\}$ have the same cardinality because I can match them as follows, while the sets $C = \{1, 2, 3\}$ and $D = \{4, 5\}$ do not, because I cannot find a unique target for at least one element of C:



Question 1·19 .

(a) Show that S and T of Question 1.15 have the same cardinality. *Don't* just count the elements; exhibit a unique matching. Is there more than one matching? If so, list a couple more. How many do you think there are?

(b) Show that $\mathbb{E} = \{0, 2, 4, 6, \dots\}$ and $\mathbb{O} = \{1, 3, 5, 7, \dots\}$ have the same cardinality. In this case, the number of elements is infinite, so you can't count them, nor draw a complete picture, so use words to describe the matching, or even a formula.

(c) Show that an arbitrary set S has the same size as itself. This may seem silly, but it forces you to think about using the *definition* of cardinality, since you don't know what the elements of S are. Don't forget to think about the case where S is empty.

(d) Show that $\mathbb{N}$ and $\mathbb{Z}$ have the same cardinality. It helps if you map negative integers to $\mathbb{O}$ and positive integers to $\mathbb{E}$. This is a little weird, because $\mathbb{N} \subsetneq \mathbb{Z}$, so you wouldn't expect them to be the same size, but weird things do happen when you start mucking around in infinite sets.

1·3 Orderings

> *The mathematical sciences particularly exhibit order, symmetry, and limitation; and these are the greatest forms of the beautiful.*
> — Aristotle

A **relation** between two sets S and T is a subset of $S \times T$. For instance, the pairings of fingers is a relation on $F \times F$, where the set of fingers, while $S \times T$ is itself a relation.

A **function** is any relation $F \subsetneq S \times T$ such that every $s \in S$ corresponds to exactly one $(s, t) \in F$. Put another way, any two (a, b) and (c, d) in F satisfy $a \neq c$ (but $b = d$ is okay). If F is a function, we write $F : S \to T$ instead of $F \subseteq S \times T$, and $F(s) = t$ instead of $(s, t) \in F$.

Two kinds of relations are essential to algebra. The first is a **homomorphism**, which is a special kind of function; we talk about those later on, so pretend I didn't mention them for now. The second is a special subset of $S \times S$, called an **ordering** on S. There are several types of orderings, so it's important to make precise the kind of ordering you mean.

Partial orderings

A **partial ordering** on S is an ordering P that satisfies three properties. Let $a, b, c \in S$ be arbitrary.

Reflexive? Every element is related to itself; that is, $(a, a) \in P$.

Antisymmetric? Symmetry implies equality; that is, if $(a, b) \in P$ and $(b, a) \in P$, then $a = b$.

Transitive? If $(a, b) \in P$ and $(b, c) \in P$, then $(a, c) \in P$.

Suppose we let P be the ordering of your fingers from left to right, or in set-builder notation,

$$P = \{(x, y) \in F \times F : x \text{ lies to the left of } y\}.$$

Then (thumb,middle) $\in P$ and (ring,pinky) $\in P$ but (index,index) $\notin P$. This is a partial ordering.

It is highly inconvenient to write orderings this way, so usually mathematicians adopt a notation involving "ordering symbols" such as $\leqslant$, $\prec$, and so forth. This allows us to write $(a, b) \in P$ more simply as $a \prec b$, and we will do this from now on. That allows us to rewrite the properties of a partial ordering as follows, using $\preceq$ as our ordering:

Reflexive? $a \preceq a$.

Antisymmetric? If $a \preceq b$ and $b \preceq a$, then $a = b$.

Transitive? If $a \preceq b$ and $b \preceq c$, then $a \preceq c$.

Now that things are a little easier to read, we introduce a few important orderings.

One example of a partial ordering is in the subset relation. If we fix a set S, then we can view $\subseteq$ as a relation on the subsets of S. For instance, if $S = \mathbb{N}$ then $\{1, 3\}$ is "less than" $\{1, 3, 7\}$ inasmuch as $\{1, 3\} \subseteq \{1, 3, 7\}$.

Definition 1·20. For any set S, let $P(S)$ denote the set of all subsets of S. We call this the **power set** of S.

Fact 1·21. *Let S be any set. The relation on $P(S)$ defined by $\subseteq$ is a partial ordering.*

Why? Let $A, B \in P(S)$. We need to show that $\subseteq$ satisfies the three properties of a partial ordering.

Reflexive? Certainly $A \subseteq A$, since any $a \in A$ is by definition an element of A. So $\subseteq$ is reflexive.

Antisymmetric? Assume $A \subseteq B$ and $B \subseteq A$. By definition of set equality, $A = B$.

Transitive? Assume $A \subseteq B$ and $B \subseteq C$. We want to show $A \subseteq C$. The definition of $\subseteq$ tells us this is true if every $a \in A$ is also in C, so let $a \in A$ be arbitrary. We know $A \subseteq B$, so by definition $a \in B$. We know $B \subseteq C$, so by definition $a \in C$. Since a was arbitrary, $A \subseteq C$, as desired. □

Next we look at the ordering you're most accustomed to.

Definition 1·22 (The natural ordering of $\mathbb{Z}$)**.** For any $a, b \in \mathbb{Z}$, we write $a \leqslant b$ if $b - a \in \mathbb{N}$. We can also write $b \geqslant a$ for this situation. If $a \leqslant b$ but $a \neq b$, we write $a < b$, or $b > a$.

Figure 1·1 illustrates this relationship by for the relation $x \leqslant y$ on $\mathbb{N}$ by plotting on the lattice the elements of the set $\leqslant$. Elements of $\leqslant$ are the black points whose y-value equals or exceeds the x-value. White points are *not* in the set O. It's worth asking yourself: which ordering *do* those white points describe?

Fact 1·23. *The natural ordering of $\mathbb{Z}$ is a partial ordering.*

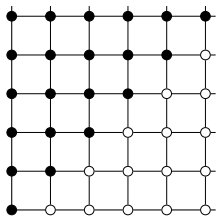


Figure 1·1: Diagram of the relation $\leqslant$ on $\mathbb{N}$.

Question 1·24 .
Fill in the blanks of Figure 1·2 to show why Fact 1·23 is true.

In the future, you can think of the $\leqslant$ ordering in the intuitive manner you're accustomed to. Use it to answer the following questions.

Question 1·25 .
One of our claims in the proof amounts is equivalent to saying that if $i, s, t \in \mathbb{Z}$, then $s \leqslant t$ if and only if $s + i \leqslant t + i$. Why is this true?

Question 1·26 .
Show that $a \leqslant |a|$ for all $a \in \mathbb{Z}$. *Hint:* You need to consider two cases: one where $|a| = a$, the other where $|a| = -a$. (Yes, the second case is quite possible! Look at some "small" integers to see why.)

Question 1·27 .
Let $a, b \in \mathbb{N}$ and assume that $0 < a < b$. Let $d = b - a$. Show that $d < b$.

Question 1·28 .
Let $a, b, c \in \mathbb{Z}$ and assume that $a \leqslant b$. Prove that

(a) $a + c \leqslant b + c$;

(b) if $c \in \mathbb{N}$, then $a \leqslant a + c$;

(c) if $c \in \mathbb{N}$, then $ac \leqslant bc$; and

(d) if $c \in \mathbb{N}^+$ *and also* $a \in \mathbb{N}^+$, then $c \leqslant ac$.

What about the lattice?

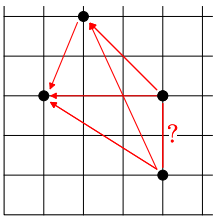
Definition 1·29 (The x-axis, y-axis, and lex orderings of the lattice)**.** For any $P, Q \in \mathbb{N}^2$, we write

- $P \prec_x Q$ if $p_1 < q_1$;

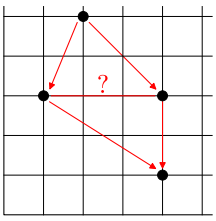
Claim: The natural ordering of $\mathbb{Z}$ is a partial ordering.
Proof:

1. We claim that $\leqslant$ is reflexive. To see why, let $a \in \mathbb{Z}$.
   (a) Observe that $a - a =$____.
   (b) This difference is an element of ____.
   (c) By definition, $a \leqslant a$.
   (d) We chose a from $\mathbb{Z}$ arbitrarily, so this is true of ____ element of $\mathbb{Z}$.
2. We claim that $\leqslant$ is antisymmetric. To see why, let $a, b \in \mathbb{Z}$.
   (a) Assume that $a \leqslant b$ and ____.
   (b) By definition, $b - a \in \mathbb{N}$ and ____.
   (c) By the distributive property, $-(b - a) =$____. (Write it as subtraction.)
   (d) In (b), we explained that $b - a \in \mathbb{N}$. In (c), we showed that $-(b - a) \in \mathbb{N}$. The only natural number whose opposite is also natural is ____.
   (e) By substitution, $b - a =$____.
   (f) By definition, $a = b$.
   (g) We chose a and b from $\mathbb{Z}$ arbitrarily, so this is true of ____ pair of elements of $\mathbb{Z}$.
3. We claim that $\leqslant$ is transitive. To see why, let $a, b, c \in \mathbb{Z}$.
   (a) Assume that $a \leqslant b$ and ____.
   (b) By definition, $b - a \in \mathbb{N}$ and ____.
   (c) Elementary properties of arithmetic tell us that ____+____$= c - a$.
   (d) The sum of any two natural numbers is ____.
   (e) By (c) and (d), then, $c - a \in$____.
   (f) By definition, ____.
   (g) We chose a, b, and c from $\mathbb{Z}$ arbitrarily, so this is true of any three elements of $\mathbb{Z}$.

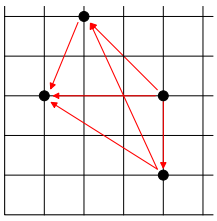
Figure 1·2: Material for Question 1.24



The ordering $\prec_x$ judges one point smaller than another if the first is further left. If the two points are on the same vertical line ($p_1 = q_1$), it makes no decision.



The ordering $\prec_y$ judges one point smaller than another if the first is below the second. If the two points lie on the same horizontal line ($p_2 = q_2$), it makes no decision.



The ordering $\prec_{\text{lex}}$ judges one point smaller than another if the first is further left. If the two points are on the same vertical line, it judges the lower point smaller.

Figure 1·3: Diagrams of the lattice orderings $\prec_x$, $\prec_y$, and $\prec_{\text{lex}}$. Arrows point from larger points to smaller ones.

- $P \prec_y Q$ if $p_2 < q_2$;
- $P \prec_{\text{lex}} Q$ if $p_1 < q_2$, or if $p_1 = q_1$ and $q_1 < q_2$.

We also write $P \preceq_x Q$, $Q \succ_x P$, $Q \succeq_x P$ with meaning analogous to $\leqslant$, $>$, and $\geqslant$; that is, $P \preceq_x Q$ if $P \prec_x Q$ or $P = Q$, and so forth.

These orderings have natural visualizations; see Figure 1·3.

Question 1·30 .
Order each set of lattice points according to the $\prec_x$, $\prec_y$, and $\prec_{\text{lex}}$ orderings. Indicate when the ordering cannot decide which of two points is smaller.

(a) $\{(7,2), (1,3), (0,8), (2,2)\}$

(b) $\{(2,4), (1,5), (5,1), (0,6)\}$

The first question we want to consider is whether the orderings are partial orderings. Determining whether an object has a certain property is very important in mathematics; *explaining* why it has that property is fundamental. Let's consider that a moment.

Theorem 1·31. *The ordering $\preceq_{lex}$ is a partial ordering of the lattice. The orderings $\preceq_x$ and $\preceq_y$ are not.*[12]

[12] As you should know, a **theorem** asserts that a claim is always true. This is also true about **lemma**s, **proposition**s and **fact**s. Most of the assumptions involved are implicit rather than explicit. If we cannot explain convincingly that a claim is always true, we call it a **conjecture**. If you get far enough in your studies, you'll find that a lot of conjectures are themselves widely believed, though remain unproven, and mathematicians use in day-to-day life. Students, however, are not generally allowed to do this on purpose!

Proof. Let $P, Q, R \in \mathbb{N}^2$.

Reflexive? It is easy to verify that $P \preceq_x P$, $P \preceq_y P$, and $P \preceq_{\text{lex}} P$, so the orderings are reflexive.

Antisymmetric? Suppose $P \preceq_{\text{lex}} Q$ and $Q \preceq_{\text{lex}} P$. By definition of the ordering, $p_1 < q_1$ or $p_1 = q_1$ and $p_2 \leqslant q_2$. Similarly, $Q \preceq_{\text{lex}} P$ gives $q_1 < p_1$ or $q_1 = p_1$ and $q_2 \leqslant p_2$. We consider several cases. If $p_1 < q_1$, then $Q \not\preceq_{\text{lex}} P$, contradicting a hypothesis. Similarly, if $q_1 < p_1$, then $P \not\preceq_{\text{lex}} Q$, contradicting a hypothesis. That leaves $p_1 = q_1$ and $p_2 \leqslant q_2$ and $q_2 \leqslant p_2$. By antisymmetry of the natural ordering, $p_2 = q_2$, so $P = Q$.

As for $\preceq_x$ and $\preceq_y$, antisymmetry is the property they both fail. We leave it to you to find a counterexample.

Transitive? Suppose $P \preceq_x Q$ and $Q \preceq_x R$. Then $p_1 \leqslant q_1$ and $q_1 \leqslant r_1$. As in the antisymmetric case, previous work implies $p_1 \leqslant r_1$, so $P \preceq_x Q$. We assumed that $P \preceq_x Q$ and $Q \preceq_x R$ and found that $P \preceq_x R$, so $\preceq_x$ is transitive. A similar argument shows that $\preceq_y$ and $\preceq_{\text{lex}}$ are transitive. □

Question 1·32 .

(a) In the proof of Theorem 1·31, we claimed that neither $\preceq_x$ nor $\preceq_y$ are antisymmetric. To verify this claim, find $P, Q \in \mathbb{N}^2$ such that $P \preceq_x Q$ and $Q \preceq_x P$, but $P \neq Q$.

(b) In the proof of Theorem 1·31, we claimed that the reason $\preceq_x$ is transitive is similar to the reasons $\preceq_y$ and $\preceq_{\text{lex}}$ are transitive. Show this explicitly for $\preceq_{\text{lex}}$.

Question 1·33 .

Define an ordering $\preceq_{x,y}$ on $\mathbb{N}^2$ as follows. We say that $P \preceq_{x,y} Q$ if $p_1 \leqslant q_1$ and $p_2 \leqslant q_2$. Is this a partial ordering? Why or why not?

Question 1·34 .

Define an ordering $\prec_{\text{sums}}$ as follows. We say that $P \prec_{\text{sums}} Q$ if $p_1 + p_2 < q_1 + q_2$ or $p_1 + p_2 = q_1 + q_2$ and $p_1 < q_1$.

(a) Order each set of lattice points according to the $\prec_x$, $\prec_y$, and $\prec_{\text{lex}}$ orderings. Indicate when the ordering cannot decide which of two points is smaller.

(i) $\{(7,2), (1,3), (0,8), (2,2)\}$

(ii) $\{(2,4), (1,5), (5,1), (0,6)\}$

(b) Is $\prec_{\text{sums}}$ a partial ordering? Why or why not?
Hint: Try to look at it geometrically. In the spirit of Figure 1·3, pick a not-too-large point P, then determine which points are smaller than P.

Linear orderings

You can see from Figure 1·3 that there is some ambiguity in the first two orderings, but not in the last one — or not with the points diagrammed, at any rate. The absence of ambiguity is always useful.

Definition 1·35. An ordering $\preceq$ on a set S is **linear** if for any $s, t \in S$ we can decide whether $s \preceq t$ or $t \preceq s$ (or both).

Fact 1·36. *The ordering $\leqslant$ on $\mathbb{N}$ is linear.*

Why? Subtraction of naturals gives us an integer, and the opposite of a non-natural integer is a natural integer. So, for any $m, n \in \mathbb{N}$, we know that either $m-n \in \mathbb{N}$ or $n-m = -(m-n) \in \mathbb{N}$. In other words, either $n \leqslant m$ or $m \leqslant n$. □

We can extend the ordering $\leqslant$ on $\mathbb{N}$ to an ordering on $\mathbb{Z}$ by using the same definition. For example, we can argue that $-5 \leqslant 3$ because $3-(-5) = 8$, and 8 is natural. On the other hand, $-10 \not\leqslant -15$ because $-15-(-10)$, and -5 is not natural.

Fact 1·37. *The ordering $\leqslant$ on $\mathbb{Z}$ is also linear.*

The reasoning is identical, so we omit it.

Question 1·38 .
Show that the ordering $<$ of $\mathbb{Z}$ generalizes "naturally" to an ordering $\prec$ of $\mathbb{Q}$ that is also a linear ordering. *Hint:* Think of how you would decide that $^{24}/_{35} < {}^{20}/_{28}$, or that $^{3}/_{51} < {}^{4}/_{53}$, and go from there.

On the other hand, the orderings $\preceq_x$ and $\preceq_y$ are *not* linear, since $\preceq_x$ cannot decide if $(4,1) \preceq (4,3)$ or $(4,3) \preceq (4,1)$, and $\preceq_y$ cannot decide if $(1,3) \preceq (4,3)$ or $(4,3) \preceq (1,3)$.

The lex ordering *is* able to sort the points diagrammed in Figure 1·3, but is this true for *any* set of points?

Theorem 1·39. *The lex ordering is a linear ordering on the lattice.*

Proof. Let $P, Q \in \mathbb{N}^2$. If $p_1 < q_1$, then $P \preceq_{\text{lex}} Q$, and we are done. If $p_1 > q_1$, then $Q \preceq_{\text{lex}} P$, and we are done. So suppose that $p_1 = q_1$; we consider p_2 and q_2, instead. If $p_2 < q_2$, then $P \preceq_{\text{lex}} Q$, and we are done. If $p_2 > q_2$, then $Q \preceq_{\text{lex}} P$, and we are done. So suppose that $p_2 = q_2$. We now have $p_1 = q_1$ and $p_2 = q_2$, so $P = Q$. This satisfies the definition of $P \preceq_{\text{lex}} Q$, so we are done. □

Question 1·40 .
In the proof of Theorem 1·39, we used implicitly the fact that $\leqslant$ is a linear ordering of the natural numbers. We really ought to give some flesh to that argument, so fill in the blanks of Figure with the correct reasons. (Notice that we actually prove it for $\mathbb{Z}$, a superset of $\mathbb{N}$. This automatically proves it for $\mathbb{N}$. It is often a good idea to prove a fact for a superset, if you can succeed at doing so.)

Question 1·41 .
Is the ordering $\preceq_{x,y}$ of Question 1.33 a linear ordering? Why or why not?

Let $a, b \in \mathbb{Z}$.

1. Suppose $b - a \in \mathbb{N}$. By _____, $a \leqslant b$.
2. Otherwise, $b - a \notin \mathbb{N}$. We know from previous work that $b - a \in \mathbb{Z}$. That means $-(b - a) \in$ _____.
   (a) By _____, $-(b - a) = a - b$.
   (b) By _____, $a - b \in \mathbb{N}$.
   (c) By _____, $b \leqslant a$.
3. We assume that $a, b \in \mathbb{Z}$, and showed that $a \leqslant b$ or $b \leqslant a$. By _____, we are done.

Figure 1·4: "Flesh" for Question 1.40.

Question 1·42 .
Is $<_{\text{sums}}$ a linear ordering? Why or why not? *Hint:* Try to look at it geometrically. In the spirit of Figure 1·3, pick a not-too-large point P, then figure out which points are smaller than P, shade that region, then ask yourself: "Do I know that all the unshaded points must be larger than P?" That should give you some insight into how to answer the question.

1·4 Well ordering and division

Can you do division? Divide a loaf by a knife — what's the answer to that?
— Lewis Carroll

Well ordering

You know from experience that the ordering $\leqslant$ has a smallest element in $\mathbb{N}$; namely, 0. Rather interestingly, every subset of $\mathbb{N}$ has a smallest element. There is no largest element, but The fact that any subset of $\mathbb{N}$ has a smallest element is very interesting.

Definition 1·43. A **well ordering** on a set S is a linear ordering on S for which each subset of S has a smallest element.

You might assume that we are going to prove that $\mathbb{N}$ is well-ordered by $\leqslant$, and in a way we will, but in another way we won't.

Axiom 1·44 (The Well-Ordering Principle)**.** *$\mathbb{N}$ is well-ordered by $\leqslant$.*

An "Axiom" is a statement you assume without proof. So, we are only going to *assume* this property. In fact, it is impossible to prove it, unless you assume something else.

That "something else" is the proof-by-dominoes technique, also called **induction**.

Axiom 1·45 (The Induction Principle)**.** *Let S be a subset of $\mathbb{N}$ that satisfies the following properties.*

(inductive base) $0 \in S$*; and*

(inductive step) for any $s \in S$, we also have $s + 1 \in S$.

Then $S = \mathbb{N}$.

Now, why should induction be true? You can't prove *that,* unless you assume the well-ordering of $\mathbb{N}$. Do you see where this is going?

Fact 1·46. *Axiom 1·44 is logically equivalent to Axiom 1·45; that is, you can't have one without the other.*

We put off an actual proof of this to the end of the section, and in fact you need not concern yourself too much with it. Typically you won't read that in this text, and I'm afraid that you can't appeal to such a judge yourself, but believe you me, this has been something mathematicians hashed out pretty thoroughly in the early 20th century. Some things you just have to accept on faith — which, contrary to popular belief, is *not* the opposite of reason, since these things work out pretty well in practice, and it's pretty reasonable to infer that things that work out in practice really are true.

Example 1·47. We will define a different ordering $\preceq$ on $\mathbb{N}$ according to the following rule:

- even numbers are always smaller than odd numbers;
- otherwise, if a and b are both even or both odd, then $a \preceq b$ if and only if $a \leqslant b$ in the natural ordering.

This ordering sorts the natural numbers roughly so:

$$0, 2, 4, 6, \ldots, 1, 3, 5, 7, \ldots.$$

Is $\preceq$ a well ordering? Indeed it is. Why?

First we show $\preceq$ is a partial ordering:

- Is the ordering reflexive? Let $a \in \mathbb{N}$; we need to show that $a \preceq a$. We use the second part of the rule here, since $b = a$: since $a \leqslant a$ in the natural ordering, $a \preceq a$.
- Is the ordering symmetric? Let $a, b \in \mathbb{N}$, and assume $a \preceq b$ and $b \preceq a$. If both numbers are even or both numbers are odd, then our rule tells us $a \leqslant b$ and $b \leqslant a$ in the natural ordering; since that is symmetric, we infer $a = b$. Otherwise, $a \preceq b$ implies a is even while b is odd, whereas $b \preceq a$ implies a is odd while b is even. That is a contradiction, so $a = b$ is indeed the only possibility.
- Is the ordering transitive? Let $a, b, c \in \mathbb{N}$, and assume $a \preceq b$ and $b \preceq c$. We consider several subcases:

– a even?

  Either c is odd, in which case $a \preceq c$, or c is even. If c is even, then b must also be even; to be otherwise would contradict $b \preceq c$. All three numbers are even, in which case our ordering tells us the natural ordering applies: $a \leqslant b$ and $b \leqslant c$. The natural ordering is transitive, so $a \leqslant c$.

– a odd?

  In this case, $a \preceq b$ implies b is odd, and $b \preceq c$ implies c is odd. All three numbers are odd, in which case our ordering tells us the natural ordering applies: $a \leqslant b$ and $b \leqslant c$. The natural ordering is transitive, so $a \leqslant c$.

Now we show $\preceq$ is a linear ordering. Let $a, b \in \mathbb{N}$; we need to show that $a \preceq b$ or $b \preceq a$. Without loss of generality, we may assume that a is even. If b is odd then our rule tells us $a \preceq b$, and we are done. Otherwise, b is even; in this case, our rule tells us to look at the natural ordering. The natural ordering *is* linear, so $a \leqslant b$ or $b \leqslant a$. By the definition of our rule, then, $a \preceq b$ or $b \preceq a$.

Finally, we show $\preceq$ is a well ordering. Let $S \subseteq \mathbb{N}$; we need to show that S has a least element. Let E be the set of even elements of S, and O the set of odd elements. Observe that $E, O \subsetneq \mathbb{N}$.

- If $E \neq \varnothing$, the well-ordering property tells us that it has a least element; call it e. Let $s \in S$; if s is even, then $s \in E$ and by our choice of e, $e \leqslant s$, so $e \preceq s$; otherwise, s is odd, and our rule tells us $e \preceq s$.

- Otherwise, $E = \varnothing$. The well-ordering property tells us that O has a least element; call it o. Let $s \in S$; if s is even, then $s \in E$, a contradiction to $E = \varnothing$, so s is odd, which puts $s \in O$, and by our choice of o, $o \leqslant s$, so $o \preceq s$.

As S was an arbitrary subset of $\mathbb{N}$, and we found a smallest element with respect to the new ordering, every subset of $\mathbb{N}$ has a smallest element with respect to the new ordering.

What about the set $\mathbb{Z}$? The ordering $\leqslant$ has neither smallest nor largest element, since $\cdots \leqslant -3 \leqslant -2 \leqslant -1 \leqslant 0 \leqslant 1 \leqslant \cdots$. It *is* possible to order $\mathbb{Z}$ a different way, so that it *does* have a smallest element, and in some cases that might be useful. That's an interesting question to ponder, and we leave it to you to pursue.

Question 1·48 .

Devise a *different* ordering of $\mathbb{Z}$ for which every subset of $\mathbb{Z}$ has a smallest element. Call this ordering $\prec$, and prove that it really is a well ordering on $\mathbb{Z}$.

So the definition depends on both the ordering and the set; change one of the two, and the property may fail.

Let's turn to a different set, the lattice $\mathbb{N}^2$. We have three different orderings to choose from; we'll start with $\preceq_x$. Do subsets of $\mathbb{N}^2$ necessarily have smallest elements? Clearly not, as $\preceq_x$ is not even a *linear* ordering! We already saw that $\preceq_x$ fails to order two points on a vertical line, such as $(2, 0)$ and $(2, 1)$. Elements like these are incomparable, so subsets containing them lack a smallest element.

What if we try a different ordering? Again, $\preceq_y$ is not linear, so that's out. On the other hand, $\preceq_{\text{lex}}$ *is* linear, so it stands a chance of being a well-ordering.

Question 1·49 .
Show that the lex ordering $\preceq_{\text{lex}}$ is a well ordering of the lattice $\mathbb{N}^2$. *Hint:* Use the Well-Ordering Principle in one dimension to find a subset of elements that are smallest from a particular point of view. Then use the Well-Ordering Principle in the other dimension to polish it off.

Question 1·50 .
While Question 1.49 refers to a two-dimensional lattice, explain that it doesn't really matter; you can use the same basic proof to show that $\mathbb{N}^n$ is well-ordered by a similar ordering. Also describe the ordering.

Here's another useful consequence of well ordering.

Fact 1·51. *Let S be a set well ordered by $\preceq$, and $s_1 \succeq s_2 \succeq \cdots$ be a nonincreasing sequence of elements of S. The sequence eventually stabilizes; that is, at some index i, $s_i = s_{i+1} = \cdots$.*

Why? Let $T = \{s_1, s_2, \dots\}$. By definition, $T \subseteq S$. By the definition of a well-ordering, S has a least element; call it t. Let $i \in \mathbb{N}^+$ such that $s_i = t$, and let $j > i$. The sequence decreases, which means $s_i \succeq s_j$. By substitution, $t \succeq s_j$. Remember that t is the *smallest* element of T; by definition, $s_j \succeq t$. We have $t \succeq s_j \succeq t$, which is possible only if $t = s_j$. We chose $j > i$ arbitrarily, so every element of the sequence after t must equal t. In other words, $s_i = s_{i+1} = \cdots$, as claimed. □

Question 1·52 .
We asserted that $t \succeq s_j \succeq t$ "is possible only if $t = s_j$." This isn't necessarily obvious, but it is true. Why ? *Hint:* It's one of the properties of the ordering. As to which property, you may need to look further afield than the properties of well orderings; remember that a well ordering is also a linear ordering, which is also a partial ordering. Those three give you a few properties to consider!

We can use this fact to show one of the desired properties of the game.

Dickson's Lemma. *Ideal Nim terminates after finitely many moves.*[13]

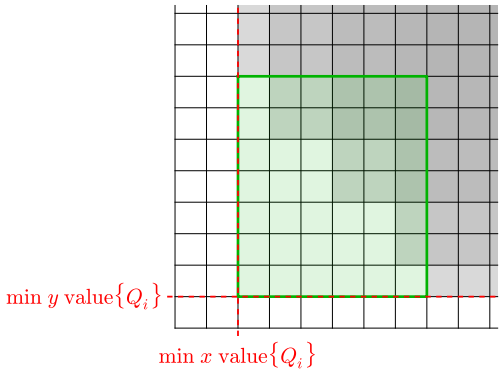
Before going into the details, let's point out a basic, geometrically intuitive argument. Let $P = (a, b)$ be the first position chosen, and Q_1, Q_2, … the subsequent positions chosen. According to the rules, no move $Q = (c, d)$ can satisfy $c \geqslant a$ and $d \geqslant b$, so $c < a$ or $d < b$. In the first case, Q is closer to the x-axis than P, or, $Q \prec_x P$. The set of their x-coordinates would be a nonincreasing sequence of natural numbers, which allows us to apply Fact 1·51. In the second case, Q is closer to the y-axis than P, or, $Q \prec_y P$. That also allows us to apply Fact 1·51.

Superficially, then, it looks as if only finitely many moves are possible. However, if we play enough games, we see that players can sometimes choose positions Q_1, Q_2, … such that $Q_1 \succ_x Q_2 \cdots \succ_x Q_i$, but $Q_i \prec_x Q_{i+1}$. If $Q_j \succ_y Q_{i+1}$ for each $j = 1, 2, \dots, i$, then, as mentioned

[13] Dickson actually proved an equivalent statement.

already, we're dealing with Fact 1·51. As long as we're dealing with one of the two cases, we can see that the game is ending.

What if some $Q_j \prec_y Q_{i+1}$? In this case, Q_{i+1} decreases neither the minimum x-coordinate nor the minimum y-coordinate, and the chain is no longer a nondecreasing sequence. This is really a temporary problem, though; sketch such a game on paper, and we see that any such Q_{i+1} must lie in a rectangle:

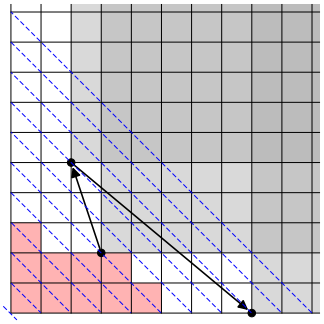


This rectangle has only finitely many positions, and that finiteness means the players will eventually have to break out, at which point either the smallest x-value or the smallest y-value will decrease anew. Writing this precisely is a bit of a bear, but intuitively, it works well.

That said, it's simpler to try the following approach, which works both intuitively and precisely. Essentially, we count the number of positions left. There can be infinitely many positions left, so we organize the points in finite-sized bins. How? Use diagonals of the lattice.

Proof. For any points P of the lattice, let $d(P) = p_1 + p_2$ be the **degree** of P. Basically, $d(P)$ tells you how far away P is from the lower left corner, using lines of slope -1. Recall that the game is defined by a finite set of points F, which defines the red, forbidden region of the gameboard. Let m be the sum of largest x and y values of points in F; notice that $m \geqslant \deg Q$ for any point $Q \in F$.

Suppose we are the beginning of the ith turn of the game. Define H_i as the function on $\mathbb{N}$ such that $H(n)$ is the number of playable points P whose degree is n.[14] For instance, in the game illustrated by



[14]This function is related to an important function in commutative algebra, called the **Hilbert function**, which measures a different phenomenon which we can visualize in a fashion similar to this one.

the number of moves available on each blue diagonal, where $d(P)$ is constant, tells us

$$\begin{aligned} H_1(n) &= (0,0,0,2,3,6,7,8,9,9,9,\dots) \\ H_2(n) &= (0,0,0,2,3,\mathbf{5},5,5,5,5,5,\dots) \\ H_3(n) &= (0,0,0,2,3,5,\mathbf{4},4,4,4,4,\dots) \\ H_4(n) &= (0,0,0,2,3,5,4,\mathbf{3},2,2,2,\dots). \end{aligned}$$

Suppose that on the ith turn, a player chooses position P. Let $m = d(P)$; since we have *removed* available positions, $H_i(m) < H_{i-1}(m)$. Let's focus on a fixed $n \in \mathbb{N}$. The game's rules make it clear that no move can *add* playable positions, which means that $H_j(n) \leqslant H_i(n)$ whenever $j > i$. In other words, n satisfies

$$H_1(n) \geqslant H_2(n) \geqslant \cdots.$$

This is a nondecreasing sequence, so Fact 1·51 tells us it must stabilize eventually. We made no assumption on n, so $H_i(n)$ stabilizes for *every* value of n.

We are not quite done; it is possible that, for some n, we can find $i,k \in \mathbb{N}^+$ such that $H_i(n) = 0$ but $H_i(n+k) \neq 0$, and j,ℓ such that $H_j(n+k) = 0$ but $H_j(n+k+\ell) \neq 0$, and so forth. In this case, the game could proceed indefinitely. Let's call such values of n *irregular degrees*. To see why there are only finitely many irregular degrees, suppose that we can find such $i,j,k,\ell,\dots$. Let (a,b) be the last point of degree n chosen in the game, which occurs on the ith turn; at this point, $H_i(n) = 0$. The fact that $H_i(n+k)$ has not stabilized yet means that at least one point of degree $n+k$ is still in play; call it (c,d). It cannot lie northeast of (a,b), so $c < a$ or $b < d$. Likewise, once $H_j(n+k) = 0$, the fact that $H_j(n+\ell) \neq 0$ means that at least one point of degree $n+\ell$ is still in play; call it (e,f). It cannot lie northeast of (a,b) or of (c,d), so $f < a$ or $e < b$ and $f < c$ or $e < d$. We see that the x- and y-values of these points give us two nonincreasing sequences of natural numbers. Fact 1·51 tells us these sequences must stabilize eventually. Were there infinitely many irregular degrees, we could proceed through these degrees from left to right indefinitely, which would prolong these sequences indefinitely; so, there must be finitely many irregular degrees.

Once we exhaust the last irregular degree, on the ith turn, there are finitely many degrees n with $H_i(n) \neq 0$. As noted, these must all stabilize eventually, which is possible only if the game ends, since whenever $H_i(n) \neq 0$, the players can choose at least one position that would decrease $H_i(n)$. □

Division

Four mathematicians are talking about a problem. They have 11 sheets of scratch paper between them. How many pages will each mathematician get, and how many will be left over? If you answered two sheets for each, with three sheets left over, then you were not only correct,[15] but you were, of course, performing division: 4 is the **divisor**, 3 the **quotient**, and 2 the **remainder**. This illustrates a big difference between division and the other arithmetic

[15] Not really. In my experience, the actual answer would be "two each, *more or less,*" but as often happens in mathematics, we care more about the truth than about reality. That is not a typo!

operations. Addition, subtraction, and multiplication all give *one* result, but division gives *two:* a quotient and a remainder.

It probably won't surprise you that we can always divide two integers.[16]

The Division Theorem. *Let n and d (the* **divisor***) be two integers. If $d \neq 0$, we can find exactly one integer q (the* **quotient***) and exactly one natural number r (the* **remainder***) satisfying the two conditions*

D1) $n = qd + r$, *and*

D2) $r < |d|$.

Try to remember the meaning of "divisor", "quotient", and "remainder", since I'll use them quite a bit from now on. Also try to remember the second criterion, since students have a habit of forgetting it, especially in those moments when it's most useful.

Example 1·53. Division of 12 by 7 gives us a quotient of 1 and a remainder of 5. Division of -12 by 7 gives us a quotient of -2 and a remainder of 2. (You can't use a quotient of -1 and a remainder of -5 because the Division Theorem wants a *nonnegative* remainder.)

Question 1·54 .
Identify the quotient and remainder when dividing:

(a) 10 by -5;

(b) -5 by 10;

(c) -10 by -4.

Proof of the Division Theorem. The proof relies on some concepts we just discussed, such as the well ordering of $\mathbb{N}$. Since it's often easier to think about positive numbers, we consider two cases: $d \in \mathbb{N}^+$ (positive), and $d \in \mathbb{Z}\backslash\mathbb{N}$ (negative). First we consider $d \in \mathbb{N}^+$; by definition of absolute value, $|d| = d$. We must show two things: first, that we can find a quotient q and remainder r; second, that r is unique. We work on each claim separately.

Existence of q and r: First we show that we *can* find q and r that satisfy (D1). Again, we split this into two cases: n nonnegative, and n negative.

First assume n is nonnegative; that is, $n \in \mathbb{N}$. We create a sequence of natural numbers in the following way. Let $r_0 = n$. For $i \in \mathbb{N}^+$ we define

$$r_{i+1} = \begin{cases} r_i - d, & d \leqslant r_i; \\ r_i, & \text{otherwise.} \end{cases} \tag{1.1}$$

We claim this sequence is nondecreasing. Why? If $r_{i+1} \neq r_i$, then by definition $d \leqslant r_i$, in which case

$$r_{i+1} = r_i - d \in \mathbb{N}, \quad \text{which we rewrite as} \quad r_i - r_{i+1} = d \in \mathbb{N}, \quad \text{so} \quad r_i \geqslant r_{i+1}.$$

[16] That's a lie. Find the lie. (*Hint:* It's a subtle detail.)

Fact 1·51 tells us that this sequence of r's must stabilize with a minimal element, r. This must satisfy $r < d$, since otherwise $d \leqslant r$, which would allow us to create a subsequent, *different* r_{i+1}, contradicting the choice of r as the stable one. In addition, the definition of the sequence requires $r \in \mathbb{N}$. Combining them, we see that r satisfies (D2). Let q be the index such that $r_q = r$; a proof by induction shows that $n = qd + r$, satisfying (D1).

Question 1·55 .
Provide this proof of induction. Use induction on q to show that the sequence of natural numbers defined in formula 1.1 satisfies the property $n = qd + r_q$. You'll want to start with $q = 0$.

Proof (continued). Now suppose $n \in \mathbb{Z}\backslash\mathbb{N}$, so n is negative. As $|n|$ is nonnegative, we can apply the previous argument to find q' and r' satisfying (D1) and (D2) for $|n|$. Unfortunately, we need these statements for n, not $|n|$. Fortunately, $n = -|n|$, so we can write

$$n = -|n| = -(q'd + r') = (-q')\,d - r'.$$

Let $q = -(q' + 1)$ and $r = d - r'$; we now have

$$qd + r = [-(q' + 1)]\,d + (d - r') = [(-q')\,d + d] + (d - r') = (-q')\,d - r' = n.$$

Written backwards and condensed, this equation says $n = qd + r$, satisfying (D1) for n. Certainly q is an integer by definition of $\mathbb{Z}$, while $r = d - r'$ is natural because $r' \leqslant d$. So we have $0 \leqslant r$, and $r < d$ from Question 1.27. Combining them, we have $0 \leqslant r < d$, satisfying (D2).

Uniqueness of q and r: Here we have to show that no other combination of an integer q' and a natural number r' satisfy both (D1) and (D2). Suppose to the contrary that there exist $q', r' \in \mathbb{Z}$ such that $n = q'd + r'$ and $0 \leqslant r' < d$. By substitution,

$$\begin{aligned} r' - r &= (n - q'd) - (n - qd) \\ &= (q - q')\,d. \end{aligned} \tag{1.2}$$

Subtraction of integers is closed, so $r' - r \in \mathbb{N}$ and $(q - q')\,d$ are both integers. If $0 = q - q'$, then substitution into equation (1.2) shows that $r - r' = 0$, as desired. If $0 \neq q - q'$, we consider two cases. If $q - q' \in \mathbb{N}^+$, then Question 1.28 tells us that $d \leqslant (q - q')\,d$ (replacing a by d and b by $q - q'$). This gives us

$$0 \leqslant r' - r \leqslant r < d \leqslant (q - q')\,d = r' - r,$$

a contradiction, so $q - q' \notin \mathbb{N}^+$. Likewise, if $q - q'$ is negative, we have $q' - q \in \mathbb{N}^+$, so we play the same game with $r - r'$ to obtain a contradiction. (That is, we negate both sides of equation (1.2).) Hence $q - q' = 0$ and $r - r' = 0$.

We have shown that if $d \in \mathbb{N}^+$, then there exist unique $q, r \in \mathbb{Z}$ satisfying (D1) and (D2). We still have to show that this is true for $d \in \mathbb{Z}\backslash\mathbb{N}$. In this case, $|d| \in \mathbb{N}^+$, so we can apply the former case to find unique $q, r \in \mathbb{Z}$ such that $n = q\,|d| + r$ and $0 \leqslant r < |d|$. By properties of arithmetic, $q\,|d| = q\,(-d) = (-q)\,d$, so $n = (-q)\,d + r$. □

Question 1·56 .
Another way to prove the existence part of the Division Theorem is to form two sets $S = \{n - qd : q \in \mathbb{Z}\}$ and $R = S \cap \mathbb{N}$, prove that $R \neq \varnothing$, and then use the well-ordering property to identify the smallest element of R, which is the remainder from division. Fill in the blanks of Figure 1·5 to see why R is nonempty.

Question 1·57 .
If a and b are both natural numbers, and $0 \leqslant a - b$, then (a) why is $b \leqslant a$? Similarly, if $|d| \leqslant r$, then why are (b) $0 \leqslant r - |d|$ and (c) $r - |d| \leqslant r$?

Notation. If the Division Theorem tells us that the remainder is zero, then we write $d \mid n$. This is shorthand for saying, d **divides** n. For instance, $2 \mid 6$. Try not to confuse this with $6/2$, which means something 6 **divided by** 2. That is a completely different idea.

Question 1·58 .
Prove that if $a \in \mathbb{Z}$, $b \in \mathbb{N}^+$, and $a \mid b$, then $a \leqslant b$.

Definition 1·59. We define lcm, the **least common multiple** of two integers, as

$$\operatorname{lcm}(a, b) = \min\left\{n \in \mathbb{N}^+ : a \mid n \text{ and } b \mid n\right\}.$$

This is a set-builder expression of the definition that you should already be familiar with: it's the smallest (min) positive ($n \in \mathbb{N}^+$) multiple of a and b ($a \mid n$, and $b \mid n$).

Question 1·60 .

(a) Fill in each blank of Figure 1·6 with the justification.

(b) One part of the proof claims that "A similar argument shows that $b \mid r$." State this argument in detail.

The equivalence of the Well-Ordering Principle and Induction

Fact 1·46 claims that the Well-Ordering Principle is equivalent to the Induction Principle.

Why? First we show the Induction Principle implies the Well-Ordering Principle. Assume that the Induction Principle is true, and let S be any subset of $\mathbb{N}$. Recall that $\leqslant$ is a linear ordering of $\mathbb{N}$, so we can compare any two elements of S. If S is finite with n elements, then we can enumerate its elements as $s_1, \ldots, s_n$, and sort them according to $\leqslant$, so we can find a smallest element.

Otherwise, suppose S is infinite. We proceed by induction. If $0 \in S$, then for any $s \in S$, we know that $s - 0 = s$, and s is a natural number, so $0 \leqslant s$. That makes 0 a minimal element. Now let $i \in \mathbb{N}$, and suppose that none of $0, \ldots, i - 1$ is in S, but i is. We claim that i is a

Let $n, d \in \mathbb{Z}$, where $d \in \mathbb{N}^+$. Define $S = \{n - qd : q \in \mathbb{Z}\}$ and $R = S \cap \mathbb{N}$.
Claim: $R \neq \varnothing$.
Proof: We consider two cases.

1. First suppose $n \in \mathbb{N}$.
   (a) Let $q =$____. By definition of $\mathbb{Z}$, $q \in \mathbb{Z}$.
   (You can infer this answer by looking down a couple of lines.)
   (b) By properties of arithmetic, $qd =$____.
   (c) By ____, $n - qd = n$.
   (d) By hypothesis, $n \in$____.
   (e) By ____, $n - qd \in \mathbb{N}$.
2. It's possible that $n \notin \mathbb{N}$, so now let's assume that, instead.
   (a) Let $q =$____. By definition of $\mathbb{Z}$, $q \in \mathbb{Z}$.
   (Again, you can infer this answer by looking down.)
   (b) By substitution, $n - qd =$____.
   (c) By ____, $n - qd = -n(d - 1)$.
   (d) By ____, $n \notin \mathbb{N}$, but it is in $\mathbb{Z}$. Hence, $-n \in \mathbb{N}^+$.
   (e) Also by ____, $d \in \mathbb{N}^+$, so arithmetic tells us that $d - 1 \in \mathbb{N}$.
   (f) Arithmetic now tells us that $-n(d - 1) \in \mathbb{N}$. (pos×natural=natural)
   (g) By ____, $n - qd \in \mathbb{N}$.
3. In both cases, we showed that $n - qd \in \mathbb{N}$. By definition of ____, $n - qd \in S$.
4. By definition of ____, $n - qd \in S \cap \mathbb{N}$.
5. By definition of ____, $S \cap \mathbb{N} \neq \varnothing$. Hence $R \neq \varnothing$.

Figure 1·5: Material for Question 1.56

Let $a, b, c \in \mathbb{Z}$.
Claim: If a and b both divide c, then $\text{lcm}(a, b)$ also divides c.
Proof:

1. Let $d = \text{lcm}(a, b)$. By _____, we can choose q, r such that $c = qd + r$ and $0 \leqslant r < d$.
2. By definition of _____, both a and b divide d.
3. By definition of _____, we can find $x, y \in \mathbb{Z}$ such that $c = ax$ and $d = ay$.
4. By _____, $ax = q(ay) + r$.
5. By _____, $r = a(x - qy)$.
6. By definition of _____, $a \mid r$. A similar argument shows that $b \mid r$.
7. We have shown that a and b divide r. Recall that $0 \leqslant r < d$, and _____. By definition of lcm, $r = 0$.
8. By _____, $c = qd = q\text{lcm}(a, b)$.
9. By definition of _____, $\text{lcm}(a, b)$ divides c.

Figure 1·6: Material for Question 1.60

minimal element of S. To see why, consider the set $T = \{s - i : s \in S\}$. This is also a subset of $\mathbb{N}$, because the definition of subtraction tells us $s - i \notin \mathbb{N}$ only when $s \in \{0, \ldots, i-1\}$, and none of those numbers is in S by hypothesis. In addition, $0 \in T$ because $i \in S$, so putting $s = i$ in the definition of T gives us $i - i \in T$. We already showed that any subset of $\mathbb{N}$ that contains 0 has 0 as a least element, so 0 is the least element of T. We return to S. Let $s \in S$; we claim that $i \leqslant s$. To see why, consider $i - i = 0$ and $s - i$. As previously discussed, both elements are in T, and $0 \leqslant s - i$. This is true if and only if $0 + i \leqslant (s - i) + i$, or, $i \leqslant s$. Since s was arbitrary, i is indeed the smallest element of S.

We have shown that any subset S of $\mathbb{N}$ has a smallest element with respect to $\leqslant$. This proves that $\mathbb{N}$ is well ordered by $\leqslant$.

Now we show that the Well-Ordering Principle implies the Induction Principle. Assume that the Well-Ordering Principle is true, and let $S \subseteq \mathbb{N}$, satisfying both the inductive hypothesis and the inductive step. Let N be the set of *all* such natural numbers that are not in S. If $N = \varnothing$, we are done. Otherwise, the Well-Ordering Principle tells us N has a smallest element, which we call n. We cannot have $n = 0$, as that would violate the inductive hypothesis, which we assumed was true. Hence $n \neq 0$, which means $n - 1 \in \mathbb{N}$. The choice of n as the *smallest* element of N implies that $n - 1 \in S$, since after all $n - 1 < n$ (this is easy to see if you think about the definition). However, we also assumed S satisfies the inductive step, so $(n-1) + 1 \in S$, but $n = (n-1) + 1$, contradicting the hypothesis that $n \in N$. Hence $N = \varnothing$, and $S = \mathbb{N}$. □

1·5 Division on the lattice (optional)

> *Algebra is nothing more than geometry, in words;*
> *geometry is nothing more than algebra, in pictures.*
> — Sophie Germain

We have shown that we can divide both integers and natural numbers to obtain a quotient and remainder. Can we divide on the lattice, identifying a quotient *and* a remainder? If so, is the result unique?

We can in fact perform division on the lattice. To do that, we first have to think about the other operations: addition, subtraction, and multiplication. Let $P = (p, q)$ and $R = (r, s)$ be points on L. We'll define addition in a natural way,

$$P + R = (p + r, q + s)\,.$$

For subtraction, use

$$P - R = (p - r, q - s)\,,$$

but notice that this doesn't always give us a point in the *natural* lattice. So, let's expand our view to the *integer* lattice, $\mathbb{Z}^2$; as with division of natural numbers, we can work first in $\mathbb{Z}^2$, then switch back to $\mathbb{N}^2$ once that's out of the way.

What of multiplication? Since the lattice is two-dimensional, we'd like multiplication to move us in two dimensions. We adopt the following convention:

- $(p, q) \cdot (c, 0) = (pc, qc)$, the point on the line that connects the origin to (p, q), but with a length c times that from the origin to (p, q);

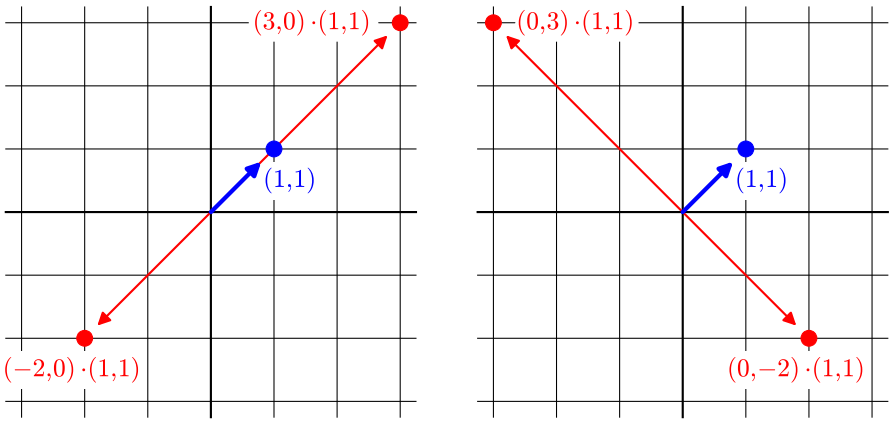


Figure 1·7: Multiplication on the lattice.

- $(p,q)\cdot(0,d) = (-qd,pd)$, the point on the line *perpendicular* to the line that connects the origin to (p,q), but with a length d times that from the origin to (p,q);
- $(p,q)\cdot(c,d) = (p,q)\cdot(c,0)+(p,q)\cdot(0,d) = (pc-qd,pd+qc)$, the *vector sum* of the previous two.

See Figure 1.61. This may look odd, but it extends well to other problems, as you will learn later.

Question 1·61 .
Suppose $P = (3,1)$.

(a) Calculate $P\cdot(c,0)$ for several different values of c. Sketch the resulting points on $\mathbb{Z}^2$. Observe how the results conform to the description in the text.

(b) With the same value of P, calculate $P\cdot(0,d)$ for several different values of d. Sketch the resulting points on $\mathbb{Z}^2$. Observe how the results conform to the description in the text.

(c) With the same value of P, calculate $P\cdot(c,d)$ for several different combinations of values of c and d that you used in parts (a) and (b). Sketch the resulting points on $\mathbb{Z}^2$. How would you describe the results geometrically?

As with division of natural numbers, the goal of dividing $P = (p,q)$ by $D = (c,d)$ will be to move D "closer and closer" to P via subtraction from P, until the remaining distance is so small that subtraction no longer makes it smaller. If we measure our distance with integers, we can then apply the Well Ordering Principle via Fact 1·51 to guarantee the division ends.

But how can we measure distance with *integers*? The traditional distance formula is based on the Pythagorean Theorem, and relies on radicals:

$$\text{the distance between } (p,q) \text{ and } (r,s) \quad \text{is} \quad \sqrt{(p-r)^2+(q-s)^2}.$$

That's bad, because Fact 1·51 does not apply to radicals. For instance, the sequence

$$\sqrt{\frac{1}{2}} > \sqrt{\frac{1}{3}} > \sqrt{\frac{1}{4}} > \sqrt{\frac{1}{5}} > \cdots$$

consists of positive numbers, and continues indefinitely.

Don't let that discourage you! It's actually easy to get around this; we'll just use a different distance formula, modifying the traditional one so that it *doesn't* use radicals,

$$\text{the "square distance" between } (p,q) \text{ and } (r,s) \quad \text{is} \quad (p-r)^2 + (q-s)^2 .$$

The square distance is always natural, opening the way to use Fact 1·51. It's a bit tedious to write "square distance" all the time, so we'll write $\operatorname{sqd}(P,Q)$ for the square distance between P and Q. We consider the distance from a point to the origin to be its *size,* much like *absolute value,* so we will write $\|R\|_{\text{sq}}$ to indicate this value.

The Division Theorem for the lattice. *Let N and D be two points of $\mathbb{Z} \times \mathbb{Z}$. If $D \neq 0$, we can find $Q \in \mathbb{Z} \times \mathbb{Z}$ (the* **quotient***) and $R \in \mathbb{N} \times \mathbb{N}$ (the* **remainder***) satisfying the two conditions*

D1) $N = QD + R$, and

D2) $0 \leqslant \|R\|_{\text{sq}} < \|D\|_{\text{sq}}$.

. *However, the points Q and R may not be unique.*

Proof. Let $S_0 = N$, and for $i \in \mathbb{N}^+$ define $S_i = N - (i,0)\,D$. Let $\mathcal{S} = \left\{ \|S_i\|_{\text{sq}} : i \in \mathbb{N} \right\}$. This is a set of natural numbers, so by the well ordering of $\mathbb{N}$, $\mathcal{S}$ has a smallest element, corresponding to a particular S_a. Let $T_0 = S_a$. For $j \in \mathbb{N}^+$, define $T_j = S_0 - (a,j)\,D$. Let $\mathcal{T} = \left\{ \|T_j\|_{\text{sq}} : j \in \mathbb{N} \right\}$. It also has a smallest element, corresponding to a particular T_b. Let $Q = (a,b)$ and $R = T_b$; by definition and substitution, we have $R = N - Q \cdot D$, or $N = QD + R$. This satisfies (D1).

To show that Q and R also satisfy (D2), suppose the contrary, that is, $\|R\|_{\text{sq}} \geqslant \|D\|_{\text{sq}}$. The set $\mathcal{U} = \{[Q \pm (1,0)] \cdot D, [Q \pm (0,1)] \cdot D\}$ is finite, so one of its points has a distance to N that is no larger than the other three. Now consider Figure 1·8. At least two points of $\mathcal{U}$ form angles with the line $N - QD$ that are no larger than 90°.

We consider two cases. If $\alpha = \beta = 45^\circ$, the Law of Cosines tells us

$$a^2 = b^2 = d^2 + (d+k)^2 - \sqrt{2} \cdot d\,(d+k) .$$

We assumed $\|R\|_{\text{sq}} \geqslant \|D\|_{\text{sq}}$. By substitution, $\|N - QD\|_{\text{sq}} \geqslant \|D\|_{\text{sq}}$. We chose Q to minimize the square distance between QD and N, so $\|N - QD\|_{\text{sq}} \leqslant a^2$. By substitution,

$$\left(d + k^2\right) \leqslant d^2 + (d+k)^2 - \sqrt{2} \cdot d\,(d+k) .$$

Rewrite this as

$$0 \leqslant d^2 - \sqrt{2} \cdot d\,(d+k) .$$

Since d is positive, we can rewrite again as

$$0 \leqslant d - (d+k)\sqrt{2},$$

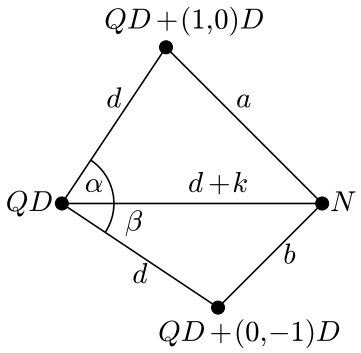


Figure 1·8: Illustration of the proof of existence for the Division Theorem in $\mathbb{N}^2$. Let the Euclidean distance from QD to $QD + (1,0)\,D$ and to $QD + (0,-1)\,D$ be d. Suppose d is smaller than the square distance from QD to N, which is then $d + k$ with some positive k. In this diagram, α and β form acute angles between two extensions from QD to the segment joining QD and N. Our task is to show that one of a or b is less than $d + k$, contradicting the choice of Q to minimize this distance.

but $k \geqslant 0$ implies that $d - (d+k)\sqrt{2} < 0$, contradicting the choice of Q.

In the case that $\alpha, \beta \neq 45^\circ$, one of the segments lengthens, while the other shortens, making it smaller than $d + k$; hence, one of them is closer to N than QD, contradicting the choice of Q.

Figure 1·8 also hints at why we might have *two* distinct quotients and remainders of the same size. If two possible remainders are $(1,0)$ and $(0,1)$, with $\|D\|_{\text{sq}} > 1$, we cannot get closer, and either solution works. □

We can thus extend the notion of "division" that we gave above to *anything* we can "view" as an integer lattice.

Question 1·62 .

Suppose $N = (10,4)$.

(a) Let $D = (3,1)$, and $R = N - (3,0) \cdot D$. Show that $\|R\|_{\text{sq}} < \|D\|_{\text{sq}}$.

(b) Let $D = (1,3)$, and $R = N - (3,-3) \cdot D$. Show that $\|R\|_{\text{sq}} < \|D\|_{\text{sq}}$.

(c) Explain how the results of parts (a) and (b) conform to those described in the text.

(d) Suppose $N = (10,4)$ and $D = (2,2)$. Find $Q \in L$ such that if $R = N - Q \cdot D$, then $\|R\|_{\text{sq}} < \|D\|_{\text{sq}}$. Sketch the geometry of N, D, QD, and R.

(e) Is the result unique? That is, could you have found $Q' \in L$ such that $R = N - Q' \cdot D$, $\|R\|_{\text{sq}} < \|D\|_{\text{sq}}$, and $Q' \neq Q$?

(f) Show that for any $N, D \in L$ with $D \neq (0,0)$, you can find $Q, R \in L$ such that $N = Q \cdot D + R$ and $\|R\|_{\text{sq}} < \|D\|_{\text{sq}}$. Again, try to build on the geometric ideas you gave in (e).

1·6 Polynomial division

He who can properly define and divide is considered to be a god.
— Plato

You may be wondering how the material we've studied so far is related to algebra, which you probably associate more with polynomials than with games. Take a look back at Ideal Nim's playfield, the lattice $\mathbb{N}^2$. This game is related to polynomials, or at least to **monomials**, which are products of variables.

- Any point (a, b) on the lattice corresponds in unique fashion to a monomial in two variables, $x^a y^b$.
- The choice (a, b) disqualifies other points (c, d); we called them Gone from Gameplay. The rule was that (c, d) is Gone from Gameplay if $a \leqslant c$ and $b \leqslant d$. In this case, the corresponding monomial $x^c y^d$ is *divisible* by $x^a y^b$.
- Just as the lex ordering $\preceq_{\text{lex}}$ is a well ordering of $\mathbb{N}^2$, it is a well ordering of monomials in two variables.

By this reasoning, we could extend division with quotient and remainder on the lattice to define division with quotient and remainder of monomials. Whether such a division with remainder is useful, we leave to others to ponder; we merely point out that it exists.

Question 1·63 .
If you are so inclined, however, translate the results of Question 1.62 to monomials. "Multiplication" and "subtraction" in the Division Theorem actually translate to what operations on monomials?

We turn instead to division with quotient and remainder of polynomials. When one polynomial is a multiple of another, we would like the quotient and remainder to be consistent with previous choices. For instance, dividing $(x + 1)(x - 1)$ by $x - 1$ should clearly give us a quotient of $x + 1$ and a remainder of 0.

We would also like to replicate the distinct behavior of integer division: that is, the remainder r should in some manner be "smaller" than the divisor g. This isn't too hard to grasp if you think about what comes naturally: we want to subtract multiples of g in such a way as to make f smaller.

Example 1·64. Suppose $f = x^2 + 1$ and $g = x - 1$. A natural way to make f "smaller" is to multiply $x - 1$ by x and subtract:

$$f - xg = \left(x^2 + 1\right) - x(x - 1) = x + 1.$$

We end up with $x + 1$ as a remainder.

That hardly seems complete, as we can subtract *another* multiple of $x - 1$:

$$(f - xg) - g = (x + 1) - (x - 1) = 2.$$

Putting them together, we have

$$f = (x + 1)g + 2.$$

We now have 2 as our remainder, and it is not possible to remove any more multiples of $x - 1$.

What the example should show you is that we make f "smaller" by reducing its largest exponent. We call this largest exponent the **degree** of a polynomial. The degree makes a "natural" target, not only because it seems to shrink the polynomial, but also because it relates polynomial division to the Well Ordering Principle, which we used to set up division on both the integers and the lattice.

We have to be a little careful here: what is deg 0? You might be tempted to say that $\deg 0 = 1$ because 0 is a constant, so $0 = 0 \cdot 1 = 0 \cdot x^0$, but we could just as easily say that $0 = 0 \cdot x^1$ or $0 = 0 \cdot x^2$ or … You get the idea. To avoid this pickle, we agree that the term "degree" applies only to *nonzero* polynomials, and that *the zero polynomial has no degree.*

Another complication lies hidden in the weeds. It isn't too hard to divide 7 by 5, but what do we do with $7x$ and $5x$? Writing $7x = 5 \cdot x + 2x$ does not decrease the degree, and the joy of decreasing the coefficient evaporates when we realize that we can decrease the coefficient even more by writing $7x = 5 \cdot x + 0x$. In any case, this problem grows even more annoying when dividing binomials, trinomials, and so forth. We will content ourselves to restrict divisors to polynomials whose leading coefficients is 1. We call such polynomials **monic**.

The Division Theorem for polynomials. *Let f, g be polynomials in one variable with integer coefficients, with the leading coefficient of g being 1. We can find exactly one polynomial q and exactly one polynomial r, also with integer coefficients, satisfying the conditions*

D1) $f = qg + r$, and

D2) $r = 0$ or $0 \leqslant \deg r < \deg g$.

Proof. First we show that *some* sort of quotient and remainder exist. If $\deg f < \deg g$, then let $r = f$ and $q = 0$; this satisfies both properties. For the case $\deg f \geqslant \deg g$, we proceed by induction on the difference in degree.

Question 1·65 .
Suppose that $f = 3x^2 + 2$ and $g = x^2 - x + 2$. These have the same degree, so we can subtract from f a constant multiple of g to obtain a remainder of smaller degree. Do that, then use the result to follow through the next paragraph of the proof.

Continuation of proof. For the inductive base, assume $\deg f - \deg g = 0$; that is, the polynomials have the same degree. Write c for the leading coefficient of f. Let $q = c$, and $r = f - cg$. We have

$$qg + r = cg + (f - cg) = f,$$

satisfying (D1). In addition, if $\deg f = d$, we can write $f = cx^d + f'$ and $g = x^d + g'$, where the degrees of f' and g' are smaller than those of f and g, respectively. That gives us

$$f - cg = \left(cx^d + f'\right) - c\left(x^d + g'\right) = \left(\cancel{cx^d} + f'\right) - \left(\cancel{cx^d} + cg'\right) = f' - cg'.$$

We may not know the degree of $f' - cg'$ with precision, but we can say that it's smaller than d. Since $\deg g = \deg f = d$, either $r = 0$ or $0 \leqslant \deg r < \deg d$, satisfying (D2).

Question 1·66 .
Suppose that $f = 2x^3 + x^2 + 4x + 2$ and $g = x^2 - x + 2$. These have different degree. Subtract from f a polynomial multiple of g to obtain a remainder of smaller degree. Do that, then use the result to follow through the next paragraph of the proof. The remainder should look quite familiar.

Continuation of proof. Now assume that the claim holds for $\deg f - \deg g = i$ whenever $i = 0, 1, 2, \ldots, n-1$. What about $i = n$? Again, write c for the leading coefficient of f. Let $q' = cx^n$, and $r' = f - q'g$. As before, if $\deg f = d$, we can write $f = cx^d + f'$, where $\deg f' < d$. If $\deg g = a$, we can write $g = x^a + g'$, where $\deg g' < a$. We have

$$r' = \left(cx^d + f'\right) - cx^n\left(x^a + g'\right) = \left(cx^d + f'\right) - \left(cx^{a+n} + cx^n g'\right).$$

Recall that $a + n = \deg g + (\deg f - \deg g) = \deg f = d$, so substitution gives us

$$r' = \left(\cancel{cx^d} + f'\right) - \left(\cancel{cx^d} + cx^n g'\right) = f' - cx^n g'.$$

We already pointed out that $\deg f' < \deg d$; we also have $\deg\left(cx^n g'\right) = n + \deg g' < n + a = d$. Again, $r = 0$ or $\deg r' < \deg f = n$. If $r = 0$, we are done, so suppose $r \neq 0$, in which case $\deg r < n$. By the inductive hypothesis, we can find q'' and r'' such that $r' = q''g + r''$ and $\deg r'' < \deg g$. By substitution and rewriting,

$$f = q'g + r' = q'g + \left(q''g + r''\right) = \left(q' + q''\right)g + r''.$$

Let $q = q' + q''$ and $r = r''$, and we satisfy both (D1) and (D2).

How about the result's *uniqueness?* Suppose we can find polynomials q_1, q_2, r_1, and r_2 such that

$$f = q_1 g + r_1 = q_2 g + r_2 \quad \text{and} \quad \text{for } i = 1, 2\ r_i = 0 \text{ or } 0 \leqslant \deg r_i < \deg g.$$

Rewrite the first equations as

$$\left(q_1 - q_2\right)g = r_2 - r_1.$$

If the polynomial on the left is nonzero, then its degree is no smaller than $\deg g$. If the polynomial on the right is nonzero, then its degree is smaller than $\deg g$. It's not possible to have a nonzero polynomial with two different degrees — the definition of degree is unambiguous — so the polynomials must be zero. That means $r_1 = r_2$, which forces $q_1 = q_2$; otherwise, the degree on the left would be nonzero. □

As with integer division, the proof of this theorem outlines an algorithm to compute the quotient and remainder. (An **algorithm** is a finite list of instructions with a well-specified output, which is guaranteed to terminate after finitely many operations.) We know that the method will end after finitely many steps, because the degrees of the remainders form a decreasing sequence of natural numbers, and the well-ordering applies. Indeed, this algorithm is sometimes called "long division" of polynomials.

Question 1·67 .
Divide $f = 10x^6 - 3x^4 + 1$ by $g = x^3 + x + 1$.

In all these questions, both f and g are polynomials with integer coefficients.

Question 1·68 .

Sometimes we *can* divide f by a non-monic g, if we're willing to surrender the requirements that the resulting quotient and remainder have integer coefficients. Can you find an example where g is non-monic, but the quotient and remainder *do* have integer coefficients? Try to find a non-trivial example; that is, you should have $f \neq 0$ and $f \neq qg$ for any polynomial q.

Question 1·69 .

The **Factor Theorem** claims that if we divide f by g and have a zero remainder, then any root of g is a root of f. Why is this true?

(A **root** of a polynomial g is any value a of x such that $g(a) = 0$. So the problem is really asking why $g(a) = 0$ implies $f(a) = 0$ under the given hypothesis.)

Question 1·70 .

The **Remainder Theorem** claims that if we let a be any integer, and divide f by $g = x - a$, the remainder is a constant that has the same value as $f(a)$. Why is this true?

Question 1·71 .

The Division Theorem requires that the polynomials be in one variable only. What if two polynomials have two or more variables? You first have to decide how to determine a leading monomial; for instance, what should be the leading monomial of $x^2 + xy + y^2$?

(a) Describe a way of choosing a leading monomial.

(b) Try to divide polynomials in several variables. Use several examples. Are you able to identify a quotient and remainder that satisfy (MD1) $f = qg + r$ and (MD2) $r = 0$ or r is somehow smaller than g? (You have to explain how r is smaller.)

(c) If it does work, describe a Multivariate Polynomial Division Theorem, and try to prove it. If it doesn't work, explain why not.

Question 1·72 .

Where in this chapter did Noetherian behavior show up? List as many places as you can; I can think of four off the top of my head. (Go ahead and count those occasions where we explained how one system could be viewed as a more fundamental system, since that really does count.)

Chapter 2

Algebraic systems and structures

In the previous chapter, we using decreasing sequences of natural numbers to formulate division in several different contexts. We already pointed out that division is rather unusual as an operation, because rather than producing only one result, it produces *two*, the quotient *and* a remainder.

Many mathematics courses treat division differently: they ignore remainders, and treat quotient exclusively as a position on the real line. This can give students the impression that remainders are a mostly useless artifact. In fact, it is often the case that the *quotient* is useless, and what *really* matters is the remainder!

That is mostly the case in the course, and this chapter will use remainders as an example to introduce you to some very elegant properties, as well as to one of the most elegant and useful objects ever devised, the *finite field*.

2·1 From symmetry to arithmetic

> *Those who assert that the mathematical sciences say nothing of the beautiful or the good are in error. … The chief forms of beauty are order and symmetry and definiteness, which the mathematical sciences demonstrate in a special degree.*
>
> — Aristotle, *Metaphysics,* Book XIII

We return a few moments to Nim and Ideal Nim, as a fun way to help motivate some material that follows. In this section, we want to consider the question,

How do we decide what makes a "winning move" for Nim or Ideal Nim?

"Nimbers"

If you've played Nim enough, you'll notice that whenever David faces two rows with an equal number of pebbles, say

he might as well give up: if he sees a *visually* symmetric game in two rows, then Emmy can undo any move he makes. Of course, interesting games are not visually symmetric; for them Emmy wants to impose a *chronological* symmetry. By this we mean that Emmy wants her first move to change the game in such a way that whatever David does, she can undo — not visually, but in such a way that she always has an advantage, and can explain why, in the same way that the two equal rows above are visually symmetric.

To do this, we'll assign values to different game configurations. We call these values **nimbers** because they are numbers that correspond to different configurations of nim.

It makes sense to say that a game with no moves left has value 0. A game with 1 pebble left is not equivalent to a game with value 0; it makes sense to call its value 1. A game with 2 pebbles in *one* row is equivalent to neither a game with 1 pebble nor a game with 2 pebbles; it makes sense to call its value 2. And so forth; we'll agree that a game with n pebbles in *one* row is has value n. This covers all games of Nim with just one row.

What about games with two rows? We've already run out of numbers, so it might help to look at the game above differently. The first thing to remember is that whenever a player sees that configuration, he might as well give up, because anything he does will break the symmetry, and the other player can easily restore the symmetry. So in some sense, that game is equivalent to a game with no moves left.

Mind the vocabulary! We did not say the game above is *equal* to a game with no moves left; it plainly is not, as it has quite a few moves left. We said it is *equivalent* to a game with no moves yet. We will dignify this situation with a special term: a **zero game** is either a game with no moves left at all, or a game where no matter what the current player does, the other player will always have a winning strategy.

Notice that if Emmy's turn ends with a zero game, then David has no way to end his turn with a zero game. (After all, if he could, then Emmy could not herself have finished her turn with a way that guaranteed her a win.)

In addition to viewing the game above as *one* game of Nim with two rows of four pebbles, we can look at it as a "sum" of *two* games, each with *one* row of *four* pebbles. This makes sense inasmuch as moving in one row doesn't affect the number of pebbles in the other row, so it really is as if you're playing two games at the same time. This point of view can help us break harder problems into smaller ones, always a goal in mathematics.

This insight gives us two guidelines, which we turn into definitions:

- A general game of Nim has value 0 if and only if it is a zero game; that is, if and only if any move the current player makes has a response by the other player that results in a zero game. If the game has value x, we indicate this by $x \equiv 0$.
- If a game of Nim has value x, another game has value x if and only if their sum (playing both games together as one) is a zero game. Using $\equiv$ to indicate the values of equivalent games, we write
$$x + y \equiv 0 \qquad \text{if and only if} \qquad x \equiv y.$$

We can use these definitions to define the value of the sum of two games.

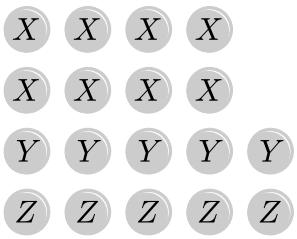
Definition 2·1. Suppose two games have values x and y. We say that $x + y \equiv z$ if and only if $(x + y) + z \equiv 0$; that is, playing all three games as one single game guarantees a win for the second player.

Example 2·2. Let X be a game of Nim with two rows of 4 pebbles each. We have already pointed out that the value of X is $x \equiv 0$.

Next let Y be a game of Nim with one row of 5 pebbles. From what we argued above, the value of Y is $y \equiv 5$.

Finally, let Z be a game of Nim with one row of 5 pebbles. From what we argued above, the value of Z is $z \equiv 5$.

Clearly $y + z = 0$, since a game with two rows of 5 pebbles each has value 0. Likewise, $x + (y + z) \equiv 0 + 0 \equiv 0$. That is, the following game has value 0:



You should be able to verify that this games is, indeed, a win for the second player: any move Emmy makes in one row of the X game, David can reply symmetrically in the other row of the X game; whereas any move she makes in the Y game, David can reply symmetrically in the Z game (and vice versa).

So far, so good. What about the following game?

In this case, we have a row of 1 pebble and a row of n pebbles. Does it makes sense to say that its *equivalent* value is $n + 1$? Not always! To start with, we already know that $1 + 1 \not\equiv 2$, because our rule $x + x \equiv 0$ implies $1 + 1 \equiv 0$. On the other hand, we can verify that $1 + 2 \equiv 3$ and, oddly enough, $1 + 3 \equiv 2$.

Question 2·3 .

Verify that $1 + 2 \equiv 3$ and $1 + 3 \equiv 2$ by playing:

(a) a game with two rows of 1 pebble and 2 pebbles at the same time as a game with 3 pebbles — or, one game with three rows of 1, 2, and 3 pebbles — and showing that no matter how Emmy starts, David can always win; and then,

(b) arguing that part (a) shows both $1 + 2 \equiv 3$ and $1 + 3 \equiv 2$, perhaps moving some rows around to make it obvious.

Question 2·4 .
Explain why addition of Nim games is always commutative; that is, if x and y are the values of games X and Y, then $x + y \equiv y + x$. It may be helpful to move rows around in a game, so as to obtain symmetry.

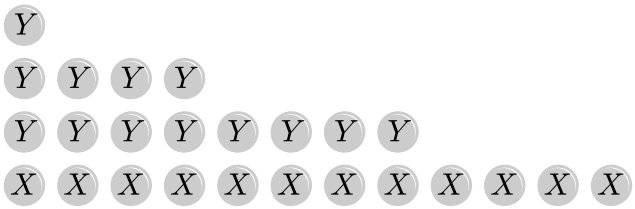
Question 2·5 .
Explain why addition of Nim games is always associative; that is, if x, y, and z are the values of games X, Y, and Z, then $(x + y) + z \equiv x + (y + z)$.

This pattern continues indefinitely; that is, if n is even, then $1+n \equiv 1+n$ and $1+(1+n) \equiv n$. Rather than prove that, however, we argue for something more general.

Nimber equivalence

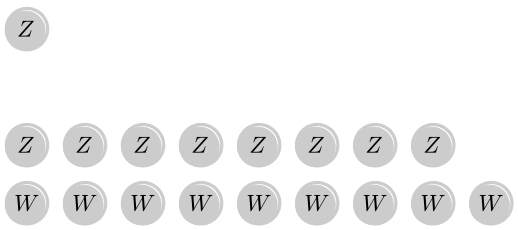
Lemma 2·6. *Suppose a Nim game X has value x. Write x as a sum of powers of 2: $x = a_n 2^n + a_{i-1}2^{i-1} + \cdots + a_1 2 + a_0$ where $a_n = 1$ and the remaining a_i satisfy $a_i \in \{0, 1\}$. Then X is equivalent to the game Y played with n rows and $a_i \times 2^i$ pebbles in the i-th row.*

We'll use an example to illustrate the idea behind the proof of the lemma. Suppose X has value $x = 13 = 8 + 4 + 1 = 2^3 + 2^2 + 2^0$. The lemma claims that the sum of X and a 3-row game Y of 1, 4, and 8 pebbles, is a zero game.

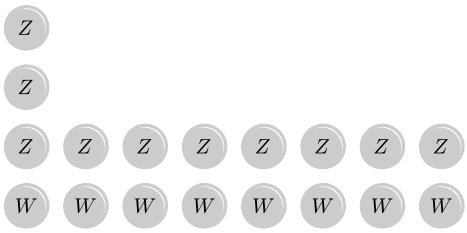


Assume we know the lemma is true for *every x less than* 13, and let's think about what happens when we remove something from X.

If Emmy removes 4 pebbles from X, she reduces X to the game W with value $w = 9 > 2^3$. David finds that there is a row in Y with 4 pebbles, and removes that entire row, obtaining a 2-row game Z of 1 and 8 pebbles. We now have the situation where $w = 9 = 8 + 1$, which corresponds precisely to Z. In other words, David leaves a zero game.

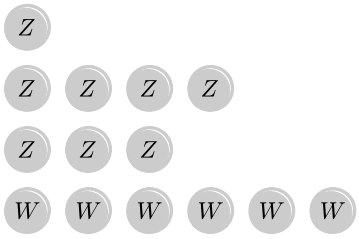


What if Emmy removes 5 pebbles from X, instead? This reduces X to the game W with value $w = 8 = 2^3$. In this case, David cannot reduce Y to one row of 8 pebbles in one move. Instead, he has to look at ways to reduce Y to a sum of rows that adds to 8; that is, he has to create some cancellations which yield 8. He can do this by remove enough pebbles from the next-smaller row of length 4 to add up to the sum of the remaining rows. In this case, he should take 3 pebbles from the middle row of Y, reducing it to the 3-row game Z of 1, 1, and 8 pebbles.



We see immediately from the game's visual symmetry that David leaves a zero game.

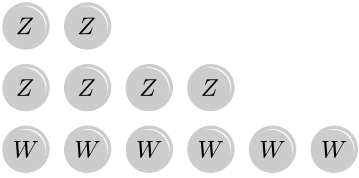
Finally, what if Emmy removes 7 pebbles from X? This reduces X to the game W with value $w = 6 < 2^3$. In this case, David can remove pebbles from the longest row of Y in such a way that it cancels with shorter rows to obtain w. In this case, he should take 10 pebbles from the longest row of Y, reducing it to the 3-row game of Z of 1, 4, and 3 pebbles.



To see that this does indeed work out to 0, we could explore every possible move, but that would take far to long. It's much better to argue that the value of Z is indeed $z = 6$. To do this, we rely on the commutative and associative properties, as well as the fact that we already know $1 + 3 \equiv 2$ and the inductive assumption that the claim above is true for every game whose value is smaller than 13:

$$z = 1 + 4 + 3 \equiv (1 + 3) + 4 \underset{\text{shown}}{\equiv} 2 + 4 = 2 + 2^2 \underset{6<13}{\equiv} 6.$$

That is, the new game is really equivalent to this one:



This process of simplifying Z from a game whose rows are not powers of 2 to one whose rows *are* powers of 2 we will call **row simplification**, and when a game is arranged in rows of powers of 2, we call it **row simplified**. Because this is a matter of game equivalence, it is free *whenever we know it is true;* a player can rearrange several rows in this fashion at any time.

Proof of Lemma 2·6. Let X and Y be the games described in the Lemma. The lemma's claim is that $X + Y \equiv 0$; that is, if Emmy and David play $X + Y$, then David always wins. To see that this is true, assume that the analogous statement is true for any game whose value is less than x, and suppose Emmy removes r pebbles from X, reducing it to the game W, which has value $w = x - r$. Before proceeding, observe that Y is currently row simplified, because it is already arranged in rows whose lengths are powers of 2.

If r is a power of 2, David can reduce Y to Z by removing r pebbles from the shortest row that is at least r pebbles long. (So if $r = 2$ and there is a row of length 2, he removes it from that row; otherwise, if there is a row of length 4, he removes it from that row, and so forth.) Recall that Y was row simplified before David moved. If he removes an entire row, as he did in the example above with $w = 9$, it is obvious that Z remains row simplified; he has simply removed a row. Otherwise, David has removed from one row fewer than half the number of pebbles, and some row simplification is in order — but this does not affect any other existing row. To see why, observe that $r = 2^k$ for some natural k, and we are removing r from a row whose length is 2^ℓ, with $\ell > k$. The result is a row whose length is $2^\ell - 2^k$; just as with subtracting 1 from a power of 10, this simplifies to $2^{\ell-1} + 2^{\ell-2} + \cdots + 2^{k+1} + 2^k$. Row simplification can only affect rows 2^ℓ through 2^k, but the only reason we chose row 2^ℓ is that there were no pebbles in rows $2^{\ell-1}$ through row 2^k, so even if we do this, Z remains row simplified.

From here on, assume r is not a power of 2. If $w = x - r \geqslant 2^i$, let $s = w - 2^i < 2^i$. This is the "surplus" in W above the longest row of Y; that is, that $w = 2^i + s$. (In the example above where $w = 9$, we have $s = w - 2^3 = 1$.) All David needs to do is remove pebbles from one of the remaining rows of Y such that he reduces it to the game Z whose value is $z = s$. The inductive hypothesis tells us the lemma is true for values less than x, and $s < x$, so David should start by considering the shortest row of Y that is longer than s; suppose its length is 2^j. He can remove enough pebbles from this row to cancel any powers of 2 smaller than s that do not already appear in Y; suppose this reduces the row 2^j to k. David should reduce the row this way only if there are no rows between this row of length 2^j and the one of length 2^i. If there are, David should instead "cascade" his choice to longer rows of Y *except* the longest, rewriting $2^{j+1} = 2^j + [k + (2^j - k)]$, $2^{j+2} = (2^{j+1} + 2^j) + [k + (2^j - k)]$, and so forth, so that now the rows of length 2^j, 2^{j+1}, etc. also cancel, leaving k as needed. This cascade will end in the last row of Y before the longest, allowing David to reduce the rows shorter than 2^i to the value s, obtaining a row-simplified game Z of value $2^i + s = w$, as claimed.

Finally, suppose $w = x - r < 2^i$. (We see this in the last example above, with $w = 6$.) In this case, remove from the longest row of Y enough pebbles so that, after row simplification with the smaller rows, the modified Y now has value w: canceling powers of 2 that appear in y but not in w, and preserving those that do appear, as claimed.

Suppose on the other hand that Emmy moves in the Y game, instead of X. David need merely the make the corresponding move in X that was described in the cases above. This covers all the cases, and we have shown that $X + Y \equiv 0$. □

Question 2·7 .
To see why this trick with powers of 2 not work with higher powers, suppose you try to write a Nim game which consists or one row of 2 pebbles as a series of rows of the powers of 3 that add up to 2. What goes wrong? (This also show why it won't work with larger powers, either.)

Nimber addition

Now we have determined that we can always write the value of a game as sums of powers of 2, it becomes easy to add and subtract them.

Nim Addition Theorem. *Let x and y be the respective values of Nim games X and Y. We know $x + x \equiv 0$, so without loss of generality, suppose $x > y$, suppose $2^i \leqslant x < 2^{i+1}$, and $2^j \leqslant y < 2^{j+1}$. Choose r and s such that $x = 2^i + r$ and $y = 2^j + s$. Then:*

- *if $i = j$, then $x + y \equiv r + s$;*
- *otherwise, $x + y \equiv 2^i + (r + y)$.*

Proof. First we claim that $r < 2^i$ and $s < 2^j$. If one of them were false, say $r \geqslant 2^i$, then we could write $r = 2^i + t$ and then $x = 2^i + (2^i + t) = 2 \times 2^i + t = 2^{i+1} + t$. This contradicts the choice of i as the largest power of 2 that is smaller than x.

We now consider the sum $x + y = (2^i + r) + (2^j + s)$. By Questions 2.4 and 2.5, the commutative and associative properties hold. If $i = j$, they allow us to rewrite the sum as

$$x + y \equiv \left(2^i + 2^j\right) + (r + s) \equiv 0 + (r + s) \equiv r + s,$$

as claimed. Otherwise, they allow us to rewrite the sum as

$$x + y \equiv 2^i + \left[r + \left(2^j + s\right)\right] \equiv 2^i + (r + y)\,,$$

as claimed. □

Example 2·8. Consider the Nim games with values 3 and 5. We can apply the Nim Addition Theorem several times on these values:

$$\begin{aligned} 5 + 3 &= (4 + 1) + 3 \\ &\equiv 4 + (1 + 3) \\ &\equiv 4 + (3 + 1) \\ &\equiv 4 + [(2 + 1) + 1] \\ &\equiv 4 + [2 + (1 + 1)] \\ &= 4 + (2 + 0) \\ &\equiv 4 + 2 \\ &\equiv 6. \end{aligned}$$

The Nim Addition Theorem gives us an easy, recursive algorithm to add the values x and y of two games: cancel or "pop out" the largest power of two. Recursive algorithms are a little difficult, but this one is easy to "flatten," as follows:

- Write x and y in terms of powers of 2.
- Cancel out equal powers.
- Simplify the result.

Example 2·9. Earlier you showed that $1 + 3 \equiv 2$. The algorithm we just defined would have you do it this way:

$$1 + 3 = 1 + (1 + 2) = (1 + 1) + 2 \equiv 0 + 2 \equiv 2.$$

We can do this more generally. Look at the original Nim game with 3, 5, and 7 pebbles in each row. Its value is

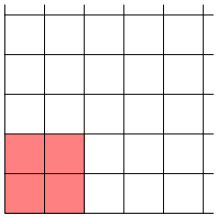
$$3 + 5 + 7 = (1 + 2) + (1 + 4) + (1 + 2 + 4) \equiv 1.$$

Question 2·10 .
Use the Nim Addition Theorem to show that if X and Y are games with values $x = 1$ and $y = n$, where n is even, then $x + y \equiv 1 + n$ and $1 + (1 + n) \equiv n$.

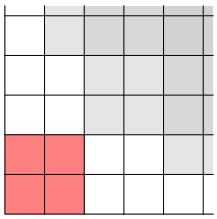
Question 2·11 .
Write a Nim Addition table for the game values from 0 to 10.

What about Ideal Nim?

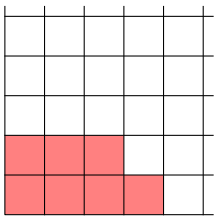
Ideal Nim exhibits a similar phenomenon. Eventually, the players divide the natural lattice into two parts. If a player divides the playing field into two, visually symmetric regions, she can force a win. For instance, suppose the game starts in this form, which you should recognize as the first game in Question 1.8:



If Emmy chooses the position $(2, 2)$, then she can reply to David's subsequent choices with a symmetric choice:



Again, not every game is *visually* symmetric, but sometimes a player can force a *chronological* symmetry. That is, you can turn the game into something that is *effectively* symmetric. For instance, suppose the game starts in this form, which you should recognize as the beginning of the game that leads to the second configuration in Question 1.10:



If Emmy chooses the position $(0, 2)$ — giving precisely the second configuration in Question 1.10 — then she can force a chronological symmetry in the following way:

- if David chooses $(a, 0)$, Emmy should choose $(a - 1, 1)$;
- if David chooses $(a, 1)$, Emmy should choose $(a + 1, 0)$.

Try this a few times to make sure you see how it works.

So the game can be won by *chronological* symmetry — in fact, the game is *always* won by chronological symmetry. Can we model this strategy arithmetically? What sort of properties should this arithmetic enjoy? The case of a *visually* symmetric game suggests the following.

Fact 2·12. *If x is a value of a configuration of the game, then x cancels itself.*

You can sort-of see this in the first game above: if we start with the choice of $(2, 2)$, then we can mirror any subsequent choice of (a, b) with (b, a), which is really the same move in an independent game.

This is not so easy to see in the second game, which is not visually symmetric. One way of seeing this property in the non-symmetric case is to define a new game, say "Dual Ideal Nim" (DIM), where players play two games of Ideal Nim, on two different lattices, *at the same time.* Were the *F*orbidden *F*rontier identical in each game, David would always win: whatever position Emmy selects in one game, David can select the exact same position in the second, showing that any move cancels itself. Again, it will help to draw a game of DIM that is based on the non-symmetric configuration to see what is going on.

As with Nim, the self-canceling symmetry of Ideal Nim implies a self-canceling arithmetic where $x + x \equiv 0$ for any value of x. It is possible, though not easy, to assign values to every game of Ideal Nim; while it bears similarities to the technique we outlined for Nim, the fact that we cancel in two directions makes it extraordinarily difficult to compute more than the simplest values. Because of this there is no feasible way to decide how to play most games of Ideal Nim.

Self-canceling arithmetic

This self-canceling property $x + x \equiv 0$ seems odd: how can you add x to itself and obtain zero, unless $x = 0$ already? Is there a more serious mathematical ground for this?

As a matter of fact, yes, and you use it every day! For instance, if you want to know the time 12 hours from now, and the time 12 hours after that, you could add 12 twice, working out the special aspects of a clock — or you could take advantage of the fact that adding $x = 12$ twice has the effect of canceling itself out! Indeed, if you want to know the time after y hours has passed, just add y and divide by 24; the remainder (!) tells you the current time.

Technically, time goes on and on without end, so we could list the hours from here to eternity as $\{0, 1, 2, \ldots\}$, which just so happens to be $\mathbb{N}$. But when you actually *compute* the hour of a day, you only work with the hours $\{1, 2, 3, \ldots, 11\}$ (if you use the conventional 12-hour clock) or $\{0, 1, 2, \ldots, 23\}$ (if you use a 24-hour clock). But aside from the fact that there are 24 hours in a day, from a mathematical point of view ***there's nothing really special about 12 or 24.*** In other situations and applications, it could be useful to play the same game with almost any other integer.

Example 2·13. The Roman general Julius Cæsar used a system called the **Cæsar cipher** to encrypt messages between separated army units. We can describe it mathematically in this fashion:

- replace the letters in the message with the numbers $A = 1, B = 2, C = 3, \ldots, Z = 26$;
- add 3 to each number in the message;
- if the value of a number is greater than 26, subtract 26 from that number;
- obtain the encrypted message by replacing the numbers with the letters $1 = A$, $2 = B$, $3 = C, \ldots, 26 = Z$.

Decryption consists of the very straightforward process of subtracting 3 from the letters' values, rather than adding. This is just like the clock, but using 26 instead of 24 (or 12).

Question 2·14 .

Can you decrypt the following message, written using the Cæsar cipher?

GDCCOHPHZLWKPDWK

Question 2·15 .

The Romans varied the Cæsar cipher by changing the second step. Rather than add 3 to each number in the message, they might add a different number instead, or even subtract. Knowing that the following message was generated using *a* Cæsar cipher, though you don't know what number was added or subtract, nor even whether it *was* added or subtracted, can you identify the precise technique and decrypt it?

THAOLTHAPJZPZAOLXBLLUVMAOLZJPLUJLZ

Hint: The most frequently used letters in English are e, t, and a. Look for a letter that appears frequently in the message, and see if assigning it to one of those three does the trick.

Clockwork arithmetic of integers

Through the rest of this chapter, d is a fixed, nonzero integer. We use $\mathbb{Z}$ for the integers, so we'll adopt $\mathbb{Z}_d$ for the set of all remainders from dividing by d. We'll use $d = 4$ for most examples, and an undetermined d for general reasoning. The Division Theorem tells us that remainders must be both nonnegative and smaller than d, so in the examples we look at $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, while in general we think about $\mathbb{Z}_d = \{0, 1, 2, \ldots, d-1\}$.

Let $a, b \in \mathbb{Z}$. Suppose the Division Theorem gives us quotients p, q and remainders r, s such that $a = pd + r$ and $b = qd + s$. What can we say about the remainder of $a + b$? On the one hand, substitution gives us

$$a + b = (p + q)\, d + (r + s)\,,$$

so we might be tempted to say that the remainder is $r + s$. Unfortunately, that's not always a remainder.

Example 2·16. With $d = 4$, $a = 7$, and $b = -22$, we have $r = 3$ and $s = 2$. The remainder of $a + b = -15$ is 1, but $r + s = 5$, which isn't even a remainder!

Let's not give up quite yet. You may have noticed a relationship between 1 (the actual remainder of $a + b$) and 5 (the sum of the remainders of a and b): the remainders are equal. If you try different values of a and b, you will observe a similar result: even if $r + s$ isn't equal to the remainder of $a + b$, the remainders of $r + s$ and $a + b$ are equal. If you try different values of d, you will observe the same phenomenon.

Theorem 2·17. *Let r and s be the remainders of dividing integers a and b by d. The remainder of $a + b$ is the same as the remainder of $r + s$ (both when divided by d).*

Proof. Let u be the remainder of $r + s$. Let q_a, q_b, q_{r+s} be the quotients of division of a, b, and $r + s$ by d. By definition,

$$r + s = q_{r+s}d + u.$$

By substitution,

$$\begin{aligned} a + b &= (q_a d + r) + (q_b d + s) \\ &= (q_a + q_b)\, d + (r + s) \\ &= (q_a + q_b)\, d + (q_{r+s}d + u) \\ a + b &= (q_a + q_b + q_{r+s})\, d + u. \end{aligned} \tag{2.1}$$

Closure of addition means $q_a + q_b + q_{r+s}$ is an integer, so line (2.1) satisfies criterion (D1) of the Division Theorem. But u is a remainder, so it also satisfies criterion (D2)! Division of integers gives us a *unique* remainder, so the remainder of $a + b$ is u, the remainder of $r + s$. □

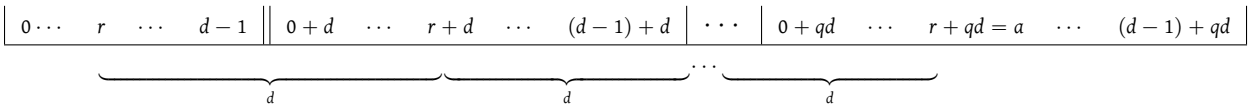
It's nice to know that the remainders of $a + b$ and $r + s$ are the same, but the theorem doesn't tell us any *relationship* between the $a + b$ and $r + s$, or at least not an obvious one. In fact, we can specify this relationship with precision.

Example 2·18. The remainders in the Example 2·16 were 1 and 5. Notice that $5 - 1 = 4$. In fact, for any number c and its remainder r, their difference $c - r$ will be a multiple of 4.

Again, this applies to any non-zero integer d. This is almost obvious, since we can rewrite criterion (D1) of the Division Theorem as

$$dq = c - r,$$

an explicit statement that d divides $c - r$. We can visualize this in the following way:



Division by d involves repeated subtraction of d. Each of the values r, $r + d$, …, $r + qd$ is a distance of d values from the next. So, *of course* d divides the difference of a number and its remainder from division by d.

This relationship further extends to any two numbers with the same remainder.

Theorem 2·19. *Two integers a and b have the same remainder after division by d if and only if d divides $a - b$.*

Proof. Assume that a and b have the same remainder r after division by d. The Division Theorem tells us that we can find integers p, q such that $a = pd + r$ and $b = qd + r$. By substitution,

$$a - b = (pd + r) - (qd + r) = (p - q)\, d.$$

By definition, d divides $a - b$.

Conversely, assume that d divides $a - b$. Let r and s be the remainders after dividing a and b by d, respectively. Find $p, q \in \mathbb{Z}$ such that $a = pd + r$ and $b = qd + s$, and choose $m \in \mathbb{Z}$ such that $dm = a - b$. By substitution and a little algebra,

$$\begin{aligned} dm &= (pd + r) - (qd + s) \\ d\,(m - p + q) &= r - s. \end{aligned}$$

The left hand side is a multiple of d. As the difference of two remainders, the right hand side is strictly between $-d$ and d; as it equals the left hand side, it must also be a multiple of d. This is possible only if $r - s = 0$, or $r = s$. So a and b have the same remainder r after division by d. □

This relationship is sufficiently important that we write $a \equiv_d b$ whenever a and b have the same remainder after division by d — or, equivalently, whenever d divides $a - b$. We call d the **modulus** of the expression $a \equiv_d b$. When the divisor is obvious, we simply write $a \equiv b$. This is sometimes pronounced, "b is **equivalent** to b (modulo d)."

This is similar to adding time. On a traditional clock, adding 8 hours to ten o'clock doesn't give you 18 o'clock; it gives you 6 o'clock: and the 6 comes from subtracting 12, the modulus. Put another way, $18 \equiv_{12} 6$.

How does this relate to Nim and Ideal Nim? Recall that we wanted an arithmetic where $x + x = 0$. Consider the set $\mathbb{Z}_2 = \{0, 1\}$; in this case, $0 + 0 \equiv 0$ and $1 + 1 \equiv 0$. This isn't large enough to model all the possible values of our games, but it does show that *at least one set* has an arithmetic where this makes sense. Eventually we will find more.

Question 2·20 .
Show that clockwork multiplication is consistent; that is, if r and s are the respective remainders of dividing integers a and b by d, then the remainder of ab is the same as the remainder of rs. In short, $ab \equiv_d rs$.

Question 2·21 .
On the other hand, show that clockwork division has the following *un*desirable behavior: *for at least one* $d \in \mathbb{N}^+$, you can find nonzero integers $a, b, c \in \mathbb{Z}_d$ such that $ab \equiv_d ac$ but $b \not\equiv_d c$. This shows that you cannot divide by a, even though it is non-zero. This will be a big deal later.

Question 2·22 .
Continuing from Question 2.21, can you find a particular $d \in \mathbb{N}^+$ where clockwork division *does* behave desirably? You're looking for a d where every nonzero $a, b, c \in \mathbb{Z}_d$ satisfying $ab \equiv ac$ also satisfy $b \equiv c$.

Hint: Neither of the previous two problems requires a large value of d.

2·2 Properties and structure

> *If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.*
> — John von Neumann

We just saw that addition of remainders is in some sense "sensible." Just how similar *are* addition of integers and addition of remainders? Both are examples of **operations**; but what are those? Let S and T be sets. A **binary operation from** S **to** T is any function $f : S \times S \to T$. If $S = T$, we say that f is a binary operation **on** S. We will call the combination of a set with one or more binary operations an **algebraic system**.

The most familiar algebraic system is the natural numbers under addition. You've met many other algebraic systems:

- polynomials under addition and multiplication;
- rational numbers under addition and multiplication;
- matrices under addition and multiplication; and just recently you met
- $\mathbb{Z}_d$ under addition.

Over the remainder of this course, you will meet and study a number of other algebraic systems.

Properties with one operation

The "fundamental" sets we've looked at so far are $\mathbb{N}$, $\mathbb{N}^+$, and $\mathbb{Z}$. Let's look at the naturals first; the operation we associate with them is addition. What do we know about that addition?

- The sum of two natural numbers is also natural. We call this **closure**, and say that $\mathbb{N}$ is **closed** under addition.
- For the sum of three natural numbers, it doesn't matter if we add the first two numbers first, or the last two numbers first; the answer is always the same. We call this the **associative property**, and say that $\mathbb{N}$ is **associative** under addition.
- The sum of 0 and a natural number n is always n. We call 0 the **identity** of $\mathbb{N}$, and say that $\mathbb{N}$ satisfies the **identity property** under addition..

Before looking at remainders, let's ask ourselves: do $\mathbb{N}^2$ and the monomials in x and y satisfy this property?

Let's look at monomials first. Right away, we see a problem: the sum of two monomials is not a monomial; if they are *un*like, we get a *binomial;* and if they are *a*like, we get a term with a coefficient. If you look back at our definition of monomials, you'll notice that we allow only the product of variables, and not a coefficient, as well. So monomial addition is *not* closed.

However, monomial exponents are natural numbers, and we *add* exponents when we *multiply* polynomials. Does *monomial multiplication* satisfy the above properties?

- The product of two monomials $t = x^a y^b$ and $u = x^c y^d$ is $v = x^{a+b} y^{c+d}$. The naturals are closed, so $a + b$ and $c + d$ are natural, so v is in fact a monomial. Since t and u were arbitrary, monomial multiplication is closed.
- The product of three monomials $t = x^a y^b$, $u = x^c y^d$, and $v = x^m y^n$ gives

  $$(tu)\, v = \left(x^{a+c} y^{b+d}\right) v = x^{(a+c)+m} y^{(b+d)+n}$$

  if we multiply the first two first, and

  $$t\,(uv) = t\left(x^{c+m} y^{d+n}\right) = x^{a+(c+m)} y^{b+(d+n)}$$

  if we multiply the second two first. These two products are equal if $(a + c) + m = a + (c + m)$ and $(b + d) + n = b + (d + n)$. These are natural numbers, which we know to be associative, so they are equal! Monomial multiplication is associative.
- What about an identity? It makes sense that the multiplicative identity should be 1, since $1 \times x^a y^b = x^a y^b$, but is 1 a monomial? Of course! $1 = x^0 y^0$, an *empty product*. So monomial multiplication has an identity.

Don't let yourself be tempted to think that the identity should be 0, as with natural number *addition*. What matters is not an element's appearance, but its behavior. "Judge not a book by its cover," they say; neither should you judge a number by its appearance. Not only is 0 not obviously a monomial, but it doesn't behave under multiplication the way an identity should behave: $0 \cdot t = 0$, but we need $0 \cdot t = t$. Fortunately, 1 fits the bill.

The correspondence between monomials and the lattice suggests that addition on the lattice also satisfies these properties:

- If we add two lattice points (a, b) and (c, d), the sum $(a + c, b + d)$ is also a lattice point. So the lattice is closed under addition.

- If we add three lattice points (a, b), (c, d), and (m, n), the sum from adding the first two first is

$$[(a, b) + (c, d)] + (m, n) = (a + c, b + d) + (m, n) = ((a + c) + m, (b + d) + n),$$

  while the sum from adding the second two first is

$$(a, b) + [(c, d) + (m, n)] = (a, b) + (c + m, d + n) = (a + (c + m), b + (d + n)).$$

  These two sums are equal on account of the associative property of natural number addition.

- What about an identity? The lattice point corresponding to the monomials' identity, $1 = x^0y^0$, is $(0, 0)$. In fact, $(0, 0) + (a, b) = (a, b) = (a, b) + (0, 0)$.

The operations on the three sets are *superficially* different: we add naturals and lattice points, but multiply monomials. Nevertheless, they share the same *substantive* structure.

You can probably remember other sets that share this structure, so it must be important. Let's give it a special name. We'll use the letter S to stand in for a generic set, and adopt the symbol $*$ to stand in for a generic operation, a symbol that combines both addition and multiplication. We say that S is a **monoid under** $*$ if together they satisfy the following operations.

closure if $s, t \in S$, then $s * t \in S$ also;

associative if $s, t, u \in S$, then $s * (t * u) = (s * t) * u$; and

identity we can find $я \in S$ such that if $s \in S$, then $я * s = s = s * я$.[1]

Take note of an important point: for closure, it's important that $s * t$ be not only *defined,* but *an element of S!* If it isn't an element of S, then S is not closed under the operation, and can't be a monoid. Also notice that we use $я$ to stand in for a generic identity element, rather than risking 1 or 0.

You may have noticed that monoids lacks some useful properties. To start with,

commutative if $s, t \in S$, then $s * t = t * s$.

While many monoids do enjoy that property, many don't. You'll meet some non-commutative monoids later on. When a monoid *is* commutative, we call it a **commutative monoid**.

What about this property?

inverse if $s \in S$, then we can find $t \in S$ such that $s * t = я = t * s$.

[1] Depending on the set and operation, the identity could be the number 0, the number 1, a matrix, a function, or something elese entirely. When we don't know (and we often don't) we will use $я$ to stand for a generic identity. This letter which looks like a backwards R is a Cyrillic letter "ya"; that already helps it stand out, but it has the added benefit that in some Slavic languages it means "I," which makes it apt for the "identity."

A monoid that enjoys the inverse property is a **group**. We usually write s^{-1} for the inverse of s, so we can rewrite the equation $s * t = я$ as $s * s^{-1} = я\ldots$ with one exception. If we know a group's operation is addition, we write its identity as 0, and the inverse of s as $-s$; in that case, we rewrite the equation $s * t = я$ as $s + (-s) = 0$.

Groups that enjoy the commutative property are usually called **abelian groups**, not commutative groups.

Question 2·23 .

Consider the set $B = \{F, T\}$ with the operation $\vee$ where

$$F \vee F = F$$
$$F \vee T = T$$
$$T \vee F = T$$
$$T \vee T = T.$$

This operation is called **Boolean or**.

Is $(B, \vee)$ a monoid? If so, is it a group? Explain how it satisfies each property.

Question 2·24 .

Consider the set $B = \{F, T\}$ with the operation $\wedge$ where

$$F \wedge F = F$$
$$F \wedge T = F$$
$$T \wedge F = F$$
$$T \wedge T = T.$$

This operation is called **Boolean and**.

Is $(B, \wedge)$ a monoid? If so, is it a group? Explain how it satisfies each property.

Question 2·25 .

Consider the set $B = \{F, T\}$ with the operation $\oplus$ where

$$F \oplus F = F$$
$$F \oplus T = T$$
$$T \oplus F = T$$
$$T \oplus T = F.$$

This operation is called **Boolean exclusive or**, or **xor** for short.

Is $(B, \oplus)$ a monoid? If so, is it a group? Explain how it satisfies each property.

Question 2·26 .

Which of the sets $\mathbb{N}^+$, $\mathbb{N}$, and $\mathbb{Q}$ are

(a) commutative monoids under addition?

(b) commutative monoids under multiplication?

(c) abelian groups under addition?

(d) abelian groups under multiplication?

Question 2·27 .

Recall that if S is a set, then $P(S)$ is the power set of S; that is, the set of all subsets of S.

(a) Suppose $S = \{a, b\}$. Compute $P(S)$, and show that it is a monoid under $\cup$ (union). Is it also a group?

(b) Let S be *any* set. Show that $P(S)$ is a monoid under $\cup$ (union). Is it also a group?

Question 2·28 .

(a) Suppose $S = \{a, b\}$. Compute $P(S)$, and show that it is a monoid under $\cap$ (intersection). Is it also a group?

(b) Let S be *any* set. Show that $P(S)$ is a monoid under $\cap$ (intersection). Is it also a group?

Definition 2·29. Let G be any group.

1. For all $x, y \in G$, define the **commutator of** x **and** y to be $x^{-1}y^{-1}xy$. We write $[x, y]$ for the commutator of x and y.
2. For all $z, g \in G$, define the **conjugation of** g **by** z to be zgz^{-1}. We write g^z for the conjugation of g by z.

Question 2·30 .

(a) Explain why $[x, y] = e$ iff x and y commute.

(b) Show that $[x, y]^{-1} = [y, x]$; that is, the inverse of $[x, y]$ is $[y, x]$.

(c) Show that $(g^z)^{-1} = (g^{-1})^z$; that is, the inverse of conjugation of g by z is the conjugation of the inverse of g by z.

(d) Fill in each blank of Figure 2.30 with the appropriate justification or statement.

Claim: $[x,y]^z = [x^z, y^z]$ for all $x, y, z \in G$.

Proof:

1. Let _____.
2. By _____, $[x^z, y^z] = [zxz^{-1}, zyz^{-1}]$.
3. By _____, $[zxz^{-1}, zyz^{-1}] = (zxz^{-1})^{-1}(zyz^{-1})^{-1}(zxz^{-1})(zyz^{-1})$.
4. By Question _____,
$$\begin{aligned}(zxz^{-1})^{-1}(zyz^{-1})^{-1}(zxz^{-1})(zyz^{-1}) = \\ = (zx^{-1}z^{-1})(zy^{-1}z^{-1})(zxz^{-1})(zyz^{-1}).\end{aligned}$$
5. By _____,
$$\begin{aligned}(zx^{-1}z^{-1})(zy^{-1}z^{-1})(zxz^{-1})(zyz^{-1}) = \\ (zx^{-1})(z^{-1}z)y^{-1}(z^{-1}z)x(z^{-1}z)(yz^{-1}).\end{aligned}$$
6. By _____,
$$\begin{aligned}(zx^{-1})(z^{-1}z)y^{-1}(z^{-1}z)x(z^{-1}z)(yz^{-1}) = \\ = (zx^{-1})ey^{-1}exe(yz^{-1}).\end{aligned}$$
7. By _____, $(zx^{-1})ey^{-1}exe(yz^{-1}) = (zx^{-1})y^{-1}x(yz^{-1})$.
8. By _____, $(zx^{-1})y^{-1}x(yz^{-1}) = z(x^{-1}y^{-1}xy)z^{-1}$.
9. By _____, $z(x^{-1}y^{-1}xy)z^{-1} = z[x,y]z^{-1}$.
10. By _____, $z[x,y]z^{-1} = [x,y]^z$.
11. By _____, $[x^z, y^z] = [x,y]^z$.

Figure 2·1: Material for Question 2.30(c)

So does addition of remainders form a monoid, or even a group?

To answer this question, we first have to make precise *what sort* of addition we mean. We have to fix a divisor, so let's go ahead and use d in general, and $d = 4$ for examples, just as before.

Remainders aren't closed under *ordinary* addition (Example 2·16), but *clockwork* addition is closed (Theorem 2·17), so let's try that. We'll use the symbol $\oplus_d$ to make it clear that we're thinking about the result of clockwork addition, or just plain $\oplus$ when no one is looking and it's clear which d we mean, which is pretty much all the time. That is, $r = a \oplus b$ means that r is the remainder from division of $a + b$ by d.

Is clockwork addition associative? Let $a, b, c \in \mathbb{Z}_d$. Suppose that $r = a \oplus b$ and $s = (a \oplus b) \oplus c$. By definition, $r \equiv a + b$, so by substitution, $s \equiv r + c$. Use Theorem 2·19 to choose $p, q \in \mathbb{Z}$ such that $dp = (a + b) - r$, and $dq = (r + c) - s$. We need to show that $s = a \oplus (b \oplus c)$, or equivalently, $r + c \equiv a + (b + c)$. The definition of congruence impels us to consider whether

$$d \mid [(a + (b + c)) - (r + c)].$$

This is true if and only if $d \mid (a + b - r)$. We have already stated that $dp = (a + b) - r$, which by definition means $d \mid (a + b - r)$.

Does clockwork addition have an identity element? It makes sense to guess that 0 is the identity of clockwork addition. Let $a \in \mathbb{Z}_d$; $a + 0 = a$, a remainder, so $a \oplus 0 = a$ and $0 \oplus a = a$, as well.

We have shown that $\mathbb{Z}_d$ is a monoid under addition! It is commutative since $a + b = b + a$, so $a \oplus b = b \oplus a$, as well. Let's see if it is also a group.

Does clockwork addition satisfy the inverse property? Let $a \in \mathbb{Z}_d$; we showed that 0 is the identity, so we need to find $b \in \mathbb{Z}_d$ such that $a + b = 0$ and $b + a = 0$. We claim that $d - a$ is the inverse. To see why, let $b = d - a$. Notice that $a + b = 0$ and $b + a = 0$, as desired. However, it's not enough for an inverse to exist *somewhere;* it must exist *in the same set!* We have to check that b is an actual element of $\mathbb{Z}_d$.

The elements of $\mathbb{Z}_d$ are $\{0, 1, 2, \ldots, d - 1\}$. If we can show that $d - a$ is one of those numbers, we're done. We know $b \in \mathbb{N}$ because $b = d - a$, and $a < d$, so that's fine. However, we do *not* have $b \in \mathbb{Z}_d$ when $a = 0$, because $d - a = d \notin \mathbb{Z}_d$! This is a mistake, but it's an important mistake to point out, because it can be easy to overlook. Fortunately, we can fix this.

Most values of a work fine with the formula $b = d - a$; the only one that fails is $a = 0$. It's easy to verify that 0 is its own inverse: $0 + 0 = 0$, done. So, one way to bridge the gap is to define $b = a$ for $a = 0$, and $b = d - a$ otherwise. A second, equivalent, way to bridge the gap: define b as the remainder of $d - a$ when you divide by d; we leave it to you to explain why this resolves the matter.

Question 2·31 .
Show that defining b as the remainder of $d - a$ when we divide by d always obtains the additive inverse of a in $\mathbb{Z}_d$.

We have now encountered finite groups, and we will encounter more. It's useful to think in terms of their size, for which we use a special term.

Definition 2·32. If a group has a finite number of elements, we say its **order** is that number of elements. If a group has an infinite number of elements, we say its order is infinite.

Question 2·33 .
The smallest group has order 1. What properties does that only element have?

Question 2·34 .
We did not indicate whether $\mathbb{Z}_d$ was a commutative monoid, and thus an abelian group. Is it?

What about structures with two operations?

So far, we've dealt only with structures that have one operation; we considered addition of numbers, clockwork addition, and monomial multiplication. You may be wondering how we classify two operations that interact. For example, how might addition and multiplication interact? You may recall the following property.

distributive if $s, t, u \in S$, then $s \times (t + u) = s \times t + s \times u$.

A **ring** is a set S where $\times$ satisfies the properties of a monoid, addition satisfies the properties of an abelian group, and the two interact via the distributive property.[2] If the multiplication is also commutative, we call S a **commutative ring**. We *always* write a generic ring's additive identity as 0, and a generic ring's multiplicative identity as 1.

What about division? A **unit** is an element of a ring with a multiplicative inverse. A **field** is a commutative ring where you can "divide" by non-zero elements, because they all have multiplicative inverses. The integers are not a field; after all, $2/3 \notin \mathbb{Z}$, in part because the multiplicative inverse of 3 is not in $\mathbb{Z}$. We fix that in the following way.

- The set of **rational numbers** is the set of all well-defined fractions of integers; in set-builder notation, we'd write,

$$\mathbb{Q} = \left\{ a/b : a \in \mathbb{Z} \text{ and } b \in \mathbb{N}^+ \right\}.$$

Just as the integers "enable" subtraction, the rationals "enable" division. That is, while you *can* subtract naturals, you aren't guaranteed a natural, but when you expand your horizon to include integers, you are always guaranteed an integer. Likewise, while you *can* divide integers, you aren't guaranteed an integer, but when you expand your horizon to include rationals, you are always guaranteed a rational. — With one exception: a number with 0 in the denominator has issues that only a nonstandard analyst can handle. This is why we qualify our fractions as "well-defined" for the same reason that the set-builder notation puts $b \in \mathbb{N}^+$.[3]

[2] Many texts do *not* assume a ring has a multiplicative identity, but others do. We side with the latter for the sake of simpler exposition and theorems.

[3] Why can't we divide by zero? Basically, it doesn't make sense. Suppose that we could find a number c such that $1 \div 0 = c$. The very idea of division means that if $1 \div 0 = c$, then $1 = 0 \cdot c$, but $0 \cdot c = 0$ for *any* integer c, so we can't have $1 = 0 \cdot c$. We could replace 1 by any nonzero integer a, and achieve the same result. Admittedly, this reasoning doesn't apply to $0 \div 0$, but even *that* offends our notion of an operation! If we were to assign some $c = 0 \div 0$, we would not be able to decide between $0 \div 0 = 1$ (since $0 = 0 \cdot 1$), $0 \div 0 = 2$ (since $0 = 0 \cdot 2$), $0 \div 0 = 3$ (since $0 = 0 \cdot 3$), and so forth. Then there is the matter of the grouping model of division; dividing $4 \div 0 = c$ implies that there are exactly c groups of 0 in 4, but no finite c satisfies this assertion.

Question 2·35 .
Which of the sets $\mathbb{N}^+$, $\mathbb{N}$, and $\mathbb{Q}$ are

(a) commutative rings under ordinary addition and multiplication?

(b) fields under ordinary addition and multiplication?

Question 2·36 .
Is $(B, \vee, \wedge)$ a ring? Is it a field? (Here we are saying that $\vee$ stands in for the addition, while $\wedge$ stands in for the multiplication.)

Question 2·37 .
Is $(B, \oplus, \wedge)$ a ring? Is it a field? (Here we are saying that $\oplus$ stands in for the addition, while $\wedge$ stands in for the multiplication.)

Question 2·38 .
Let's return to the discussion of cardinality in Question 1·18. We had concluded with the weird result that the cardinalities of $\mathbb{N}$ and $\mathbb{Z}$ are the same.

Speaking of weird results, show that $\mathbb{N}$ and $\mathbb{Q}$ have the same cardinality. This is a little harder, so we're going to cheat. First, explain why $\mathbb{Q}$ "obviously" has cardinality no smaller than $\mathbb{N}$'s, by showing that you can match every element of $\mathbb{N}$ to an element of $\mathbb{Q}$, and have infinitely many elements of $\mathbb{Q}$ left over. Then, show that $\mathbb{N}$ "obviously" has cardinality no smaller than $\mathbb{Q}$'s, because if we arrange the elements of $\mathbb{Q}$ according to the following table:

$$\begin{array}{ccccc} 0/1 & 0/2 & 0/3 & 0/4 & \cdots \\ 1/1 & 1/2 & 1/3 & 1/4 & \cdots \\ 2/1 & 2/2 & 2/3 & 2/4 & \cdots \\ 3/1 & 3/2 & 3/3 & 3/4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

then you can match every element of $\mathbb{Q}$ to an element of $\mathbb{N}$, and have infinitely many elements of $\mathbb{N}$ left over. (Think diagonally. — No, the *other* diagonally.) Since neither's cardinality is smaller than the other's, it seems reasonable to conclude they have equal cardinality.

Cayley tables

A useful tool for analyzing operations on small sets is an abstract multiplication table, sometimes called the **Cayley table**. For instance, the Cayley tables for addition and multiplication in $\mathbb{Z}_4$ look like this:

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

$\otimes$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

You may notice some interesting properties: every element of $\mathbb{Z}_4$ appears exactly once in each row or column of the first table, but not the second. Another strange phenomenon is that $2 \otimes 3 = 2 \otimes 1$ even though $3 \neq 1$.

Question 2·39 .
List the elements of $\mathbb{Z}_2$, and write its Cayley table. Notice how this starts to justify our notion of a self-canceling arithmetic.

Question 2·40 .
We observed that every element appears exactly once in each row or column of the Cayley tables above. We can write this mathematically as, if $a * c = d$ and $b * c = d$, then $a = b$.

(a) To see that the statement might *not* be true in a monoid, build the Cayley table of $\mathbb{Z}_6$ under multiplication. Show that it satisfies the properties of a monoid, but not of a group. Then identify elements $x, y, z \in \mathbb{Z}_6$ such that $x \otimes z = y \otimes z$, even though $x \neq y$. Isn't that weird?

(b) A phenomenon related to this one is that with natural numbers, if $ab = 0$, then $a = 0$ or $b = 0$. That's *not* true in an arbitrary monoid! Identify elements $a, b \in \mathbb{Z}_6$ such that $a \otimes b = 0$, but $a, b \neq 0$. How are these elements related to the result in (a)? *Hint:* You've already done this problem; it's just phrased differently. See Question 4.72.

(c) Prove that, in a group G, if $a, b, c, d \in G$, $a * c = d$ and $b * c = d$, then $a = b$. *Hint:* Since it's true in a group, but not in a monoid, you should use a property that is special to groups, but not monoids.

(d) Use (c) to explain why every element d of a group G appears exactly once in each row or column of a group's Cayley table. *Hint:* If it appears in two rows, what equation does that imply?

In a ring, multiplication by zero behaves exactly as you'd expect.

Fact 2·41. *If R is a ring and $a \in R$, then $a \times 0 = 0$ and $0 \times a = 0$.*

Why? By the identity and distributive properties, $a \times 0 = a \times (0 + 0) = a \times 0 + a \times 0$. Let $b = a \times 0$ and condense the chain to

$$b = b + b.$$

Add $-b$ to both sides, and apply some properties of rings, and we have

$$\begin{aligned} -b + b &= -b + (b + b) \\ 0 &= (-b + b) + b \\ 0 &= 0 + b \\ 0 &= b. \end{aligned}$$

By substitution, $a \times 0 = 0$. □

On the other hand, multiplication *to* zero is a bit funny — not so much "ha ha funny" so much as "strange funny."

Definition 2·42. Let R be a ring, and $a, b \in R$. If $ab = 0$ and neither $a = 0$ nor $b = 0$, then we call a and b **zero divisors**. A ring *without* zero divisors satisfies the **zero product property**; that is, if $ab = 0$, then $a = 0$ or $b = 0$. ("If the product is zero, a factor is zero.") A ring that satisfies the zero product rule is an **integral domain**.

Example 2·43. • The integers $\mathbb{Z}$ are an integral domain.

As you have just seen, $\mathbb{Z}_d$ is not always an integral domain, but sometimes it is. When?

Question 2·44 .
Carry out enough computations in $\mathbb{Z}_3$, $\mathbb{Z}_4$, $\mathbb{Z}_5$, and $\mathbb{Z}_6$ to answer the following: For which values of d will $\mathbb{Z}_d$ have zero divisors, and for which values of d is $\mathbb{Z}_d$ an integral domain? What property do the rings with zero divisors share, as opposed to the integral domains?

Question 2·45 .
Show that every field is an integral domain. Conversely, name an integral domain that is not a field.

2·3 Isomorphism

Plus ça change, plus c'est la même chose.
(The more things change, the more they stay the same.)
— French proverb

We've seen several important algebraic systems that share the same structure. For instance, $(\mathbb{N}, +)$, $(\mathbb{Z}_d, +)$, and $(\mathbb{M}, \times)$ are all monoids. When looking at two algebraic systems that share a basic structure, mathematicians sometimes ask themselves, *How similar are they?* Is the similarity more than superficial? Could it be that their Cayley tables are essentially identical, so that one of the systems are, from an algebraic view, exactly the same?

You might also look at it a different way. Two algebraic systems can have an initially different appearance, but while working with both you notice that certain behaviors are the same. It's easier to work with one system than the other; in particular, it's easier to show that a pleasant property holds for one system than for the other. If their Cayley tables are essentially identical, then you know the "difficult" system does in fact share that pleasant property, as well.

The technical word for this is *isomorphism,* and we can rephrase our question this way:

How can we decide whether two algebraic systems are isomorphic?

In general, we replace "algebraic system" with the particular structure that interests us:

How can we decide whether two monoids are isomorphic?

How can we decide whether two groups are isomorphic?

How can we decide whether two rings are isomorphic?

How can we decide whether two fields are isomorphic?

This section considers how to do this.

Question 2·46 .

Recall the structures Boolean or $(B, \vee)$, Boolean and $(B, \wedge)$, and Boolean xor $(B, \oplus)$ (Questions 2.23, 2.24, and 2.25).

All three are monoids, but inspection of the Cayley tables will show that two are more or less the same (hence, "isomorphic" in our intuitive notion of the term), but the third is different from the others. Which two are isomorphic? Why isn't the third isomorphic?

Be careful on this problem — superficially, *none* of their Cayley tables look the same. You have to look closely at the layout of the Cayley table before you notice the pattern.

The idea

Imagine two offices. How would you decide if the offices were equally suitable for a certain job? You first need to know what tasks have to be completed, and what materials you need. If the tasks require reference books, you would want a bookshelf in the office. If they require writing, you would want a desk, perhaps a computer. If they require communication, you might need a phone.

With such a list in hand, you can make an educated comparison between the offices. If both offer the needed equipment, you'd consider both suitable for the job at hand. The precise manner in which the offices satisfy these requirements doesn't matter; if one's desk is wood, and the other's is steel, that makes an aesthetic difference, but they're functionally the same. If one office lacked a desk, however, it wouldn't be up to the required job.

Deciding whether two algebraic systems are isomorphic is similar. First, you decide what structure you want to analyze. Next, you compare how the sets satisfy those structural properties. If you're looking at finite monoids, an exhaustive comparison of their Cayley tables might work, but the method is called "exhaustive" for a reason. Besides, we deal with infinite sets like $\mathbb{N}$ often enough that we need a non-exhaustive way to compare their structure. Functions turn out to be just the tool we need.

How so? Let S and T be any two sets. Recall that a **function** $f : S \to T$ is a relation that sends every input $s \in S$ to precisely one value in T, the output $f(s)$. You have probably heard the geometric interpretation of this: f passes the "vertical line test." You might suspect at this point that we are going to generalize the notion of function to something more general, just as we generalized from the lattice and monomials to monoids. To the contrary, we *specialize* the notion of a function in a way that tells us important information about a monoid.

Suppose M and N are monoids. If they are isomorphic, their monoid structure is identical, so we ought to be able to build a function that maps elements with a certain behavior in M to elements with the same behavior in N. (Table to table, phone to phone.) What does that mean? Let $a, b, c \in M$ and $x, y, z \in N$. If M and N have the same structure as monoids, with x filling in for a, y filling in for b, and z filling in for c, we would expect that

- if $ab = c$, then
- $xy = z$.

Question 2·47 .
Suppose you know only two facts about an algebraic system $(G, *)$: it forms a group, and G holds exactly two elements, $\mathfrak{s}$ (the identity) and g. You know neither the elements' internal structure, nor how the operation $*$ works. You know *only* that G is a group of two elements. Show that, regardless of this profound ignorance, the group properties force exactly one Cayley table on G. In other words, **all groups of order 2 are isomorphic!**

Hint: Try to build the Cayley table of G. You will encounter no ambiguity in the process, forcing the conclusion that only one possible table exists.

Question 2·48 .
Suppose you know only two facts about an algebraic system $(G, *)$: it forms a group, and G holds exactly three elements, $\mathfrak{s}$ (the identity), g, and h. As before, you know neither the elements' internal structure, nor how the operation $*$ works. Show that, regardless of this profound ignorance, the group properties force exactly one Cayley table on G. In other words, **all groups of order 3 are isomorphic!**

Question 2·49 .
Suppose you know only two facts about an algebraic system $(G, *)$: it forms a group, and G holds exactly four elements, $\mathfrak{s}$ (the identity), g, h, and gh. As before, you know neither the elements' internal structure, nor how the operation $*$ works. Show that, regardless of this profound ignorance, the group properties force exactly… *two* Cayley tables on G. (More than one!) In other words, (a) **not** all groups of order 4 are isomorphic, and (b) there are exactly *two* groups of order 4, **"up to isomorphism!"**

Definition 2·50. In Question 2.49, you should have encountered exactly one ambiguity while completing the Cayley table: what value can we assign $a * a$? The case where $a^2 = \mathfrak{s}$ is called the **Klein 4-group**. The case where $a^2 \neq \mathfrak{s}$ should look like another system you've played with.

The definition

Recall that our idea of isomorphism in monoids works as follows. For every $a, b, c \in M$ and every $x, y, z \in N$,

- if a corresponds to x, b corresponds to y, and c corresponds to z, *and*
- if $ab = c$, then $xy = z$.

In mathematics, we can say that "a corresponds to x" using function notation, $f(a) = x$. That principle allows us to rewrite the equation $xy = z$ as

$$f(a) f(b) = f(c).$$

But remember, $ab = c$, so substitution tells us the operation corresponds if

$$f(a)f(b) = f(ab). \tag{2.2}$$

The identity of M should also correspond to the identity of N, so we need to add the condition

$$f(\text{я}_M) = \text{я}_N. \tag{2.3}$$

When dealing with a group, the inverse of an element should correspond to the inverse its corresponding element, which gives us a third condition, $f(x^{-1}) = a^{-1}$, which we rewrite as

$$f\left(x^{-1}\right) = f(x)^{-1}. \tag{2.4}$$

If we can pull off both (2.2) and (2.3) (as well as (2.4) in a group), we say that f is a **homomorphism**, from the Greek words "homo" and "morphos", meaning "same shape". The existence of a homomorphism tells us that the Cayley table of M has the same shape as a subset of the Cayley table of N.

That's not enough to answer the question. We don't want to know merely whether something akin to M appears *in* N; we want M and N to be *essentially identical.* Just as we only need one table in any office, we want the correspondence between the elements of the monoids to be unique: in other words,

f should be one-to-one.

Finally, everything in N should correspond to something in M; if the offices are identical, we shouldn't find something useful in the second that doesn't appear in the first. In terms of f, that means

f should be onto.

We summarize our discussion up to this point with the following definition:

Definition 2·51. Let $(S, *)$ and $(T, \star)$ be monoids. If there exists a function $f : S \to T$ such that

- $f(\text{я}_S) = \text{я}_T$ *(f preserves the identity)*

and

- $f(a * b) = f(a) \star f(b)$ for all $a, b \in S$, *(f preserves the operation)*

then we call f a **monoid homomorphism**.

Now suppose $(S, *)$ and $(T, \star)$ are groups. If there exists a function $f : S \to T$ such that

- $f(\text{я}_S) = \text{я}_T$, *(f preserves the identity)*
- $f(a * b) = f(a) \star f(b)$ for all $a, b \in S$, *(f preserves the operation)*

and

- $f\left(a^{-1}\right) = f(a)^{-1}$ for all $a \in S$, *(f preserves inverses)*

then we call f a **group homomorphism**.

Finally, suppose $(R, \times, +)$ and $(S, \times, +)$ are rings. If there exists a function $f : R \to S$ such that

- f is a group homomorphism with respect to addition, and
- f is a monoid homomorphism with respect to multiplication,

then we call f a **ring homomorphism**.

If f is also a bijection, then we say M is **isomorphic** to N, write $M \cong N$, and call f an **isomorphism**. (A **bijection** is a function that is both one-to-one and onto.)

We used $(S, *)$ and $(T, \star)$ in the definition to emphasize that they could stand for *any* two algebraic systems, regardless of the operations involved.

An immediate goal, of course, is to show that the natural numbers under addition are isomorphic (as monoids) to the monomials under multiplication. We'll write $\mathbb{X}$ for the set of all natural powers of x; that is, $\mathbb{X} = \{1, x, x^2, \dots\}$. We have noticed already that monoid multiplication works like the addition of natural numbers.

Example 2·52. We claim that $(\mathbb{X}, \times)$ is isomorphic to $(\mathbb{N}, +)$. To see why, map $f : \mathbb{X} \to \mathbb{N}$ via $f\left(x^a\right) = a$. First we show that f is a bijection.

To see that it is one-to-one, let $t, u \in \mathbb{X}$, and assume that $f(t) = f(u)$. By definition of $\mathbb{X}$, we can find $a, b \in \mathbb{N}$ such that $t = x^a$ and $u = x^b$. Substituting this into $f(t) = f(u)$, we find that $f\left(x^a\right) = f\left(x^b\right)$. The definition of f allows us to rewrite this as $a = b$. However, if $a = b$, then $x^a = x^b$, and $t = u$. We assumed that $f(t) = f(u)$ for arbitrary $t, u \in \mathbb{X}$, and showed that $t = u$; that proves f is one-to-one.

To see that f is onto, let $a \in \mathbb{N}$. We need to find $t \in \mathbb{X}$ such that $f(t) = a$. Which t should we choose? We want $f\left(x^{\text{something}}\right) = a$. We know that $f\left(x^{\text{something}}\right) = \text{something}$. We are looking for a t that makes $f(t) = a$, so the "natural" choice seems to be something $= a$, or $t = x^a$. That would certainly guarantee $f(t) = a$, but can we actually find such an object t in $\mathbb{X}$? Since $x^a \in \mathbb{X}$, we can in fact make this choice! We took an arbitrary element $a \in \mathbb{N}$, and showed that f maps some element of $\mathbb{X}$ to a; that proves f is onto.

So f is a bijection. Is it also an isomorphism? First we check that f preserves the operation. Let[4] $t, u \in \mathbb{X}$. By definition of $\mathbb{X}$, $t = x^a$ and $u = x^b$ for $a, b \in \mathbb{N}$. We now manipulate $f(tu)$ using definitions and substitutions to show that the operation is preserved:

$$\begin{aligned} f(tu) &= f\left(x^a x^b\right) = f\left(x^{a+b}\right) \\ &= a + b \\ &= f\left(x^a\right) + f\left(x^b\right) = f(t) + f(u). \end{aligned}$$

[4]The definition uses the variables x and y, but those are just letters that stand for arbitrary elements of M. Here $M = \mathbb{X}$ and we can likewise choose any two letters we want to stand in place of x and y. It would be a very bad idea to use x when talking about an arbitrary element of $\mathbb{X}$, because there *is* an element of $\mathbb{X}$ called x. So we choose t and u instead.

The operation in $\mathbb{X}$ is multiplication; the operation in $\mathbb{N}$ is addition, so we should expect $f(t)+f(u)$ at the end; the operations is indeed preserved.

Does f also preserve the identity? We usually write the identity of $M = \mathbb{X}$ as 1, but this just stands in for x^0. On the other hand, the identity (under addition) of $N = \mathbb{N}$ is the number 0. We use this fact to verify that f preserves the identity:

$$f(\text{ят}_M) = f(1) = f\left(x^0\right) = 0 = \text{ят}_N.$$

(We won't usually write ят_M and ят_N, but I'm doing it here to show explicitly how this relates to the definition.)

We have shown that there exists a bijection $f : \mathbb{X} \longrightarrow \mathbb{N}$ that preserves the operation and the identity. We conclude that $\mathbb{X} \cong \mathbb{N}$.

Question 2·53 .

Earlier, you inspected the Cayley tables of $(B, \wedge)$, $(B, \vee)$, and $(B, \oplus)$, and found that two were isomorphic. Define an isomorphism f from one monoid to its isomorphic counterpart.

On the other hand, is $(\mathbb{N}, +) \cong (\mathbb{N}, \times)$? You might think this easy to verify, since the sets are the same. Let's see what happens.

Example 2·54. Suppose there *does* exist an isomorphism $f : (\mathbb{N}, +) \to (\mathbb{N}, \times)$. What would have to be true about f? Let $a \in \mathbb{N}$ such that $f(1) = a$; after all, f has to map 1 to *something!* An isomorphism must preserve the operation, so

$$\begin{aligned} f(2) &= f(1+1) = f(1) \times f(1) = a^2 \text{ and} \\ f(3) &= f(1+(1+1)) = f(1) \times f(1+1) = a^3, \text{ so that} \\ f(n) &= \cdots = a^n \text{ for } \textit{any } n \in \mathbb{N}. \end{aligned}$$

So f sends *every* integer in $(\mathbb{N}, +)$ to a power of a.

Think about what this implies. For f to be a bijection, it would have to be onto, so *every* element of $(\mathbb{N}, \times)$ would *have* to be an integer power of a. ***This is false!*** After all, 2 is not an integer power of 3, and 3 is not an integer power of 2. We have found that $(\mathbb{N}, +) \ncong (\mathbb{N}, \times)$.

Question 2·55 .

Both $\mathbb{Z}$ and $2\mathbb{Z}$ are groups under addition.

(a) Show that $f : \mathbb{Z} \to 2\mathbb{Z}$ by $f(z) = 2z$ is a group isomorphism. Hence $\mathbb{Z} \cong 2\mathbb{Z}$.

(b) Show that $\mathbb{Z} \cong n\mathbb{Z}$, as groups, for every nonzero integer n.

Question 2·56 .

Let $d \geqslant 1$. Both $\mathbb{Z}$ and $\mathbb{Z}_d$ are rings, though $\mathbb{Z}$ is a ring under ordinary addition and multiplication, while $\mathbb{Z}_d$ is a ring under modular addition and multiplication. Let $f : \mathbb{Z} \to \mathbb{Z}_d$ by $f(a) = [a]_d$, where $[a]_d$ means "the remainder of a after division by d."

(a) Show that f is a ring homomorphism.

(b) Explain why f cannot possibly be a ring isomorphism. You don't need any symbols here; the best explanation uses only a few words.

Question 2·57 .
Let $M = \{\{\}, \{a\}\}$.

(a) Show that M is a monoid under the operation $\cup$ (set union).

(b) Show that $(M, \cup)$ is isomorphic to the monoid "Boolean or".

(c) Can M be isomorphic to the monoid "Boolean xor"?

Question 2·58 .
Let

$$M = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

(a) Show that M is a monoid under matrix multiplication.

(b) Show that M is isomorphic to the monoid "Boolean xor".

(c) Can M be isomorphic to the monoid "Boolean or"?

Sometimes, less is more

As defined, a group homomorphism is a function that preserves

- the operation ($f(xy) = f(x)f(y)$),
- the identity ($f(\text{я}) = \text{я}$), and
- inverses ($f(x^{-1}) = f(x)^{-1}$).

Amazingly, we can define a group homomorphism using *only one of these three!*

Theorem 2·59. *Let G and H be groups, and suppose $f : G \to H$ is a function that preserves the operation; that is, $f(xy) = f(x)f(y)$ for all $x, y \in G$. In this case, f* automatically *preserves the identity and all inverses.*

The upshot is that to show a function is a group homomorphism, you need not check all three properties! You need check only that the operation is preserved.

Proof. We need to show that f preserves the identity and all inverses.

For the identity, let $h \in H$. Let $g \in G$, and $h = f(g)$. By hypothesis, f preserves the operation, so $f(g\text{я}_G) = f(g)f(\text{я}_G)$. By definition of an identity, $g\text{я}_G = g$, so we can rewrite the previous equation as $f(g) = f(g)f(\text{я}_G)$. By substitution, $h = h \cdot f(\text{я}_G)$. Since H is a group, h has an inverse in H, so we can multiply both sides by the inverse of h, obtaining $\text{я}_H = f(\text{я}_G)$. In other words, f preserves the identity.

For inverses, let $g \in G$, and let $h = f(g)$. Since G is a group, g has an inverse in G. By hypothesis, f preserves the operation, so $f(g \cdot g^{-1}) = f(g) \cdot f(g^{-1})$. By substitution, $f(я_G) = hf(g^{-1})$. We just showed that f preserves the identity, so we can rewrite the equation as $я_H = hf(g^{-1})$. Since H is a group, h has an inverse in H, so we can multiply both sides by the inverse of h, obtaining $h^{-1} = f(g^{-1})$. By substitution, $f(g)^{-1} = f(g^{-1})$. In other words, f preserves the inverse of g. Since g was an arbitrary element of G, f must preserve *all* inverses. □

This shortcut does not work for monoid homomorphisms!

Question 2·60 .
What aspect of the proof suggests that this shortcut does not work for monoid homomorphisms?

Question 2·61 .
Consider the monoids $M = (\mathbb{N}, \times)$ and $N = (\mathbb{N}, +)$. Let $f : M \to N$ by $f(x) = 0$. Explain why:

(a) f preserves the operation, but

(b) f does not preserve the identity.

Question 2·62 .
Let $M = \{я, a\}$, and consider the operation where $я$ is the identity and $a^2 = я$. Let $N = \{1, b\}$ and consider the operation where 1 is the identity and $b^2 = b$.

(a) Show that M and N are both monoids under this operation. Which one is not a group?

(b) Show that the map $f : M \to N$ defined by $f(я) = b$ and $f(a) = b$ preserves the operation, despite not preserving the identity.

(c) How does this show that there is no parallel to Theorem 2·59 for monoids?

Direct Products

It is easy to build new algebraic systems using a Cartesian product of algebraic systems. Let S_1, S_2, ... be a sequence of groups, a sequence of monoids, or a sequence of rings. Let $T = S_1 \times S_2 \times \cdots$. (We proceed as if we have infinitely many S, but it works just as well if there are finitely many, and the example below will have finitely many.) Define an operation $*$ on T as follows:

- for any $t, u \in T$,
- we can write $t = (t_1, t_2, \dots)$, $(u_1, u_2, \dots)$ where
  - $t_1, u_1 \in S_1, t_2, u_2 \in S_2, \dots$

so define

$$t * u = (s_1 t_1, s_2 t_2, \dots).$$

We say that the operation in T is **componentwise**: we apply the operation of S_1 to elements in the first component; the operation of S_2 to elements in the second component; and so forth.

Example 2·63. Consider $\mathbb{Z}_2$ and $\mathbb{Z}_3$ as rings under addition and multiplication, modulo 2 or 3 as appropriate. Then

$$\mathbb{Z}_2 \times \mathbb{Z}_3 = \{(0_2, 0_3), (0_2, 1_3), (0_2, 2_3), (1_2, 0_3), (1_2, 1_3), (1_2, 2_3)\}.$$

(Henceforth we leave off the 2's and 3's, since the first component is only ever in $\mathbb{Z}_2$ and the second only ever in $\mathbb{Z}_3$.) Given the operation defined above, sums of elements in $\mathbb{Z}_2 \times \mathbb{Z}_3$ are

$$(0, 2) + (1, 1) = (1, 0)$$
$$(1, 1) + (1, 1) = (0, 2)$$

while products of elements in $\mathbb{Z}_2 \times \mathbb{Z}_3$ are

$$(0, 2) \times (1, 1) = (0, 2)$$
$$(1, 1) \times (1, 1) = (1, 1).$$

Fact 2·64. *Let S_1, S_2, … be a sequence (possibly finite) of algebraic systems, and T their cartesian product, with componentwise operation(s) defined as above.*

(A) T is a monoid under the componentwise operation if all the S_i are monoids.

(B) T is a group under the componentwise operation if all the S_i are groups.

(C) T is a ring under componentwise addition and multiplication if all the S_i are rings under their respective addition and multiplication.

However, T is never an integral domain, even if all the S_i are integral domains, unless every $S_i = \{0\}$.

Why? We show (A) and (B), since that also covers (C). We leave the question of why T is not an integral domain to the reader. To see why, let $t, u \in T$.

(A) Suppose each S_i is a group. By definition, $t * u = (t_1 u_1, t_2 u_2, \dots)$. By hypothesis, each S_i is a monoid, hence closed, so each $t_i u_i \in S_i$, so $t * u \in T$. That shows closure. For associativity, let $v \in T$; again, each S_i is associative, so

$$\begin{aligned}
t * (u * v) &= t * (u_1 v_1, u_2 v_2, \dots) && (\text{def of } *)\\
&= (t_1 (u_1 v_1), t_2 (u_2 v_2), \dots) && (\text{def of } *)\\
&= ((t_1 u_1) v_1, (t_2 u_2) v_2, \dots) && (\text{each } S_i \text{ assoc})\\
&= (t_1 u_1, t_2 u_2, \dots) * v && (\text{def of } *)\\
&= (t * u) * v. && (\text{def of } *)
\end{aligned}$$

Finally, write $я_i$ for the identity of S_i, and observe that $(я_1, я_2, \dots) \in T$. We claim that this is the identity; indeed,

$$t * (я_1, я_2, \dots) = (t_1 я_1, t_2 я_2, \dots) = (t_1, t_2, \dots) = t$$

and likewise if we multiply by t on the right. So $(я_1, я_2, \dots)$ really does act as the identity for T, and we abbreviate it as $я_T$.

We have show that T is closed and associative under the componentwise operation, and that it has an identity; hence, T is a monoid.

(B) For the group property, we need merely show that every element of T has an inverse. Each component t_i of t is an element of S_i, which by hypothesis is a group, so $t_i^{-1} \in S_i$. By definition of T, $\left(t_1^{-1}, t_2^{-1}, \dots\right) \in T$. Consider its product with t:

$$t * \left(t_1^{-1}, t_2^{-1}, \dots\right) = \left(t_1 t_1^{-1}, t_2 t_2^{-1}, \dots\right) = (я_1, я_2, \dots) = я_T.$$

Hence $\left(t_1^{-1}, t_2^{-1}, \dots\right)$ is an inverse of t in T. □

Question 2·65 .

Construct Cayley tables for addition and multiplication in $\mathbb{Z}_2 \times \mathbb{Z}_3$. Indicate the zero divisors.

Question 2·66 .

The group $\mathbb{Z}_2 \times \mathbb{Z}_2$ has four elements. We already know that, up to isomorphism, there are only two groups: $\mathbb{Z}_4$ and the Klein 4-group. To which of these is $\mathbb{Z}_2 \times \mathbb{Z}_2$ isomorphic?

Question 2·67 .

Let $f : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$ by the rule $f(a) = ([a]_2, [a]_3)$. For instance, $f(4) = ([4]_2, [4]_3) = (0, 1)$.

(a) Compute all the images of f.

(b) How do you know f is one-to-one and onto?

(c) Show that f is a homomorphism.
Hint: You could show this exhaustively (only 36 pairs!) but need not do so. Instead, use a previous result on products of $\mathbb{Z}_n$.

(d) Why is $\mathbb{Z}_6 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$?

Question 2·68 .

Show that even if S_1, S_2, … are all integral domains, $T = S_1 \times S_2 \times \cdots$ is not an integral domain, unless every $S_i = \{0\}$.

Chapter 3

Common and important algebraic systems

The previous chapter introduced you to monoids, groups, rings, and fields, emphasizing primarily remainders. This chapter aims to show that these structures' elegant properties apply to other mathematical objects. These objects are of fundamental important in advanced algebra, so it seems appropriate to introduce them here.

3·1 Polynomials, real and complex numbers

God created the integers. All else is the work of man.
— Leopold Kronecker

Let R be any commutative ring. We say that x is **indeterminate over** R if x has no specific value, but we can substitute any value of R for x. Naturally, $ax = xa$. A **polynomial in** x **over** R is any finite sum of the form

$$f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where each $a_i \in R$ and $a_n \neq 0$. We call each a_i the **coefficient** of the corresponding x^i, and call a_n the **leading coefficient**.

If we're feeling lazy, which we often are, we just say f is polynomial over R, since the indeterminate is obvious. If we're feeling *especially* lazy, which we sometimes are, we just say f is polynomial, since the ring is clear from context.

We need not restrain ourselves to x; any symbol will do, as long as the meaning is clear. For instance, if t is indeterminate over $\mathbb{Z}_4$, then $2t + 3$ is a polynomial in t over $\mathbb{Z}_4$. If y is indeterminate over $\mathbb{Q}$, then $\frac{2}{3}x^2 - \frac{1}{5}x$ is a polynomial in y over $\mathbb{Q}$.

Let f be a polynomial in x whose coefficients are elements of R. We say that f is a **polynomial over** R, and we write $R[x]$ for the set of all polynomials over R. We call R the **ground ring** of $R[x]$. Addition and multiplication of polynomials over R behaves the same as addition and multiplication of polynomials over $\mathbb{Z}$; the only difference is the ground ring.

Example 3·1. Polynomials with integer coefficients are elements of $\mathbb{Z}[x]$. Polynomials with rational coefficients are elements of $\mathbb{Q}[x]$. Polynomials with coefficients modulo $d > 0$ are elements of $\mathbb{Z}_d[x]$.

Question 3·2 .
Suppose R is a commutative ring, with additive identity 0 and multiplicative identity 1. Show that $R[x]$ is also a commutative ring, with the same identities as R.

Fact 3·3. *It is also the case that if R is an integral domain, then so is $R[x]$.*

Why? If $f, g \in R[x]$ are nonzero but $fg = 0$, then the leading term of fg is zero; this leading term is the product of the leading terms of f and g. If we write at for the leading term of f and bu for the leading term of g (where $c, d \in R$ and $t, u \in \mathbb{X}$) then, by definition, $(ct)(du) = 0$. This is possible only if $cd = 0$. As they come from the leading terms of f and g, the leading coefficients must be nonzero; that is, $c, d \neq 0$. But $c, d \neq 0$ and $cd = 0$ means c and d are zero divisors, so R cannot be an integral domain. We have shown the contrapositive of the claim, and the contrapositive is equivalent to the claim itself. □

The Division Theorem for Polynomials (p. 34) tells us that we can use monic divisors to compute quotients and remainders in $\mathbb{Z}[x]$. We can actually do this with polynomials over *any* commutative ring!

On the one hand, it makes sense that a similar argument should apply for polynomials with rational, or even real coefficients, but it might not be so clear for stranger rings which you have yet to meet. Stranger yet, *we decline to write a proof* generalizing the Division Theorem for Polynomials to these other rings. Why? Sometimes, generalizing a result like this is quite hard, but in this case it does not require much convincing; go back and examine the proof. Does anything in the argument depend on the coefficients' being integers? Nothing does; the argument would have worked for any ring R. We *do* need a monic divisor, and we *do* need a ring of coefficients, since the proof required both subtraction and multiplication of coefficients. This hints that there is a larger, more interesting structure we have not named yet, but we pass over that for the time being.

Question 3·4 .
Rewrite the proof of the Division Theorem for Polynomials, replacing any instance of $\mathbb{Z}$ or "integer" with R or "ring element". Convince yourself that, yes, this is a wonderfully general result.

You will recall that we developed a class of rings, called $\mathbb{Z}_d$, by building an algebraic system on remainders of integer division. A natural question to ask is,

Can we build a consistent algebraic system on remainders of polynomial *division?*

Indeed, we can! We will also find that this gives us a concrete way of building an "imaginary" algebraic system.

Polynomial remainders

Let's look at how remainder arithmetic modulo a polynomial might work.

Example 3·5. Let $g = x^2 - 1$. Any remainder r after division by g has degree smaller than 2 (after all, $\deg g = 2$), so we can write

$$r = ax + b,$$

where a and b are integers. That's it! There are no other restrictions on r, and none on a and b, aside from their being integers.

We have already encountered one difference with integer remainders: there can be infinitely many polynomial remainders! (After all, you can choose a and b arbitrarily from the ground ring.) At least the degree of the divisor constrains them.

Will the arithmetic of polynomial remainders exhibit a "clockwork" behavior, as with integer remainders? Not with addition, since

$$(ax + b) + (cx + d) = (a + c)\,x + (b + d)\,,$$

and no matter what the values of a, b, c, and d, that sum has degree 1. With multiplication, however,

$$(ax + b)\,(cx + d) = acx^2 + (ad + bc)\,x + bd$$

ventures into forbidden territory, with degree 2. We have to reduce this polynomial modulo $x^2 - 1$.

Example 3·6. Consider the remainders $2x + 3$ and $-5x + 12$, modulo $x^2 - 1$. Their sum is

$$(2x + 3) + (-5x + 12) = -3x + 15,$$

another remainder. Their product is

$$(2x + 3)\,(-5x + 12) = -10x^2 + 9x + 36,$$

which is not a remainder, but we can reduce it modulo $x^2 - 1$ to $9x + 26$. In other words,

$$(2x + 3)\,(-5x + 12) \equiv 9x + 26.$$

Theorem 3·7. *Let R be a commutative ring, and $R[x]$ a polynomial ring. Let g be a monic polynomial of $R[x]$. The set of remainders modulo g also forms a ring under addition and multiplication, modulo g.*

Proof. Question 3.2 tells us that $R[x]$ is a commutative ring, and hence an abelian group under addition. Addition of polynomials does not change the degree, so as we saw above, the properties of $R[x]$ are preserved in the set of remainders; the sums are, in fact, identical, so the identity of addition of remainders remains the zero polynomial, which is itself a remainder, and the additive inverse of a remainder is also present. So the set of remainders preserves the abelian group property of $R[x]$.

On the other hand, multiplication of remainders risks raising the degree, so the product of two remainders might not itself be a remainder, as we saw above. However, our multiplication is modulo g, and when we divide the product by g, we obtain a remainder. This guarantees closure. The multiplicative identity of polynomial multiplication is the constant polynomial 1, which is itself a remainder. The commutative property is likewise preserved, so if the set of remainders *is* a ring, it is a commutative ring. There remain two properties to check.

What of the associative property of multiplication? Let r, s, and t be remainders. We know that $(rs)\,t = r\,(st)$ as polynomials; since the remainder of division is unique, we must also have $(rs)\,t \equiv r\,(st)$. Distribution follows similarly. □

So far, we have observed nothing strange with these remainders, but the next example *does* exhibit a very unusual behavior.

Example 3·8. Consider the remainders $x + 1$ and $x - 1$ modulo $x^2 - 1$. Their sum is

$$(x + 1) + (x - 1) = 2x,$$

another remainder. No surprise. Their product is

$$(x + 1)(x - 1) = x^2 - 1,$$

which is not a remainder, but we can reduce it modulo $x^2 - 1$ to… 0?!?

Zero divisors have returned!

Ordinary multiplication of two nonzero polynomials over an integral domain gives you a nonzero polynomial (Fact 3·3). After all, multiplication increases the degree, so you can't get 0 as a product of nonzero polynomials.

With the modular product of remainders, those guarantees vanish! Upon reflection, this makes sense, because $x^2 - 1$ factors into $x + 1$ and $x - 1$ precisely — just as $6 = 2 \times 3$. Just as with $\mathbb{Z}_d$, this has consequences for solving equations; until now, you usually solved equations under the assumption that a product of zero has a factor of zero.

Example 3·9. When we try to find integer solutions of equations such as $x^3 - 1 = 0$, we typically factor first, obtaining

$$(x - 1)\left(x^2 + x + 1\right) = 0.$$

As integer polynomials, we know that if the product is zero, a factor must be zero, helping us to find the solution $x = 1$. We enjoy no such guarantee from *remainder* arithmetic.

The introduction of zero divisors doesn't happen modulo every polynomial. With some, we get a different phenomenon.

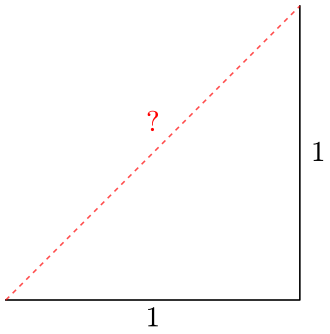
Real numbers

The set of **real numbers** is the set of all possible distances one can move along a line, with "positive length" indicating we moved in one direction, and "negative length" indicating we moved in the opposite direction. Its shorthand is $\mathbb{R}$. There are ways to write this in set-builder notation, but I'll pass over that for now.

You may wonder if $\mathbb{R} = \mathbb{Q}$. If you don't wonder it, that's okay; someone else has already wondered it, and we know the answer: *no*.

Fact 3·10. *$\sqrt{2}$ is real, but not rational.*

Why? We know that $\sqrt{2}$ is real because the Pythagorean Theorem tells us that it is the length of the hypotenuse of an isosceles right triangle whose legs have length 1.



The length of the red line is $\sqrt{2}$, so $\sqrt{2}$ is real.

However, $\sqrt{2}$ is not rational. To see why, let $a, b \in \mathbb{N}$, with $b \neq 0$, and suppose $\sqrt{2} = a/b$. (We can assume a is natural because $\sqrt{2}$ is positive.) Suppose further that a and b have no common divisors; after all, if they do, we can simplify the fraction. (The well-ordering principle means simplification can't continue indefinitely.) Rewrite $\sqrt{2} = a/b$ as $b\sqrt{2} = a$; square both sides to obtain $2b^2 = a^2$. Notice that a^2 is an even number; this is possible only if a is even, so $a = 2c$ for some integer c. Rewrite as $2b^2 = (2c)^2$, so $2b^2 = 4c^2$, so $b^2 = 2c^2$. The argument above implies that b is even. So a and b are both even, giving them a common divisor. But this contradicts the reasonable assumption above that they have *no* common divisors! Our assumption that we could write $\sqrt{2} = a/b$, where a and b are natural, is false: $\sqrt{2}$ is real, but not rational. □

We call lengths like $\sqrt{2}$ **irrational numbers**. You'll meet some of these in the exercises. Despite the unfortunate name, they are not unreasonable, and have some very important uses. Thus, we not only have

$$\mathbb{N}^+ \subseteq \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R},$$

we also have

$$\mathbb{N}^+ \subsetneq \mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

We can describe three-dimensional real space as

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(a, b, c) : a, b, c \in \mathbb{R}\};$$

people use this notation a lot in multivariate calculus.

As with the rationals, we can divide real numbers, and end up with a real number. Also with the rational, we can't divide by zero.

Question 3·11 .

We return to the question of cardinality again. We had shown that $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ have the same cardinality. They do *not* have the same cardinality as $\mathbb{R}$. To see why, suppose the contrary, that we have a matching of distinct real numbers to natural numbers, so that we can list all the real numbers in a row, $a_1, a_2, \ldots$.

Consider a real number b built by taking as its first digit after the decimal point a digit that is not the first digit after the decimal point of a_1, as its second digit after the decimal point a digit that is not the second digit after the decimal point of a_2, as its third digit after the decimal point a digit that is not the third digit after the decimal point of a_3, and so forth.

(a) How do we know that b does not appear in the list $a_1, a_2, \ldots$?

(b) You need not show that b is a real number, but it is. How does this show that $\mathbb{N}$ and $\mathbb{R}$ must have different cardinality?

(c) Why does that mean that $\mathbb{Z}$ and $\mathbb{Q}$ likewise have different cardinality from $\mathbb{R}$?

Complex numbers

The real numbers make for a lovely field, but they retain an important defect.

Fact 3·12. *There is no real solution to $x^2 + 1 = 0$; that is, $\sqrt{-1} \notin \mathbb{R}$.*

Proof. Let $a \in \mathbb{R}$. By the definitions of real arithmetic, a^2 is positive. That means $a^2 + 1$ is also positive, so $a^2 + 1 > 0$ for any real number a. Thus, no real number a can serve as a solution to $x^2 + 1 = 0$. □

Historically, we introduce a new symbol, i, to stand in for the solution to $x^2 + 1 = 0$, and say that i possesses the property that $i^2 = -1$. This is not especially appealing; small wonder mathematicians refer to it as "the imaginary number". Aside from our desire to introduce a solution to this polynomial, can we identify a concrete representation of such a number? Yes!

Let $\mathbb{C}$ be the set of all remainders when you divide a polynomial in $\mathbb{R}[x]$ by $x^2 + 1$. In other words,

$$\mathbb{C} = \{ax + b : a, b \in \mathbb{R}\}.$$

We can show without much effort that $\mathbb{C}$ is a field, where the arithmetic is addition and multiplication modulo $x^2 + 1$.

Fact 3·13. *$\mathbb{C}$ is a field.*

Proof. Theorem 3·7 tells us that $\mathbb{C}$ is a commutative ring, so we need merely show that every nonzero element of $\mathbb{C}$ has an inverse. To see this, let $z \in \mathbb{C}$ be nonzero. By definition, we can find real numbers a and b such that $z = ax + b$, and at least one of a and b is nonzero. That means $a^2 + b^2 \neq 0$. Let

$$w = -\frac{a}{a^2 + b^2} \cdot x + \frac{b}{a^2 + b^2}.$$

Notice that w has the proper form to be an element of $\mathbb{C}$. In addition,

$$zw = \left(-\frac{a^2}{a^2 + b^2}\right) \cdot x^2 + \frac{b^2}{a^2 + b^2}.$$

Reducing this modulo x^2+1, we have

$$zw \equiv \left[\left(-\frac{a^2}{a^2+b^2}\right)\cdot x^2 + \frac{b^2}{a^2+b^2}\right] - \left[\left(-\frac{a^2}{a^2+b^2}\right)\cdot x^2 - \frac{a^2}{a^2+b^2}\right] = \frac{b^2+a^2}{a^2+b^2} = 1,$$

so w is the multiplicative inverse of z, and $\mathbb{C}$ is a field. □

In the example of the previous section, we encountered zero divisors via $(x+1)(x-1) \equiv 0$. Can this happen in $\mathbb{C}$? In fact, it cannot, *precisely because* $\mathbb{C}$ *is a field.*

Question 3·14 .
Suppose that f and g are nonzero polynomials over a *field*. Why must $fg \neq 0$? *Hint:* Question 2.45 would be helpful.

You should notice that $\mathbb{R} \subseteq \mathbb{C}$, since constants are a special kind of polynomial. One element of $\mathbb{C}$ has a very special property.

Fact 3·15. $\mathbb{C}$ *contains exactly two elements that satisfy* $x^2+1 \equiv 0$.

Why? Let $i = 1x+0$. We claim that i satisfies the equation. Notice that i is, in fact, an element of $\mathbb{C}$, since it has the proper form. Substituting $x = 1x+0$ into x^2+1 shows that

$$x^2+1 = (1x+0)(1x+0)+1 \equiv 0$$

The other root is $-i = -1x+0$. We leave it to the reader to see that no other element of $\mathbb{C}$ satisfies the equation. □

Question 3·16 .
Why can no other element of $\mathbb{C}$ satisfy the equation $x^2+1=0$?

Let's summarize our accomplishment. We created a *new field* $\mathbb{C}$, which contains the real numbers as a subfield, and possesses a well-defined arithmetic that is consistent with the arithmetic of the real numbers: after all, multiplication of real numbers does not increase the degree, let alone invoke modular reduction. This new field also contains two elements that satisfy the equation given. We have constructed a number that has the properties of the imaginary number, but by its construction is clearly concrete!

Question 3·17 .
A real number a has a polynomial representation in $\mathbb{C}$ as $0x+a$. Use this to explain why "multiplication of real numbers does not increase the degree, let alone invoke modular reduction."

Although we have introduced the complex numbers using polynomial notation and congruence of remainders, we can write them in the more natural form, $a+bi$ where $a, b \in \mathbb{R}$.

Question 3·18 .
Show that there is a ring isomorphism between $\mathbb{C}$ as we have defined them, and $\mathbb{C}$ as traditionally defined. That is, show that

$$\left\{ax+b:\ a,b\in\mathbb{R},\ x^2+1\equiv 0\right\} \cong \left\{a+bi:\ a,b\in\mathbb{R},\ i^2=-1\right\}.$$

We rely on the traditional representation for future sections.

Question 3·19 .
We don't have to build $\mathbb{C}$ to obtain a ring containing the roots of $x^2 + 1$. Show that we *can* build such a ring using remainders of $\mathbb{Z}[x]$, modulo $x^2 + 1$.

We were able to construct a field containing the roots of $x^2 + 1$ using $x^2 + 1$ itself, but we cannot do this with $x^2 - 1$, because $x^2 - 1 = (x - 1)(x + 1)$, creating zero divisors. So $x^2 + 1$ is special, in that we can't rewrite it as the product of two smaller polynomials over $\mathbb{Z}$, or even over $\mathbb{R}$. That's an important property; let's give them a name. Recall that a *unit* is any element of a ring with a multiplicative inverse.

Definition 3·20. Suppose $r \in R$ is an element of a commutative ring, and r is not a unit. We say that r **factors over** R if we can find $s, t \in R$ such that $r = st$ and neither s nor t has a multiplicative inverse. Otherwise, r is **irreducible**.

Remark. If you are familiar with the notion of a "prime number", then you are likely wondering why we call r "irreducible" rather than "prime". The reason is that the algebraic meaning of "prime" is different. The two notions are compatible in the integers, but not in some other rings that you have studied, and will study later.

The definition assumes only that R is a commutative ring. That includes polynomial rings, so we've taken care of $x^2 + 1 \in \mathbb{Z}[x]$, and in fact of all irreducible polynomials over $\mathbb{Z}$.

The requirement that neither s nor t have a multiplicative inverse is important; otherwise, some smart aleck will point out that, in the integers, $2 = (-1) \times (-2)$ is a factorization of 2. Don't write off the smart aleck too quickly, though; we will see in Chapter 6 that this has important implications for factorization.

Question 3·21 .
Suppose that f is a polynomial with integer coefficients that factors into two polynomials of smaller degree, g and h, so that $f = gh$. Explain why we cannot use f to construct a *field* containing its own roots.

However, we can still build the roots of non-irreducible polynomials; it just takes a few steps.

Question 3·22 .
Suppose $f \in \mathbb{Z}[x]$ is not irreducible. How could you construct a field that contains *at least one* root of f, if not all of them? *Hint:* If f factors, the factors have lower degree. If they factor…

Question 3·23 .
Earlier, we constructed $\sqrt{2}$ as the length of the hypotenuse of a right triangle with legs of length 1. We can also construct it in the same way that we constructed the imaginary number i, using an irreducible polynomial with integer coefficients. Find such an irreducible polynomial, and show which remainder behaves the same as $\sqrt{2}$.

Irreducible polynomials play a major role in Chapter 6 on Factorization.

3·2 The roots of unity

> *The imaginary number is a fine and wonderful recourse of the divine spirit, almost an amphibian between being and not being.*
> — Gottfried Wilhelm Leibniz

Recall from Question 1.69 that a **root** of a polynomial $f(x)$ is any element a of the domain which, when substituted into f, gives us zero; that is, $f(a) = 0$. The example that motivated us to define the complex numbers was the polynomial $f(x) = x^2 + 1$, which has two roots, $\pm i$, where $i^2 = -1$.

Any root of the polynomial $f(x) = x^n - 1$ is called a **root of unity**. These are very important in the study of polynomial roots, in part because of their elegant form.

Example 3·24. The roots of $x^2 - 1$ are called the **square roots of unity**; they are $x = \pm 1$.

The roots of $x^3 - 1$ are called the **cube roots of unity**. It is clear that $x = 1$ is one such root, and the polynomial factors as

$$(x-1)\left(x^2 + x + 1\right).$$

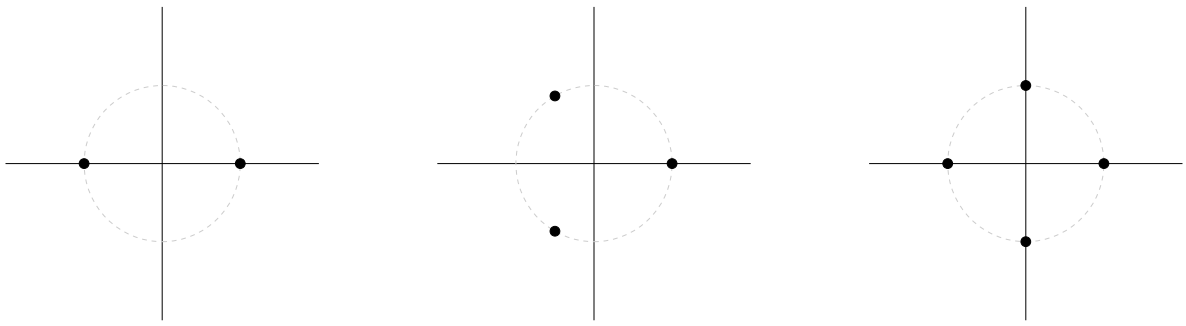
The quadratic factor contains the other cube roots of unity; by the quadratic formula, they are

$$x = \frac{-1 \pm \sqrt{1-4}}{2} = -\frac{1}{2} \pm i \cdot \frac{\sqrt{3}}{2}.$$

The roots of $x^4 - 1$ are called **the fourth roots of unity**. Since $x^4 - 1$ factors as $(x^2 - 1)(x^2 + 1)$, we already know these roots; they are $x = \pm 1, \pm i$.

A geometric pattern

It's often instructive to study the geometric behavior of a phenomenon, and this is no exception, but how shall we visualize complex numbers? Write $z = a + bi \in \mathbb{C}$, and refer to a as the **real part** of z, and b as the **imaginary part**. We'll abbreviate this in the future as $\text{real}(z) = a$ and $\text{imag}(z) = b$. Let's agree to plot z on the x-y plane using $\text{real}(z)$ for the x-coordinate, and $\text{imag}(z)$ for the y-coordinate. The graphs of the square, cube, and fourth roots of unity are as follows:



We've added the outline of a circle of radius 1 at the origin to illustrate a few interesting patterns:

- $x = 1$ is always a root.
- All the roots lie on the circle.
- The roots are, in fact, equidistant around the circle: they split the circumference of 2π into equal-sized arcs.

Question 3·25 .
Use the pattern above to sketch where the sixth roots of unity should like on the complex plane. Use that graph and some basic trigonometry to find their actual values as complex numbers. Verify that the values are correct by substituting them into the polynomial $x^6 - 1$.

If you recall your trigonometry, especially the parametric representation of the unit circle as $\cos^2 t + \sin^2 t = 1$, the observations above suggest the following.

Theorem 3·26. *Let $n \in \mathbb{N}^+$. The complex number*

$$\omega = \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)$$

is a root of $f(x) = x^n - 1$.

To prove Theorem 3·26, we need a different property of ω. We could insert it into the proof of Theorem 3·26, but it's useful enough on its own that we separate it as:

Lemma 3·27 (Powers of ω)**.** *If ω is defined as in Theorem 3·26, then*

$$\omega^m = \cos\left(\frac{2\pi m}{n}\right) + i\sin\left(\frac{2\pi m}{n}\right)$$

for every $m \in \mathbb{N}^+$.

Proof. We proceed by induction on m. For the *inductive base,* the definition of ω shows that ω^1 has the desired form. For the *inductive hypothesis,* assume that ω^m has the desired form. In the *inductive step,* we need to show that

$$\omega^{m+1} = \cos\left(\frac{2\pi(m+1)}{n}\right) + i\sin\left(\frac{2\pi(m+1)}{n}\right).$$

To see why this is true, use the inductive hypothesis to rewrite ω^{m+1} as,

$$\omega^{m+1} = \omega^m \cdot \omega \underset{\text{hyp.}}{\overset{\text{ind.}}{=}} \left[\cos\left(\frac{2\pi m}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)\right] \cdot \left[\cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)\right].$$

Distribution gives us

$$\begin{aligned}\omega^{m+1} = {} & \cos\left(\frac{2\pi m}{n}\right)\cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right)\cos\left(\frac{2\pi m}{n}\right) \\ & + i\sin\left(\frac{2\pi m}{n}\right)\cos\left(\frac{2\pi}{n}\right) - \sin\left(\frac{2\pi m}{n}\right)\sin\left(\frac{2\pi}{n}\right).\end{aligned}$$

Regroup the terms as

$$\omega^{m+1} = \left[\cos\left(\frac{2\pi m}{n}\right)\cos\left(\frac{2\pi}{n}\right) - \sin\left(\frac{2\pi m}{n}\right)\sin\left(\frac{2\pi}{n}\right)\right] + i\left[\sin\left(\frac{2\pi}{n}\right)\cos\left(\frac{2\pi m}{n}\right) + \sin\left(\frac{2\pi m}{n}\right)\cos\left(\frac{2\pi}{n}\right)\right].$$

The trigonometric sum identities $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$ and $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$, *used "in reverse"*, show that

$$\omega^{m+1} = \cos\left(\frac{2\pi(m+1)}{n}\right) + i\sin\left(\frac{2\pi(m+1)}{n}\right).$$

□

Once we have Lemma 3·27, proving Theorem 3·26 is spectacularly easy.

Proof of Theorem 3·26. Substitution and the lemma give us

$$\begin{aligned}\omega^n - 1 &= \left[\cos\left(\frac{2\pi n}{n}\right) + i\sin\left(\frac{2\pi n}{n}\right)\right] - 1\\ &= \cos 2\pi + i\sin 2\pi - 1\\ &= (1 + i\cdot 0) - 1 = 0,\end{aligned}$$

so ω is indeed a root of $x^n - 1$. □

A group!

Once we fix n, the nth roots of unity give us a nice group.

Theorem 3·28. *The nth roots of unity are $\Omega_n = \{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$, where ω is defined as in Theorem 3·26. They form a group of order n under multiplication.*

The theorem does not claim merely that Ω_n is a list of *some* nth roots of unity; it claims that Ω_n is a list of *all* nth roots of unity. Our proof is going to cheat a little bit, because we don't quite have the machinery to prove that Ω_n is an exhaustive list of the roots of unity. We will eventually, however, and you should be able to follow the general idea now.

Basically, let f be a polynomial of degree n. Suppose we know that f has n roots, named $\alpha_1, \alpha_2, \ldots, \alpha_n$. The parts you have to take on faith (for now) are twofold.

- First, there is only one way to factor f into linear polynomials. This is not obvious, and in fact it's not always true — but it is in this case, honest! The idea is called *unique factorization.*
- Second, if α_i is a root of f, then $x - \alpha_i$ is a factor of f for each α_i, so

  $$f(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)\cdot g(x),$$

  where g is yet to be determined. Each linear factor adds one to the degree of a polynomial, and f has degree n, so the product of the factors of f cannot have degree higher than n. However, we already have degree n on the right hand side of the equation, which means g can only be a constant, and the *only* roots of f are $\alpha_1, \ldots, \alpha_n$.

(You can see this in the example above with x^4-1, but the Factor Theorem will have the details (Question 1.69). You should have encountered that theorem in your precalculus studies, and since it doesn't depend on anything in this section, the reasoning is not circular.)

If you're okay with that, then you're okay with everything else.

Proof. For $m \in \mathbb{N}^+$, we use the associative property of multiplication in $\mathbb{C}$ and the commutative property of multiplication in $\mathbb{N}^+$:

$$(\omega^m)^n - 1 = \omega^{mn} - 1 = \omega^{nm} - 1 = (\omega^n)^m - 1 = 1^m - 1 = 0.$$

This shows that every positive power of ω is a root of unity. Most of these overlap, just as $(-1)^2 = (-1)^4 = (-1)^6 = \cdots$. If $\omega^m = \omega^\ell$, then

$$\cos\left(\frac{2\pi m}{n}\right) = \cos\left(\frac{2\pi \ell}{n}\right) \quad \text{and} \quad \sin\left(\frac{2\pi m}{m}\right) = \sin\left(\frac{2\pi \ell}{n}\right),$$

and we know from trigonometry that this is possible only if

$$\begin{aligned} \frac{2\pi m}{n} &= \frac{2\pi \ell}{n} + 2\pi k \\ \frac{2\pi}{n}(m-\ell) &= 2\pi k \\ m - \ell &= kn. \end{aligned}$$

That is, $m-\ell$ is a multiple of n. Since Ω_n lists only those powers from 0 to $n-1$, the powers must be distinct, so Ω_n contains n distinct roots of unity. (See also Question 3·26.) As there can be at most n distinct roots, Ω_n is a complete list of nth roots of unity.

Now we show that Ω_n is a cyclic group.

(closure) Let $x, y \in \Omega_n$; you will show in Question 3.29 that $xy \in \Omega_n$.

□

Question 3·29 .

Let $n \in \mathbb{N}^+$, and suppose that a and b are both positive powers of ω. Show that $ab \in \Omega_n$.

Proof of Theorem 3·28, continued. (associativity) The complex numbers are associative under multiplication; since $\Omega_n \subseteq \mathbb{C}$, the elements of Ω_n are also associative under multiplication.

(identity) The multiplicative identity in $\mathbb{C}$ is 1. This is certainly an element of Ω_n, since $1^n = 1$ for any $n \in \mathbb{N}^+$.

(inverses) Let $x \in \Omega_n$; you will show in Question 3.33 that $x^{-1} \in \Omega_n$.

(cyclic) Theorem 3·26 tells us that $\omega \in \Omega_n$; the remaining elements are powers of ω. Hence $\Omega_n = \langle \omega \rangle$.

□

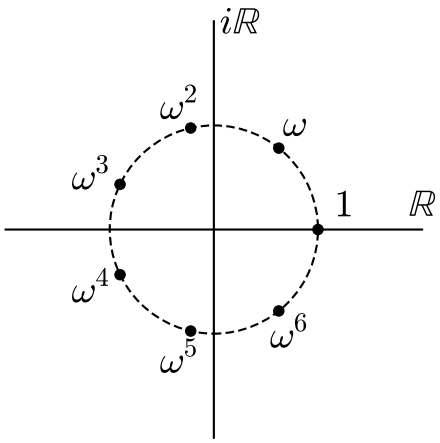


Figure 3·1: The seventh roots of unity, on the complex plane

Combined with the explanation we gave earlier of the complex plane, Theorem 3·28 gives us a wonderful symmetry for the roots of unity.

Example 3·30. Consider the case where $n = 7$. According to the theorem, the 7th roots of unity are $\Omega_7 = \{1, \omega, \omega^2, \ldots, \omega^6\}$ where

$$\omega = \cos\left(\frac{2\pi}{7}\right) + i\sin\left(\frac{2\pi}{7}\right).$$

According to Lemma 3·27,

$$\omega^m = \cos\left(\frac{2\pi m}{7}\right) + i\sin\left(\frac{2\pi m}{7}\right),$$

where $m = 0, 1, \ldots, 6$. By substitution, the angles we are looking at are

$$0, \frac{2\pi}{7}, \frac{4\pi}{7}, \frac{6\pi}{7}, \frac{8\pi}{7}, \frac{10\pi}{7}, \frac{12\pi}{7}.$$

See Figure 3·30.

Although we used $n = 7$ in this example, we used no special properties of that number in the argument. That tells us that this property is true for any n: the nth roots of unity divide the unit circle of the complex plane into n equal arcs!

Here's an interesting question: is ω is the only element of Ω_n whose powers "generate" the other elements of the group? In fact, no. A natural follow-up: are *all* the elements of Ω_n generators of the group? Likewise, no. Well, which ones are? We are not yet ready to give a precise criterion that signals which elements generate Ω_n, but they do have a special name.

Definition 3·31. We call any element of Ω_n whose powers gives us all other elements of Ω_n a **primitive nth root of unity**.

Question 3·32 .
Show that Ω_n is isomorphic to $\mathbb{Z}_n$.

Question 3·33 .

(a) Let ω be a 14th root of unity; let $\alpha = \omega^5$, and $\beta = \omega^{14-5} = \omega^9$. Show that $\alpha\beta = 1$.

(b) More generally, let ω be a primitive nth root of unity, Let $\alpha = \omega^a$, where $a \in \mathbb{N}$ and $a < n$. Show that $\beta = \omega^{n-a}$ satisfies $\alpha\beta = 1$.

(c) Explain why this shows that every element of Ω_n has an inverse.

Question 3·34 .
Suppose β is a root of $x^n - b$.

(a) Show that $\omega\beta$ is also a root of $x^n - b$, where ω is any nth root of unity.

(b) Use (a) and the idea of unique factorization that we described right before the proof of Theorem 3·28 to explain how we can use β and Ω_n to list all n roots of $x^n - b$.

Definition 3·35. Given a field $\mathbb{F}$, a **vector space** over $\mathbb{F}$ is an abelian group $(V, +)$ with an additional property called **scalar multiplication** that satisfies the following *additional* properties:

- Scalar multiplication maps $\mathbb{F} \times V$ to V, with $(a, u) \mapsto v$ abbreviated as $au = v$.
- **Closure:** for all $a \in \mathbb{F}$ and all $v \in V$, $av \in V$.
- **Compatibility:** for all $a, b \in \mathbb{F}$ and all $v \in V$, $(ab)\, v = a\,(bv)$.
- **Scalar identity:** for all $v \in V$, $1_{\mathbb{F}} v = v$.
- **Scalar distribution:** for all $a \in \mathbb{F}$ and all $u, v \in V$, $a\,(u + v) = au + av$.
- **Vectors distribution:** for all $a, b \in \mathbb{F}$ and all $v \in V$, $(a + b)\, v = av + bv$.

Question 3·36 .
Show that this section's construction of $\mathbb{C}$ satisfies the requirements of a vector space over $\mathbb{R}$.

3·3 Cyclic groups; the order of an element

> *"Well, in our country," said Alice, still panting a little, "you'd generally get to somewhere else—if you run very fast for a long time, as we've been doing."*
> *"A slow sort of country!" said the Queen. "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"*
> — Lewis Carroll

This section builds on a phenomenon we observed in a group of roots of unity to describe an important class of groups. Recall that the nth roots of unity can all be written as powers of

$$\omega = \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right);$$

that is,

$$\Omega_n = \left\{\omega, \omega^2, \ldots, \omega^n = 1\right\}.$$

Because of this, we spoke of ω as "generating" Ω_n. As you will see, we can write many other groups in this form. In addition, it will be of interest to look at groups generated by an element. Since we're dealing with repeating the operation of a group on one element, we'd best shore up some properties of exponents first.

Exponents

In essence, we claim that the usual arithmetic holds for exponents and multiples, regardless of the underlying group or ring; that is:

- for any integers $a, b \in \mathbb{Z}$, we define $g^a g^b = g^{a+b}$;
- if the set has an identity, then we define $g^0 = \text{я}$;
- if the set has multiplicative inverses, then we define $g^{-a} = (g^{-1})^a = (g^a)^{-1}$.

We have to make sure these definitions are reasonably well defined in any group or ring.

Question 3·37 .
We're going to start off deciding that g^0 is just shorthand for the group identity, я. If the operation of the group is addition, we'll usually write $0 \times g = 0$. Why do these notations make sense? *Hint:* $\text{я} = gg^{-1}$.

Question 3·38 .
Suppose $a \in \mathbb{N}^+$. Why can we say $g^{-a} = (g^{-1})^a = (g^a)^{-1}$? Are we sure that $(g^{-1})^a$ and $(g^a)^{-1}$ are always the same? *Hint:* Think about the definitions. The meaning of $(g^a)^{-1}$ is, "the inverse of g^a." What, then, has to be true for us to be able to say that $(g^{-1})^a = (g^a)^{-1}$? Show that *that* is true.

Lemma 3·39. *Let G be a group, $g \in G$, and $m, n \in \mathbb{Z}$. Each of the following holds:*

(A) $g^m g^{-m} = \text{ϧ}$; *that is,* $g^{-m} = (g^m)^{-1}$.

(B) $(g^m)^n = g^{mn}$.

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

The proof of Lemma 3·39 is not especially hard, but it does involve tedious notation. Originally, I included it here, but decided to remove it, on the grounds that (a) it distracts from the point of this section, which is to introduce you to cyclic groups, and (b) you really ought to be able to show it on your own (especially if your plan is to teach one day). So:

Question 3·40 .
Suppose $m \in \mathbb{Z}$ (not just $a \in \mathbb{N}^+$ as before). Why can we say $g^{-m} = (g^m)^{-1}$? *Hint:* What makes this different from before is that we're now dealing with *negative* exponents. Try considering different cases when $m \in \mathbb{N}^+$ (which we've already discussed, actually) and $n < 0$.

Question 3·41 .
Building on the previous question: let $n \in \mathbb{Z}$. Why can we say $(g^m)^n = g^{mn}$? *Hint:* As before, you need to consider separate cases for m or n negative.

Cyclic groups and generators

Some groups enjoy the special property that *every* element is a power of one, special element.

Definition 3·42. Let G be a group. If there exists $g \in G$ such that every element $x \in G$ has the form $x = g^n$ for some $n \in \mathbb{Z}$, then G is a **cyclic group** and we write $G = \langle g \rangle$. We call g a **generator** of G.

The idea of a cyclic group is that it has the form

$$\{\ldots, g^{-2}, g^{-1}, \text{ϧ}, g^1, g^2, \ldots\}.$$

If the group's operation is addition, we would of course write

$$\{\ldots, -2g, -g, 0, g, 2g, \ldots\}.$$

Example 3·43. Let's look at $\mathbb{Z}$ first. Any $n \in \mathbb{Z}$ has the form $n \cdot 1$, such as $2 = 2 \cdot 1$, $-5 = (-5) \cdot 1$, and so forth. We see that $\mathbb{Z}$ is cyclic, and write $\mathbb{Z} = \langle 1 \rangle$.

In addition, n has the form $(-n) \cdot (-1)$, so $\mathbb{Z} = \langle -1 \rangle$ as well. Both 1 and -1 are generators of $\mathbb{Z}$.

Question 3·44 .
Show that any group of 3 elements is cyclic.

Question 3·45 .
Is the Klein 4-group (Question 2.49 on page 61) cyclic? What about the cyclic group of order 4?

Question 3·46 .

Show that $\mathbb{Q}$ is not cyclic as an additive group. *Hint:* Suppose it were; then you could find a rational number q such that $\mathbb{Q} = \{\ldots, -2q, -q, 0, q, 2q, \ldots\}$. Surely you can find some $r \in \mathbb{Q}$ that isn't listed.

Question 3·47 .

Let $n \in \mathbb{Z}$, and consider the ring $\mathbb{Z}_n$.

(a) Show that its additive group is cyclic.

(b) Show that if $n = 7$, the subset $\{1, 2, \ldots, 6\}$ is a cyclic group under *multiplication*.

Hint: It's not enough to show that all the elements are generated by one element, though you do have to start there. You also have to check the properties of a group, *especially* that every element has an inverse.

(c) Show that if $n = 6$, the subset $\{1, 2, \ldots, 5\}$ is *not* a cyclic group under multiplication.

(d) Look at the subsets $\{1, 2, \ldots, n-1\}$ in some other finite rings $\mathbb{Z}_n$, where $n \geqslant 5$. Try at least two more and determine whether they are cyclic groups under multiplication.

(e) Do you notice a pattern to which values of n work and which don't?

Question 3·48 .

Suppose that G and H are groups, and $G \cong H$. Show that if G is cyclic, then so is H, because the generator of G is a generator of H.

In Definition 3·42 we referred to g as *a* generator of G, not as *the* generator. There could in fact be more than one generator; we see this in Example 3·43 from the fact that $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$. Another example is Ω_3, where ω and ω^2 both generate the group.

An important question arises here. Given a group G and an element $g \in G$, define $\langle g \rangle$ as the set of all integer powers of g. That is,

$$\langle g \rangle = \left\{\ldots, g^{-2}, g^{-1}, \text{ﻉ}, g, g^2, \ldots\right\}.$$

We call this the **group generated by** g, and call g the **generator** of this group. When we're feeling a little lazy, which is actually pretty common, we simply say the **group generated by** g. Every cyclic group has the form $\langle g \rangle$ for some $g \in G$. Is the converse also true that $\langle g \rangle$ is a group for any $g \in G$? As a matter of fact, yes!

Theorem 3·49. *For every group G and for every $g \in G$, $\langle g \rangle$ is an abelian group.*

Proof. We show that $\langle g \rangle$ satisfies the properties of an abelian group. Let $x, y, z \in \langle g \rangle$. By definition of $\langle g \rangle$, there exist $a, b, c \in \mathbb{Z}$ such that $x = g^a$, $y = g^b$, and $z = g^c$. We will use Lemma 3·39 implicitly.

- By substitution, $xy = g^a g^b = g^{a+b} \in \langle g \rangle$. So $\langle g \rangle$ is closed.

- By substitution, $x(yz) = g^a\left(g^b g^c\right)$. These are elements of G by inclusion (that is, $\langle g\rangle \subseteq G$ so $x, y, z \in G$), so the associative property *in* G gives us

$$x(yz) = g^a\left(g^b g^c\right) = \left(g^a g^b\right)g^c = (xy)z.$$

- By definition, $\mathfrak{s} = g^0 \in \langle g\rangle$.
- By definition, $g^{-a} \in \langle g\rangle$, and $x \cdot g^{-a} = g^a g^{-a} = e$. Hence $x^{-1} = g^{-a} \in \langle g\rangle$.
- Using the fact that $\mathbb{Z}$ is commutative under addition,

$$xy = g^a g^b = g^{a+b} = g^{b+a} = g^b g^a = yx.$$

□

Question 3·50 .
Find all the generators of Ω_8. *Hint:* In Question 3.32 you showed that $\Omega_n \cong \mathbb{Z}_n$, so the generators of $\mathbb{Z}_8$ must correspond to the generators of Ω_8. The mapping you used in the isomorphism will tell you which ones.

The order of an element

Given an element and an operation, Theorem 3·49 links them to a group. It makes sense, therefore, to link an element to the order of the group that it generates.

Definition 3·51. Let G be a group, and $g \in G$. We say that the **order** of g is the order of the group it generates; $\operatorname{ord}(g) = |\langle g\rangle|$. If $\operatorname{ord}(g) = \infty$, we say that g has **infinite order**.

We can write an element in different ways when its order is finite.

Example 3·52. Consider $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. Since $4 \equiv_4 0$, we can write 1 as $1 \times 4 + 1$, $2 \times 4 + 1$, $3 \times 4 + 1$, etc.

Example 3·53. Recall $\Omega_7 = \{1, \omega, \omega^2, \ldots, \omega^6\}$. Since $\omega^7 = 1$, we can write ω^2 as ω^2, ω^9, ω^{16}, etc.

The example suggests that if the order of an element G is $n \in \mathbb{N}$, then we can write

$$\langle g\rangle = \left\{\mathfrak{s}, g, g^2, \ldots, g^{n-1}\right\}.$$

This explains why we call $\langle g\rangle$ a *cyclic* group: once they reach $\operatorname{ord}(g)$, the powers of g "cycle". To prove this in general, we have to show that for a finite cyclic group $\langle g\rangle$ with $\operatorname{ord}(g) = n$,

- n is the smallest positive power that gives us the identity; that is, $g^n = \mathfrak{s}$, and
- for any two integers between 0 and n, the powers of g are different; that is, if $0 \leqslant a < b < n$, then $g^a \neq g^b$.

Theorem 3·54 accomplishes that, and a bit more as well.

Theorem 3·54. *Let G be a finite group, $g \in G$, and* $\operatorname{ord}(g) = n$.

(A) $я, g, g^2, \ldots, g^{n-1}$ *are all distinct.*

(B) $g^n = я$;

(C) n *is the smallest positive integer* d *such that* $g^d = я$; *and*

(D) For any $a, b \in \mathbb{Z}$, $n \mid (a-b)$ *if and only if* $g^a = g^b$.

Proof. The meat of the theorem is (A). The remaining assertions are consequences.

(A) By way of contradiction, suppose that there exist $a, b \in \mathbb{N}$ such that $0 \leqslant a < b < n$ and $g^a = g^b$; then $я = (g^a)^{-1} g^b$. By Lemma 3·39, we can write

$$я = g^{-a} g^b = g^{-a+b} = g^{b-a}.$$

Let $d = b - a$. Recall that $a < b$, so $d = b - a \in \mathbb{N}^+$. By the Division Theorem, for any integer m we can find $q, r \in \mathbb{Z}$ such that $m = qd + r$ and $0 \leqslant r < d$. Applying Lemma 3·39 again, we have

$$g^m = g^{qd+r} = \left(g^d\right)^q g^r = я^q g^r = g^r,$$

so *any* power of g can be written as a remainder after division by d. In other words,

$$\langle g \rangle = \left\{я, g, g^2, \ldots, g^{d-1}\right\}.$$

This implies that $|\langle g \rangle| = d$, which contradicts the assumption that $n = \operatorname{ord}(g) = |\langle d \rangle|$.

(B) We know that $\operatorname{ord}(g) = n$, so there are n distinct elements of $\langle g \rangle$. By part (a), the n powers $g^0, g^1, \ldots, g^{n-1}$ are all distinct, so

$$\langle g \rangle = \left\{g^0, g^1, \ldots, g^{n-1}\right\}.$$

This implies that $g^n = g^d$ for some $d = 0, 1, \ldots, n-1$. Which one?

Using Lemma 3·39, we find that $g^{n-d} = я$. Recall that $0 \leqslant d < n$, so $0 < n - d \leqslant n$. By (A), $g^a \neq я$ for $a = 1, 2, \ldots, n-1$, so $n - d = n$, so $d = 0$. By substitution, $g^n = g^d = g^0 = я$.

(C) let S is the set of all positive integers m such that $g^m = я$; this is a subset of $\mathbb{N}$, so it has a smallest element. Let the smallest element be d; by (B), $g^n = я$, so $n \in S$. Hence $d \leqslant n$. On the other hand, (A) tells us that we cannot have $d < n$; otherwise, $g^d = g^0 = я$. Hence, $n \leqslant d$. We already had $d \leqslant n$, so the two must be equal.

(D) Let $a, b \in \mathbb{Z}$. Assume that $n \mid (a-b)$. Let $q \in \mathbb{Z}$ such that $nq = a - b$. Substitution, Lemma 3·39 and some arithmetic tell us that

$$\begin{aligned} g^b &= g^b \cdot я = g^b \cdot я^q \\ &= g^b \cdot (g^n)^q = g^b \cdot g^{nq} \\ &= g^b \cdot g^{a-b} = g^{b+(a-b)} = g^a. \end{aligned}$$

Conversely, if we assume that $g^b = g^a$, then Lemma 3·39 implies that $g^{b-a} = я$. Use the Division Theorem to choose $q, r \in \mathbb{Z}$ such that $b - a = nq + r$ and $0 \leqslant r < n$. By substitution and Lemma 3·39,

$$я = g^{b-a} = g^{qn+r} = (g^n)^q g^r = я^q g^r = g^r.$$

Recall that $0 \leqslant r < n$. By (C), r cannot be positive, so $r = 0$. By substitution, $b - a = qn$, so $n \mid (b - a)$.

□

We conclude that, at least when they are finite, cyclic groups are aptly named: increasing powers of g generate new elements until the power reaches n, in which case $g^n = я$ and we "cycle around."

Question 3·55 .

Complete the proof of Lemma 3·39(C).

Question 3·56 .

Fill in each blank of Figure 3.56 with the justification or statement.

Let G be a group, and $g \in G$. Let $d, n \in \mathbb{Z}$ and assume $\text{ord}(g) = d$.
Claim: $g^n = я$ if and only if $d \mid n$.
Proof:

1. Assume that $g^n = я$.
   (a) By _____, there exist $q, r \in \mathbb{Z}$ such that $n = qd + r$ and $0 \leqslant r < d$.
   (b) By _____, $g^{qd+r} = я$.
   (c) By _____, $g^{qd}g^r = я$.
   (d) By _____, $\left(g^d\right)^q g^r = я$.
   (e) By _____, $я^q g^r = я$.
   (f) By _____, $яg^r = я$. By the identity property, $g^r = я$.
   (g) By _____, d is the *smallest* positive integer such that $g^d = я$.
   (h) Since _____, it cannot be that r is positive. Hence, $r = 0$.
   (i) By _____, $n = qd$. By definition, then $d \mid n$.
2. Now we show the converse. Assume that _____.
   (a) By definition of divisibility, _____.
   (b) By substitution, $g^n =$_____.
   (c) By Lemma 3·39, the right hand side of that equation can be rewritten as _____.
   (d) Recall that $\text{ord}(g) = d$. By Theorem 3·54, $g^d = я$, so we can rewrite the right hand side again as _____.
   (e) A little more simplification turns the right hand side into _____, which obviously simplifies to e.
   (f) By _____, then, $g^n = я$.
3. We showed first that if $g^n = я$, then $d|n$; we then showed that _____. This proves the claim.

Figure 3·2: Material for Question 3.56

3·4 An introduction to finite rings and fields

> *Our minds are finite, and yet even in these circumstances of finitude we are surrounded by possibilities that are infinite, and the purpose of life is to grasp as much as we can out of that infinitude.*
>
> — Alfred North Whitehead

The rings and fields you're most familiar with are infinite: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. A natural question to ask is, "Do finite rings or fields exist?"

We'll look at rings first. You saw in Section 2·2 that $\mathbb{Z}_d$ is an abelian group under addition, one of the requirements of a ring.

Theorem 3·57. *For any nonzero integer d, the set $\mathbb{Z}_d$ is a commutative ring under modular addition and multiplication.*

Proof. Let $d \in \mathbb{Z}$ be nonzero, and let $a, b, c \in \mathbb{Z}_d$. We already know that $\mathbb{Z}_d$ makes an abelian group under modular addition, so we need merely show that modular multiplication satisfies the requirements of a commutative monoid. Closure is guaranteed by property (D2) of the Division Theorem. The multiplicative identity is 1, itself a remainder and thus an element of $\mathbb{Z}_n$. The associative property follows from multiplication of the integers and from the uniqueness of remainders: since $a(bc) = (ab)c$ as integers, the unique remainders of $a(bc)$ and $(ab)c$ must also be equal, so $a(bc) \equiv (ab)c$. □

So $\mathbb{Z}_d$ is a finite ring for every nonzero value of d.

As for finite *fields,* ah, uhm… well! You already met zero divisors of finite rings in Question 2.40, so at least one of our finite rings are not good candidates for finite fields. Other finite rings work dandily.

Example 3·58. Recall $\mathbb{Z}_7 = \{0, 1, 2, \ldots, 6\}$. We can see that this is a field by verifying that every nonzero element has a multiplicative inverse: $1 \otimes 1 = 1$, $2 \otimes 4 = 8 \equiv 1$, $3 \otimes 5 = 15 \equiv 1$, and $6 \otimes 6 = 36 \equiv 1$.

So $\mathbb{Z}_7$ is a field, but $\mathbb{Z}_6$ is not.

Question 3·59 .
What difference between $\mathbb{Z}_6$ and $\mathbb{Z}_7$ makes the latter a field, while the former is not?

Don't draw *too* hasty a conclusion! You might be tempted to think that the *only* finite field are those of the $\mathbb{Z}_d$, where d has the "correct" form. In fact, there *can* be other fields of size d!

Example 3·60. Consider $g = x^2 + 1$, in the ring $\mathbb{Z}_3[x]$. Let $\mathbb{F}_9$ be the set of remainders possible when dividing by g. Arithmetic is modulo *both* 3 *and* $x^2 + 1$, so its elements are

$$\mathbb{F}_9 = \{0, 1, 2, x, x+1, x+2, 2x, 2x+1, 2x+2\}.$$

This set is not the same as $\mathbb{Z}_9$!

You already know from Theorem 3·7 that $\mathbb{F}_9$ forms a ring. It is routine to verify that:

$$\begin{array}{ll} 1^{-1} = 1 & (x+1)^{-1} = x+2 \\ 2^{-1} = 2 & (2x+1)^{-1} = 2x+2 \\ x^{-1} = 2x & \end{array}$$

For example,

$$(2x+1)(2x+2) = 4x^2 + \cancelto{0}{6x} + 2 \equiv x^2 + 2 \equiv 1.$$

So all its nonzero elements have inverses.

Even though it has nine elements, $\mathbb{F}_9$ is a field! So we can't just look at whether the number of elements in a set factors. That said, it's not completely unrelated.

Characteristics of finite rings

Definition 3·61. Let R be a ring. The **characteristic** of R is the smallest positive integer n such that

$$0 \equiv nr = \underbrace{r + r + \cdots + r}_{n \text{ times}}$$

for every $r \in R$. If there is no such number, we say that R has **characteristic 0**.

If n is the characteristic of R, we write $\text{char}R = n$.

Example 3·62. In $\mathbb{Z}_7$, the characteristic of 1 is 7, since $7 \times 1 \equiv 0$, and any smaller multiple of 1 is non-zero. In fact, $7 \times r \equiv 0$ for any nonzero element r, and no smaller value of n gives $n \times r \equiv 0$. The one exception is $r = 0$, in which case $1 \times 0 \equiv 0$, but 1 doesn't work for the other elements ($1 \times 2 \not\equiv 0$), whereas 7 works for 0 ($7 \times 0 \equiv 0$), so the characteristic of $\mathbb{Z}_7$ is indeed 7.

In $\mathbb{Z}_6$, the relationship $2 \times 3 \equiv 0$ suggests that the characteristic could be 2 or 3, but neither number is the characteristic of every element, since $2 \times 1 \equiv 2 \not\equiv 0$ and $3 \times 5 \equiv 3 \not\equiv 0$. We have to try something larger, and in fact neither $4 \times 5 \equiv 0$ nor $5 \times 5 \equiv 0$; we find $6 \times 5 \equiv 0$. Similarly, $6 \times 1 \equiv 0$, so the characteristic of $\mathbb{Z}_6$ is 6.

Don't jump too quickly into thinking the characteristic of a ring is simply the number of elements! In $\mathbb{F}_9$, we get a different answer, because everything is modulo 3, so $3 \times r \equiv 0$ for every $r \in \mathbb{F}_9$. Smaller numbers won't work as $1 \times 1 \not\equiv 0$, and $2 \times 1 \not\equiv 0$, so the characteristic must in fact be 3.

While the characteristic of a ring is defined in terms of *every* element, it actually depends on *only one* element!

Theorem 3·63. *The characteristic of a ring is either zero or the smallest positive number n such that $n \times 1 = 0$, where 1 is the multiplicative identity of R.*

Proof. Let R be a ring, and $r \in R$. If $n \times 1 \neq 0$ for any $n \in \mathbb{N}^+$, then by definition of characteristic, $\text{char}R = 0$. Otherwise, R has positive characteristic, and there exists $n \in \mathbb{N}^+$ such that $n \times 1 =$

0; use the Well-Ordering Principle to choose the smallest such n. By closure, $n \in R$, so we can apply the associative property to see that

$$n \times r = n \times (1 \times r) = (n \times 1) \times r = 0 \times r = 0.$$

(Notice the use of Fact 2·41 in the last step.) Thus, any $n \in \mathbb{N}^+$ satisfying $n \times 1 = 0$ also satisfies $n \times r = 0$. By choice of n, no smaller positive m satisfies $m \times 1 = 0$, so $n \times r = 0$ for all $r \in R$, and is the smallest such. The characteristic of R depends entirely on its multiplicative identity. □

It turns out that the *characteristic* is the key property distinguishing fields from mere rings.

Theorem 3·64. *The characteristic of a field is either zero or irreducible.*

Proof. Let $\mathbb{F}$ be a field of characteristic n. Suppose to the contrary that $n = pq$, where p and q are integers, but neither is 1. Notice that $1 < p, q < n$. Let $S = \{0, 1, 2 \times 1, \ldots, (n-1) \times 1\}$. By closure of multiplication, $S \subseteq \mathbb{F}$. In addition, S is a set with n *distinct* elements; otherwise, we would contradict Theorem 3·63. (Keep in mind that 2×1 means $1 + 1$, $3 \times 1 = 1 + 1 + 1$, etc.)

Of course, 1 is the multiplicative identity, so $S = \{0, 1, 2, \ldots, n-1\}$. Recall that $p, q < n$. That means $p, q \in S$; by inclusion, $p, q \in \mathbb{F}$. Closure of multiplication forces $p \times q \in \mathbb{F}$. By the definition of characteristic, $p \times q = n \equiv 0$, so that $\mathbb{F}$ has zero divisors. This contradicts Question 2.45 and the hypothesis that $\mathbb{F}$ is a field!

The only questionable assumption we have made is that neither p nor q is 1, so it must be that one of them *is* 1, and n is irreducible. □

But what if n *is* irreducible?

Fact 3·65. *If n is irreducible, then $\mathbb{Z}_n$ is a field of characteristic n.*

Proof. Certainly $\mathbb{Z}_n$ is a *ring* of characteristic n, since $i \times 1 \not\equiv 0$ for any $i = 1, \ldots, n-1$. Why must it be a field? We claim that for any nonzero $r \in \mathbb{Z}_n$ we can find $s \in \mathbb{Z}_n$ such that $rs \equiv 1$. To see why, we need the following lemma.

Bézout's Lemma. *If d is the largest integer that divides two integers m and n, then we can find integers x and y such that $mx + ny = d$. In fact, d is the* smallest *positive integer for which we can find such an expression.*

The equation $mx + ny = d$ is sometimes called **Bézout's Identity**. The integer d of Bézout's Lemma is the **greatest common divisor** of m and n, and is abbreviated $\gcd(m, n)$.

Proof of Bézout's Lemma. Let $S = \{mx + ny : x, y \in \mathbb{Z}\}$, and let $L = S \cap \mathbb{N}^+$. By the Well-Ordering Principle, L has a smallest element; call it ℓ, and choose x and y such that $mx + ny = \ell$. By hypothesis, d divides both m and n; say $m = ad$ and $n = bd$. By substitution,

$$(ad)\,x + (bd)\,y = \ell.$$

We can rewrite this as

$$d\,(ax + by) = \ell,$$

so $d \mid \ell$. We know that this means $d \leqslant \ell$.

On the other hand, choose a quotient q and remainder r such that $m = q\ell + r$ satisfies the Division Theorem. Rewrite this equation as

$$r = m - q\ell = m - q\,(mx + ny) = m\,(1 - qx) + n\,(-qy)\,.$$

With $r = m\,(1 - qx) + n\,(-qy)$, we see that $r \in S$. As a remainder, $r \in \mathbb{N}$, so either $r = 0$ or $r \in S \cap \mathbb{N} = L$. If $r \neq 0$, the choice of ℓ as the smallest element of L implies $\ell \leqslant r$. But r is a remainder from division by d, so $r < d$, and we saw above that $d \leqslant \ell$; it doesn't make sense to have $\ell \leqslant r < d \leqslant \ell$! The only way to avoid a contradiction is if $r = 0$, so ℓ divides m. A similar argument shows that ℓ divides n. We now have ℓ dividing both m and n; recall that d is the largest integer that divides both m and n, so $\ell \leqslant d$.

We are now finished: the first paragraph concluded that $d \leqslant \ell$, and the second paragraph concluded that $\ell \leqslant d$. This is only possible if $\ell = d$, and we have shown the claim.

Question 3·66 .
The first paragraph of the proof of Bézout's Theorem concludes with the assertion that if d, ℓ are both positive integers, and d divides ℓ, then $d \leqslant \ell$. Why must this be true?

We return to our main question.

Proof of Fact 3·64 (continued). Let $m \in \mathbb{Z}_n$. By hypothesis, n is irreducible, so the greatest common divisor of m and n is 1. Well, then, Bézout's Lemma gives us integers x, y such that $mx + ny = 1$. Rewrite this as $ny = 1 - mx$, and Theorem 2·19 shows that $1 \equiv_n mx$. In other words, x is the multiplicative inverse we sought for m. □

Evaluating positions in the game

We return to the question of evaluating the value of a position in Nim and Ideal Nim. One way to do this is to count the number of possible moves remaining. For instance, if we have only a single row of m boxes, we would call that a row of value m. How can we model this? Let's start with these first two principles to keep in mind:

Principle the first: A choice's value must satisfy $0 \leqslant m$.

Principle the second: Any choice is its own inverse, so $m \oplus m = 0$.

To a seasoned mathematician, the self-inverse property indicates that we're working in a ring of characteristic 2. We'll aim for a field, if we can get it. Unfortunately, the basic field of characteristic 2 is $\mathbb{Z}_2$, which has only two values. By themselves, 0 and 1 won't model our game, so we'll have to extend our ring. Nothing stops us from extending it in a fashion similar to the one we used to build the complex numbers, so let's try that.

Fact 3·67. *The polynomial* $f = x\,(x-1)\,(x-a_1)\cdots(x-a_{n-2}) + 1$ *has no roots in the finite field* $\mathbb{F}_n = \{0, 1, a_1, \ldots, a_{n-2}\}$.

Why? We can see this by simple substitution; $f\,(b) \equiv 1$ for any element b of $\mathbb{F}_n$. □

Fact 3·68. *Any factorization of f that uses coefficients only in $\mathbb{F}_n$ has no linear components.*

Why? The alternative would set up a contradiction between the Division Theorem for Polynomials and the previous fact: for any hypothetical linear factor of f, the definition of f would have a remainder of 1, while the factorization would have a remainder of 0. □

In other words, while f may factor, its irreducible factors are not linear.

Fact 3·69. *Defining a ring $\mathbb{E}$ as $\mathbb{F}_n[x]$ modulo an irreducible factor of f actually gives us a field.*

Why? If not, there must be some nonzero element a of $\mathbb{E}$ that does not have a multiplicative inverse. Since $\mathbb{E}$ is finite, we can list all products ax for $x \in \mathbb{E}$. The fact that $ax \neq 1$ means there must be distinct elements $x, y \in \mathbb{E}$ whose products give $ax = ay$. Rewrite this as $a(x - y) = 0$. Let $z = x - y$; with distinct x and y, we must have $z \neq 0$. That means $az = 0$ even though $a, z \neq 0$; we have found zero divisors.

This is a contradiction! To see why, let g be the irreducible factor of f. Both a and z are polynomials with degree smaller than $\deg g$. The statement "$az = 0$ in $\mathbb{E}$" translates to "$az \equiv 0$ in $\mathbb{F}_n[x]$ modulo g," but since $0 \equiv g$, we have found a factorization of an irreducible polynomial! □

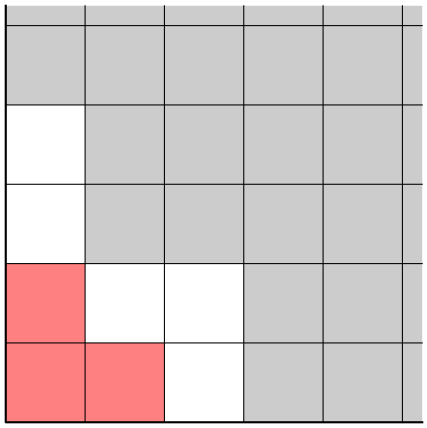
Satisfied that $\mathbb{E}$ is in fact a field, we can now build successively larger fields

$$\{0, 1\} \subsetneq \{0, 1, 2, 3\} \subsetneq \{0, 1, 2, 3, 4, 5, 6, 7\} \subsetneq \cdots$$

where 2^n represents x^n in the corresponding extension, $2^n + 1$ represents $x^n + 1$, $2^n + 2$ represents $x^n + x$, etc. These are not your ordinary 2, 3, ... because here, $1 + 1 \not\equiv 2$; after all, $1 + 1 \equiv 0$. The addition of the remainders, modulus the irreducible polynomial, corresponds precisely to integer addition using powers of 2, also called **binary notation**. This allows us to model the Nim and Ideal Nim.

It is not enough to form a winning strategy for Ideal Nim, because a winning strategy for this game is *currently unknown!* However, we can still evaluate the values of many games using the implication of symmetry that $x + x \equiv 0$.

Example 3·70. Suppose a game of Ideal Nim has led to the following configuration:



Suppose it is Emmy's turn. As this is visually non-symmetric, you might conclude that the Emmy has an advantage. Upon further inspection, you'd discover that if David has any sense at all, then no, *Emmy will in fact lose.* For instance, if Emmy chooses $(0, 3)$, David could choose $(1, 1)$, leaving a visually symmetric game; if Emmy chooses $(0, 2)$ instead, David could choose $(2, 1)$, again leaving a visually symmetric game. *Either choice is a win for David!* The remaining choices for the next player are similarly parried.

Our conclusion from this is that the value of the upside-down L on the right is equivalent to the value of the two blocks in a line on the left; both blocks have value 2.

Question 3·71 .

Operating under the assumption that a line of m blocks has value m, use a technique analogous to the one in the previous example to show that the values of the following configurations are 1, 3, 4, 1 (again!), and 5.

3·5 Matrices

> **matrix**, *n. 1.* (Latin) *the womb. 2.* (mathematics) *A rectangular array of numeric or algebraic quantities subject to mathematical operations.*
>
> — from *The American Heritage Dictionary of the English Language* (4th edition)

Let R be a commutative ring, and $m, n \in \mathbb{N}^+$. An $m \times n$ **matrix** M **over** R is a list of m lists (**rows**) of n elements of R. We say the **dimension** of the matrix is $m \times n$. We call R the **base ring** of M. If $m = n$, we call the matrix **square**, and say that the **dimension** of the matrix is m. The set of all $m \times n$ matrices over R is $R^{m \times n}$.

Notation 3·72. We write the jth element of row i of the matrix A as a_{ij}. We often omit 0's from the matrix, not so much from laziness as from a desire to improve readability. (It really does help to omit the 0's when there are a lot of them.) If the dimension of A is $m \times n$, then we write $\dim A = m \times n$.

Example 3·73. If

$$A = \begin{pmatrix} 1 & & 1 \\ & 1 & \\ & 5 & 1 \end{pmatrix},$$

then $a_{21} = 0$ while $a_{32} = 5$. Notice that A is a 3×3 matrix; or, $\dim A = 3 \times 3$. As a square matrix, we say its dimension is 3.

Definition 3·74. The **transpose** of a matrix A is the matrix B satisfying $b_{ij} = a_{ji}$. In other words, the jth element of row i of B is the ith element of row j of A. A **column** of a matrix is a row of its transpose.

Notation 3·75. We often write A^T for the transpose of A.

Example 3·76. If A is the matrix of the previous example, then

$$A^T = \begin{pmatrix} 1 & & \\ & 1 & 5 \\ 1 & & 1 \end{pmatrix}.$$

We focus mostly on square matrices, with the exception of $m \times 1$ matrices, also called **column vectors**, or just plain "vectors" if we feel lazy, as we often do. The **dimension** of an $m \times 1$ vector is m. We write R^m for the set of all column vectors of dimension n with entries from a ring R. This looks the same as the Cartesian product $R \times R \times \cdots R$, because it is: a column vector $(r_1 \ \cdots \ r_m)^T$ is merely a different representation of writing the tuple $(r_1, \ldots, r_m)$.

Matrix arithmetic

The two major operations for matrices are addition and multiplication. Addition is componentwise; we *add* matrices by adding entries in the same row and column. Multiplication is not componentwise.

- If A and B are $m \times n$ matrices and $C = A + B$, then $c_{ij} = a_{ij} + b_{ij}$ for all $1 \leqslant i \leqslant m$ and all $1 \leqslant j \leqslant n$. Notice that C is also an $m \times n$ matrix.
- If A is an $m \times r$ matrix, B is an $r \times n$ matrix, and $C = AB$, then C is the $m \times n$ matrix whose entries satisfy

  $$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj} \ ;$$

  that is, the jth element in row i of C is the sum of the products of corresponding elements of row i of A and column j of B.

This definition of multiplication, while odd, satisfies certain useful properties: in particular, relating matrix equations to systems of linear equations.

Example 3·77. If A is the matrix of the previous example and

$$B = \begin{pmatrix} 1 & 5 & -1 \\ & 1 & \\ & -5 & 1 \end{pmatrix},$$

then

$$\begin{aligned} AB &= \begin{pmatrix} 1\cdot 1 + 0\cdot 0 + 1\cdot 0 & 1\cdot 5 + 0\cdot 1 + 1\cdot -5 & 1\cdot -1 + 0\cdot 0 + 1\cdot 1 \\ 0\cdot 1 + 1\cdot 0 + 0\cdot 0 & 0\cdot 5 + 1\cdot 1 + 0\cdot -5 & 0\cdot -1 + 1\cdot 0 + 0\cdot 1 \\ 0\cdot 1 + 5\cdot 0 + 1\cdot 0 & 0\cdot 5 + 5\cdot 1 + 1\cdot -5 & 0\cdot -1 + 5\cdot 0 + 1\cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}. \end{aligned}$$

On the other hand, if $\mathbf{x} = (x\ y\ z)^T$ and $\mathbf{b} = (0\ 0\ 2)^T$, the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

simplifies to

$$\begin{pmatrix} x & & z \\ & y & \\ & 5y & z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix},$$

which corresponds to the system of equations

$$\begin{cases} x & & +z = 0 \\ & y & = 0 \\ & 5y + z = 2 \end{cases}.$$

Question 3·78 .

Recall the definition of zero divisors from Definition 2·42. Show that matrix multiplication has zero divisors by finding two square matrices A and B such that $A \neq \mathbf{0}$ and $B \neq \mathbf{0}$, but $AB = \mathbf{0}$. You can start with 2×2 matrices, but try to make it a general formula, and describe how one could build such matrix zero divisors regardless of their size. *Hint:* Don't overthink this; there is a very, *very* simple answer.

Question 3·79 .

In this problem, pay careful attention to which symbols are **thickened**, as they represent matrices.

Let i denote the imaginary number, so that $i^2 = -1$, and let Q_8 be the set of **quaternions**, defined by the matrices $\{\pm\mathbf{1}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$ where

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$
$$\mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ \mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

(a) Show that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}$.

(b) Show that $\mathbf{ij} = \mathbf{k}$, $\mathbf{jk} = \mathbf{i}$, and $\mathbf{ik} = -\mathbf{j}$.

(c) Show that $\mathbf{ij} = -\mathbf{ji}$, $\mathbf{ik} = -\mathbf{ki}$, and $\mathbf{jk} = -\mathbf{kj}$.

(d) Use these properties to construct the Cayley table of Q_8. *Hint:* If you use the properties carefully, along with what you know of linear algebra, you can fill in the remaining spaces without performing a single matrix multiplication.

(e) Show that Q_8 is a group under matrix multiplication.

(f) Explain why Q_8 is *not* an abelian group.

Question 3·80 .
The following exercises refer to elements of the quaternions (Question 3.79).

(a) Determine the elements of $\langle -\mathbf{1} \rangle$ and $\langle \mathbf{j} \rangle$ in Q_8.

(b) Verify that $H = \{\mathbf{1}, -\mathbf{1}, \mathbf{i}, -\mathbf{i}\}$ is a cyclic group. Which elements actually generate H?

(c) Show that Q_8 is not cyclic.

Question 3·81 .
In each of the following, compute the order of the element $a \in Q_8$.

(a) $a = \mathbf{i}$

(b) $a = -\mathbf{1}$

(c) $a = \mathbf{1}$

Question 3·82 .
We sometimes allow matrices which proceed indefinitely in two directions. Here are two such matrices which are mostly zero, though we highlight the zeros on the main diagonal:

$$D = \begin{pmatrix} \mathbf{0} & 1 & & & & \\ & \mathbf{0} & 2 & & & \\ & & \mathbf{0} & 3 & & \\ & & & \mathbf{0} & 4 & \\ & & & & \ddots & \ddots \end{pmatrix} \quad S = \begin{pmatrix} \mathbf{0} & & & & & \\ \frac{1}{2} & \mathbf{0} & & & & \\ & \frac{1}{3} & \mathbf{0} & & & \\ & & \frac{1}{4} & \mathbf{0} & & \\ & & & \frac{1}{5} & \mathbf{0} & \\ & & & & \ddots & \ddots \end{pmatrix}.$$

Let R be a ring. A polynomial in $R[x]$ corresponds to a **coefficient vector** via the map

$$r_n x^n + \cdots r_1 x + r_0 \mapsto \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_n \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

Choose several random polynomials p, write their coefficient vectors $\mathbf{p}$, then compute $D\mathbf{p}$ and $S\mathbf{p}$ for each. What are the results? How would you characterize the effect of multiplying D and S to a "polynomial vector"?

Definition 3·83. The **kernel** of a matrix M is the set of vectors $\mathbf{v}$ such that $M\mathbf{v} = \mathbf{0}$. In other words, the kernel is the set of vectors whose product with M is the zero matrix.

Notation 3·84. We write $\ker M$ for the kernel of M.

Example 3·85. Let $R = \mathbb{Z}$, and

$$M = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} -5 \\ 0 \\ 1 \end{pmatrix}.$$

Since

$$M\mathbf{x} = \begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix} \quad \text{and} \quad M\mathbf{y} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \mathbf{0},$$

we see that $\mathbf{x}$ is not in the kernel of M, but $\mathbf{y}$ is. In fact, it can be shown (you will do so in a moment) that

$$\ker M = \left\{ v \in R^3 : \ v = \begin{pmatrix} -5c \\ 0 \\ c \end{pmatrix} \ \exists c \in \mathbb{F} \right\}.$$

The kernel has important and fascinating properties, which we explore later on.

Question 3·86 .

Let $R = \mathbb{Z}$, and

$$M = \begin{pmatrix} 1 & & 1 \\ & 1 & \\ & 5 & -1 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Show that

$$\ker M = \{\mathbf{0}\},$$

and

$$\ker N = \left\{ v \in R^3 : \ v = \begin{pmatrix} -5c \\ 0 \\ c \end{pmatrix} \ \exists c \in R \right\}.$$

Question 3·87 .

What are the kernels of the matrices D and I of Question 3.82? *Hint:* In that problem, we asked you to "characterize the effect" of D and S on a "polynomial vector." If you know the effect, you can use that to make an educated guess at what appears in the kernel, then prove it.

Properties of matrix arithmetic

We now explore some properties of arithmetic of square matrices, so as to find a structure that describes them.

Fact 3·88. *For a fixed dimension of square matrices, matrix addition and multiplication are closed.*

Why? The hypothesis is that we have a *fixed* dimension of square matrices, say $n \times n$. *From the definition of the operations,* you see immediately that both addition and multiplication of matrices result in an $n \times n$ matrix. Thus, any $A, B \in R^{m\times n}$ satisfy $A+B \in R^{m\times n}$ and $AB \in R^{m\times n}$. □

Recall A and B from Examples 3·73 and 3·77. If we write I_3 for a 3×3 matrix of three 1's on the diagonal (and zeroes elsewhere), something interesting happens:

$$AI_3 = I_3A = A \quad \text{and} \quad BI_3 = I_3B = B.$$

The pattern of this matrix ensures that the property remains true for *any* matrix, as long as you're working in the correct dimension. That is, I_3 is an "identity" matrix. In particular, it's the identity of **multiplication**. Is there a second identity matrix?

Don't confuse "the identity matrix" with a matrix filled with zeros; that is the identity matrix for *addition*. Can there be another second matrix for *multiplication?* In fact, there *cannot*. You will see why in a moment.

Notation 3·89.

- We write **0** (that's a **bold zero**) for any matrix whose entries are all zero.
- We write I_n for the $n \times n$ matrix satisfying
  - $a_{ii} = 1$ for any $i = 1, 2, \ldots, n$; and
  - $a_{ij} = 0$ for any $i \neq j$.

Theorem 3·90. *The zero matrix* **0** *is an identity for matrix addition. The matrix I_n is an identity for multiplication of $n \times n$ matrices.*

When reading theorems, you sometimes have to read between the lines. Here, you have to infer that $n \in \mathbb{N}^+$ and **0** is a matrix whose dimension is appropriate to the other matrix. We should *not* take it to mean an $m \times 4$ matrix with zero entries is an identity for matrices of dimension $m \times 2$, as the addition would be undefined. Similarly, you have to infer that I_n is an identity for square matrices of dimension n; it wouldn't make sense to multiply I_n to a 3×5 matrix.

Question 3·91 .
Can you find a multiplicative identity for 3×5 matrices? If so, what it is? If not, why not?

Proof of Theorem 3·90. Let A be a square matrix of dimension $m \times n$. By definition, the jth element in row i of $A + \mathbf{0}$ is $a_{ij} + 0 = a_{ij}$. This is true regardless of the values of i and j, so if we choose **0** to be an $m \times n$ matrix with zero entries, $A + \mathbf{0} = A$. A similar argument shows $\mathbf{0} + A = A$. Since A is arbitrary, **0** really is an additive identity.

As for I_n, we point out that the jth element of row i of AI_n is (by definition of multiplication)

$$\begin{pmatrix} & & \text{col } j & & \\ & & \vdots & & \\ \text{row } i & \cdots & \text{this element?} & \cdots & \\ & & \vdots & & \end{pmatrix} = \begin{pmatrix} \text{row } i & a_1 & \cdots & a_{ij} & \cdots a_{in} \end{pmatrix} \begin{pmatrix} 1 & & \text{col } j & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = a_{ij} \cdot 1 + \sum_{\substack{k=1,\ldots,m \\ k \neq j}} a_{ik} \cdot 0.$$

Simplifying this gives us a_{ij}. This is true regardless of the values of i and j, so $AI_n = A$. A similar argument shows that $I_n A = A$. Since A is arbitrary, I_n really is a multiplicative identity. □

Given a matrix A, an **additive inverse** of A is any matrix B such that $A + B = \mathbf{0}$. A **multiplicative inverse** of A is any matrix B such that $AB = I_n$. Additive inverses always exist, and it is easy to construct them. Multiplicative inverses *do not* exist for some matrices, even when the matrix is square. Because of this we call a matrix **invertible** if it has a multiplicative matrix, and if we merely speak of the "inverse" of a matrix, we mean its multiplicative inverse.

Notation 3·92. We write the additive inverse of a matrix A and $-A$, and the multiplicative inverse of A as A^{-1}.

Example 3·93. The matrices A and B of the previous example are inverses; that is, $A = B^{-1}$ and $B = A^{-1}$. The non-zero matrix

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$$

is *not* invertible, because any matrix satisfying

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = I_2$$

must satisfy the system of equations

$$\begin{cases} a = 1 \\ b = 0 \\ 2a = 0 \\ 2b = 1 \end{cases},$$

an impossible task.

Question 3·94 .

A matrix A is **orthogonal** if its transpose is also its inverse. Let $n \in \mathbb{N}^+$ and $\mathrm{O}(n)$ be the set of all orthogonal $n \times n$ matrices.

(a) Show that this matrix is orthogonal, regardless of the value of α:

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

(b) Find some other orthogonal matrices. (Their entries can consist of numbers alone.) Compute their determinant. Do you notice a pattern? See if you can prove it.

Hint: The easiest way to show this requires some properties of determinants. Since you may not remember them, or may not even have *seen* them (it could depend on the class, on the teacher, on which universe you existed in the day they were presented…) here are the ones you need: for any matrix that has an inverse, $\det A = \det A^T$, $\det(AB) = (\det A)(\det B)$, and $\det I_n = 1$ for every $n \in \mathbb{N}^+$.

We want one more property.

Theorem 3·95. *Matrix multiplication is associative. That is, if A, B, and C are matrices, then* $A(BC) = (AB)C$.

Proof. Let A be an $m \times r$ matrix, B an $r \times s$ matrix, and C an $s \times n$ matrix. By definition, the ℓth element in row i of AB is

$$(AB)_{i\ell} = \sum_{k=1}^{r} a_{ik} b_{k\ell}.$$

Likewise, the jth element in row i of $(AB)C$ is

$$((AB)C)_{ij} = \sum_{\ell=1}^{s} (AB)_{i\ell} c_{\ell j} = \sum_{\ell=1}^{s} \left[\left(\sum_{k=1}^{r} a_{ik} b_{k\ell} \right) c_{\ell j} \right].$$

Notice that $c_{\ell j}$ is multiplied to a sum; we can distribute it and obtain

$$((AB)C)_{ij} = \sum_{\ell=1}^{s} \sum_{k=1}^{r} (a_{ik} b_{k\ell}) c_{\ell j}. \tag{3.1}$$

We turn to the other side of the equation. By definition, the jth element in row k of BC is

$$(BC)_{kj} = \sum_{\ell=1}^{s} b_{k\ell} c_{\ell j}.$$

Likewise, the jth element in row i of $A(BC)$ is

$$(A(BC))_{ij} = \sum_{k=1}^{r} \left(a_{ik} \sum_{\ell=1}^{s} b_{k\ell} c_{\ell j} \right).$$

This time, a_{ik} is multiplied to a sum; we can distribute it and obtain

$$(A(BC))_{ij} = \sum_{k=1}^{r} \sum_{\ell=1}^{s} a_{ik} \left(b_{k\ell} c_{\ell j} \right).$$

By the associative property of the entries,

$$(A(BC))_{ij} = \sum_{k=1}^{r} \sum_{\ell=1}^{s} (a_{ik} b_{k\ell}) c_{\ell j}. \tag{3.2}$$

The only difference between equations (3.1) and (3.2) is in the order of the summations: whether we add up the k's first or the ℓ's first. That is, the sums have the same terms, but those terms appear in different orders! We assumed the entries of the matrices were commutative under addition, so the order of the terms does not matter; we have

$$\left((AB)\,C\right)_{ij} = \left(A\,(BC)\right)_{ij}\ .$$

We chose arbitrary i and j, so this is true for all entries of the matrices. The matrices are equal, which means $(AB)\,C = A\,(BC)$. □

We now have enough information to classify two useful and important structures of *square* matrices. First, suppose the entries come from a general ring.

Theorem 3·96. *For any commutative ring R, the set $R^{n\times n}$ of $n\times n$ matrices over R is a noncommutative ring.*

Proof. We have shown that matrix addition satisfies most of the properties of an abelian group; the only one we have not shown is the commutative property of *addition*, which is easy to show.

Question 3·97 .
Why is matrix addition commutative?

Proof of Theorem 3·96 (continued). We have also shown that matrix multiplication satisfies the properties of a monoid; see Fact 3·88 and Theorems 3·90 and 3·95. So we need merely show that matrix multiplication distributes over addition. Let $n \in \mathbb{N}^+$ and $A, B, C \in R^{n\times n}$.

$$\begin{aligned}
\left[A\,(B+C)\right]_{ij} &= \sum_{k=1}^{n} \left[a_{ik}\left(b_{kj}+c_{kj}\right)\right] \\
&= \sum_{k=1}^{n} \left(a_{ik}b_{kj} + a_{ik}c_{kj}\right) \\
&= \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} \\
&= \left(AB\right)_{ij} + \left(AC\right)_{ij}.
\end{aligned}$$

This shows the elements in row i and column j are equal whenever we fix i and j between 1 and n. All the entries of $A\,(B+C)$ and $AB+AC$ are equal, so $A\,(B+C) = AB+AC$; the distributive property holds. □

Usually the multiplication does *not* commute.

Question 3·98 .
Look back at Question 3.79. Find two quaternion matrices A and B such that $AB \neq BA$.

Question 3·99 .
Suppose $n > 1$ and $R^{n\times n}$ is the set of all $n \times n$ matrices whose entries are elements of R. Find matrices A and B such that $AB \neq BA$.

Hint: Since the ring R is arbitrary, it has to work even when $R = \mathbb{Z}_2$, which limits your options in a way that is surprisingly useful. So, try finding two 2×2 matrices A and B whose entries are elements of $\mathbb{Z}_2$, and $AB \neq BA$. Once you find them, generalize your answer to any dimension $n \geqslant 2$.

So if the entries of our matrices merely come from a ring, the set of square matrices forms another ring, though most sets of matrices form a *noncommutative* ring. Nice!

Suppose we go further, using a field for our base ring. Except for the additive identity, multiplication in a field satisfies the inverse property. Will this be true of the matrices over that field? We've already seen this isn't true: Question 3.78 shows that zero divisor matrices exist over *every* ground ring R, which includes fields, and Question 2.45 tells us that fields cannot have zero divisors. So most rings of matrices will not be fields.

Can we build a field using *invertible* matrices? We need closure of multiplication.

Fact 3·100. *The product of two invertible matrices is also invertible.*

Question 3·101 .
Why is Fact 3·100 true? In other words, if A and B are invertible matrices, why is AB invertible?
Hint: Try to construct an inverse using the inverses of A and B.

We already know that matrix multiplication has an identity, which is invertible, and is associative. That's all we need; the set of invertible matrices forms a group!

Definition 3·102. Let $\mathbb{F}$ be a field. We call the set of invertible $n \times n$ matrices with elements in $\mathbb{F}$ the **general linear group over $\mathbb{F}$ of dimension** n, abbreviated $\mathrm{GL}_n(\mathbb{F})$. The operation is multiplication. Ordinarily we work with $\mathbb{F} = \mathbb{R}$ and fixed degree n, so when the meaning is clear and we're feeling somewhat lazy (which we usually are), we will refer simply to the **general linear group**.

Unfortunately, the set of invertible matrices *still* won't form a field, for several reasons.

Question 3·103 .
Find at least three properties of a field that $\mathrm{GL}_n(\mathbb{F})$ does not satisfy.

Question 3·104 .
Recall from Question 3.94 the orthogonal matrices $O(n)$.

(a) Show that if A and B are orthogonal matrices, then AB is also orthogonal.
Hint: You will need the additional matrix properties $(AB)^T = B^T A^T$.

(b) Show that $O(n)$ is a group under matrix multiplication.

We now return to the question we first posed above: why can't there be a different identity, either for addition or multiplication?

Fact 3·105. *The identity of a monoid is unique.*

Notice the claim: we don't say merely that the identity matrix is unique, whether that be the identity of addition or multiplication. We say that the identity of *any* monoid is unique. This covers matrices, whether under addition or multiplication, *and every other monoid possible.* We don't need even the full monoid structure! Pay attention to the explanation, and see if you can identify which properties aren't required.

Why? Let M be a monoid, and $я \in M$ an identity. Suppose that $e \in M$ is also an identity; perhaps $я = e$, but perhaps $я \neq e$; we are not sure. (Merely having a different name does not imply a different substance.) By the fact that $я$ is an identity, we know that $яe = e$. On the other hand, the fact that e is an identity tells us that $яe = я$. By substitution, $я = e$. Since our choice of identities was arbitrary, and they turned out equal, it must be that the identity of a monoid is unique. □

Question 3·106 .
Which property (-ies) of a monoid did we not use in the explanation above?

Question 3·107 .
Suppose G is a group, and $x \in G$. We know that x has an inverse; call it $y \in z$. Can x have another inverse, $z \in G$? *Hint:* As in the explanation for Fact 3·105, it helps to show that $y = z$, but the trick is a little different.

Question 3·108 .
Use a fact from linear algebra to explain why $\mathrm{GL}_m(\mathbb{R})$ is not cyclic.

Example 3·109. Let

$$G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \subsetneq \mathrm{GL}_m(\mathbb{R}).$$

It turns out that G is a group; both the second and third matrices generate it. For example,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Question 3·110 .
For the matrices in Example 3·109, let

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Express A as a power of the other non-identity matrices of the group.

Example 3·111. Recall Example 3·109; we can write

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^4 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^8 = \cdots.$$

Since multiples of 4 give the identity, let's take any power of the matrix, and divide it by 4. The Division Theorem allows us to write any power of the matrix as $4q + r$, where $0 \leqslant r < 4$. Since there are only four possible remainders, and multiples of 4 give the identity, positive powers of this matrix can generate only four possible matrices:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{4q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{4q+1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{4q+2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{4q+3} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We can do the same with negative powers; the Division Theorem still gives us only four possible remainders. Let's write

$$g = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Thus

$$\langle g \rangle = \left\{ I_2, g, g^2, g^3 \right\}.$$

3·6 Symmetry in polygons

> *What is it indeed that gives us the feeling of elegance in a solution, in a demonstration? It is the harmony of the diverse parts, their symmetry, their happy balance; in a word it is all that introduces order, all that gives unity, that permits us to see clearly and to comprehend at once both the ensemble and the details.*
>
> — Henri Poincaré

(a) (b)

Figure 3·3: Rotation and reflection of the triangle

A *geometric* phenomenon with mathematical structure is called "the symmetries of a regular polygon." This mouthful of words requires some explanation. For the sake of simplicity, we stick with a triangle, but the basic ideas here work with any number of sides, and we touch on this briefly at the end of the section.

In general, the set of symmetries of a regular polygon with n sides is called D_n, so we will be looking at D_3, but you should pause from time to time and think of D_4 or D_5, because you're going to face them sooner or later, too.

Intuitive development of D_3

To describe D_3, start with an equilateral triangle in $\mathbb{R}^2$, with its center at the origin. A "symmetry" is a transformation of the plane that leaves the *triangle* in the same location, even if its *points* are in different locations. For instance, if you rotate the triangle 120° over its center, the triangle ends up in the same location, even thought all the points have moved; this is not true if you rotate by 30° or 60°. Likewise, if you reflect the triangle about the y-axis, the triangle ends up in the same location, even though most of the points have moved. We'll call that rotation ρ, and that reflection φ. See Figure 3·6.

"Transformations" include actions like rotation, reflection (flip), and translation (shift). Translating the plane in some direction certainly won't leave the triangle intact, but rotation and reflection can.

It is helpful to observe two important properties.

Theorem 3·112. *If φ and ρ are as specified, then $\varphi\rho = \rho^2\varphi$.*

For now, we consider intuitive proofs only. Detailed proofs appear later in the section. It'll help if you sketch the arguments.

Intuitive proof. The expression $\varphi\rho$ means to apply ρ first, then φ; after all, these are functions, so $(\varphi\rho)(x) = \varphi(\rho(x))$. Rotating 120° moves vertex 1 to vertex 2, vertex 2 to vertex 3, and vertex 3 to vertex 1. Flipping through the y-axis leaves the top vertex in place; since we performed the rotation first, the top vertex is now vertex 3, so vertices 1 and 2 are the ones swapped. Thus, vertex 1 has moved to vertex 3, vertex 3 has moved to vertex 1, and vertex 2 is in its original location.

On the other hand, $\rho^2\varphi$ means to apply φ first, then apply ρ twice. Again, it will help to sketch what follows. Flipping through the y-axis swaps vertices 2 and 3, leaving vertex 1 in the same place. Rotating twice then moves vertex 1 to the lower right position, vertex 3 to the top position, and vertex 2 to the lower left position. This is the same arrangement of the vertices as we had for $\varphi\rho$, which means that $\varphi\rho = \rho^2\varphi$. □

You might notice a gap in the reasoning: we showed that each *vertex* of the triangle moved to a position that previously held a *vertex*, but said nothing of the *points in between*. That requires a little more work, which is why we provide detailed proofs later.

By the way, did you notice what Theorem 3·112 did *not* claim?

Question 3·113 .
Show that D_3 is non-commutative: $\varphi\rho \neq \rho\varphi$.

Another "obvious" symmetry of the triangle is the transformation where you *do nothing* – or, if you prefer, where you effectively *move every point back to itself,* as in a 360° rotation. We'll call this symmetry ι. It gives us the last property we need to specify the group, D_3.

Question 3·114 .
Compute the cyclic group generated by a in D_3.

(a) $a = \varphi$

(b) $a = \rho^2$

(c) $a = \rho\varphi$

Theorem 3·115. *In D_3, $\rho^3 = \varphi^2 = \iota$.*

Intuitive proof. Rotating 120° three times is the same as rotating 360°, which leaves points in the same position as if they had not rotated at all. Likewise, φ moves any point (x, y) to $(x, -y)$, and applying φ again moves $(x, -y)$ back to (x, y), which is the same as not flipping at all. □

We are now ready to specify D_3.

Theorem 3·116. *The set of symmetries of a regular triangle, $D_3 = \{\iota, \varphi, \rho, \rho^2, \rho\varphi, \rho^2\varphi\}$, is a group under composition of functions.*

We can prove most of these by mere inspection of the Cayley table, will you will compute in Question 3.117. However, we can also give geometric reasoning. As long as that isn't too complicated, we add a geometric argument, as well.

Proof. We prove this by showing that all the properties of a group are satisfied. We only start the proof, leaving it you to finish in Question 3.117.

Closure: In Question 3.117, you will compute the Cayley table of D_3. There, you will see that every composition is also an element of D_3.

Associative: In Section 7·1, we show that composition of functions is associative. Symmetries are functions that map any point in $\mathbb{R}^2$ to another point in $\mathbb{R}^2$, with no ambiguity about

where the point goes. Proving the associative property once for an *arbitrary* function over an *arbitrary* set takes care of particular functions (D_3) over a particular set ($\mathbb{R}^2$).

Identity: In Question 3.117, you will compute the Cayley table of D_3. There, you will find that $\iota\sigma = \sigma\iota = \sigma$ for every $\sigma \in D_3$.

(Alternately, let $\sigma \in D_3$ be any symmetry. Apply σ to the triangle. Then apply ι. Since ι leaves everything in place, all the points are in the same place they were after we applied σ. In other words, $\iota\sigma = \sigma$. The proof that $\sigma\iota = \sigma$ is similar.)

Inverse: In Question 3.117, you will compute the Cayley table of D_3. There, you will find that for every $\sigma \in D_3$, the row labeled σ contains ι in exactly one column. The element at the top of that row is σ^{-1} by definition.

(Alternately, it is clear that rotation and reflection are one-to-one-functions; after all, if a point P is mapped to a point R by either, it doesn't make sense that another point Q would also be mapped to R. Since one-to-one functions have inverses, every element σ of D_3 must have an inverse function σ^{-1}, which undoes whatever σ did. But is $\sigma^{-1} \in D_3$ — that is, is σ^{-1} a *symmetry?* Since σ maps every point of the triangle onto the triangle, σ^{-1} will undo that map: every point of the triangle will be mapped back onto another point of the triangle, as well. So, yes, $\sigma^{-1} \in D_3$.) □

Question 3·117 .

The multiplication table for D_3 has at least this structure:

$\circ$	ι	φ	ρ	ρ^2	$\rho\varphi$	$\rho^2\varphi$
ι	ι	φ	ρ	ρ^2	$\rho\varphi$	$\rho^2\varphi$
φ	φ		$\rho^2\varphi$			
ρ	ρ	$\rho\varphi$				
ρ^2	ρ^2					
$\rho\varphi$	$\rho\varphi$					
$\rho^2\varphi$	$\rho^2\varphi$					

Complete the multiplication table, writing every element in the form $\rho^m\varphi^n$, never with φ before ρ. *Do not use matrix multiplication;* instead, use Theorems 3·112 and 3·115.

Question 3·118 .

The set D_4 of symmetries of a square is also a group, though it has 8 elements. It, too, can be built using only a rotation and a reflection. Choose such a rotation and reflection that allow you to list all 8 elements as products of them, in a manner similar to what we did with D_3. Identify properties of its elements that resemble the properties found for the rotation and reflection of D_3, and use them to build a Cayley table for D_4

Question 3·119 .

Find a geometric figure (not a polygon) that is preserved by at least one rotation, at least one reflection, *and* at least one translation. Keep in mind that, when we say "preserved", we mean that the points of the figure end up on the figure itself — just as a 120° rotation leaves the triangle on itself.

Detailed proof that D_3 contains all symmetries of the triangle

To prove that D_3 contains *all* symmetries of the triangle, we need to make some notions more precise. First, what is a symmetry? A **symmetry** of *any* polygon is a distance-preserving function on $\mathbb{R}^2$ that maps points of the polygon back onto itself. Notice the careful wording: the *points* of the polygon can change places, but since they have to be mapped back onto the polygon, the polygon itself has to remain in the same place.

Let's look at the specifics for our triangle. What functions are symmetries of the triangle? To answer this question, we divide it into two parts.

1. What are the distance-preserving functions that map $\mathbb{R}^2$ to itself, and leave the origin undisturbed? Here, distance is measured by the usual metric,

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

   (You might wonder why we don't want the origin to move. Basically, if a function α preserves both distances between points and a figure centered at the origin, then the origin *cannot* move, since its distance to points on the figure would change.)

2. Not all of the functions identified by question (1) map points on the triangle back onto the triangle; for example, a 45° degree rotation does not. Which ones do?

Lemma 3·120 answers the first question.

Lemma 3·120. *Let $\alpha : \mathbb{R}^2 \to \mathbb{R}^2$. If*

- *α does not move the origin; that is, $\alpha(0,0) = (0,0)$, and*
- *the distance between $\alpha(P)$ and $\alpha(R)$ is the same as the distance between P and R for every $P, R \in \mathbb{R}^2$,*

then α has one of the following two forms:

$$\rho = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \quad \exists t \in \mathbb{R}$$

or

$$\varphi = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix} \quad \exists t \in \mathbb{R}.$$

The two values of t may be different.

Proof. Assume that $\alpha(0,0) = (0,0)$ and for every $P, R \in \mathbb{R}^2$ the distance between $\alpha(P)$ and $\alpha(R)$ is the same as the distance between P and R. We can determine α precisely merely from how it moves two points in the plane! We'll choose two "easy" points to manipulate.

Consider $P = (1,0)$ as the first point. Let $Q = \alpha(P)$; that is, Q is P's destination when α moves it. Write $Q = (q_1, q_2)$. The distance between P and the origin is 1. By hypothesis α, does not move the origin, so the distance between Q and the origin will also be 1. In other words,

$$1 = \sqrt{q_1^2 + q_2^2},$$

or

$$q_1^2 + q_2^2 = 1.$$

The only values for Q that satisfy this equation are those points that lie on the circle whose center is the origin. We can describe any point on this circle as

$$(\cos t, \sin t)$$

where $t \in [0, 2\pi)$ represents an angle. Hence, $\alpha(P) = (\cos t, \sin t)$.

Consider $R = (0, 1)$ as the second point. Let $S = \alpha(R)$; that is, S is R's destination when α moves it. Write $S = (s_1, s_2)$. An argument similar to the one above shows that S also lies on the circle whose center is the origin. Moreover, the distance between P and R is $\sqrt{2}$, so the distance between Q and S is also $\sqrt{2}$. That is,

$$\sqrt{(\cos t - s_1)^2 + (\sin t - s_2)^2} = \sqrt{2},$$

or

$$(\cos t - s_1)^2 + (\sin t - s_2)^2 = 2. \tag{3.3}$$

Recall that $\cos^2 t + \sin^2 t = 1$. That means we can rewrite (3.3) as

$$-2(s_1 \cos t + s_2 \sin t) + \left(s_1^2 + s_2^2\right) = 1. \tag{3.4}$$

To solve this, recall that the distance from S to the origin must be the same as the distance from R to the origin, which is 1. Hence

$$\begin{aligned} \sqrt{s_1^2 + s_2^2} &= 1 \\ s_1^2 + s_2^2 &= 1. \end{aligned}$$

Substituting this into (3.4), we find that

$$\begin{aligned} -2(s_1 \cos t + s_2 \sin t) + s_1^2 + s_2^2 &= 1 \\ -2(s_1 \cos t + s_2 \sin t) + 1 &= 1 \\ -2(s_1 \cos t + s_2 \sin t) &= 0 \\ s_1 \cos t &= -s_2 \sin t. \end{aligned} \tag{3.5}$$

You can guess two solutions to this equation: $S = (\sin t, -\cos t)$ and $S = (-\sin t, \cos t)$ is another. Are there more?

Recall that $s_1^2 + s_2^2 = 1$, so $s_2 = \pm\sqrt{1 - s_1^2}$. Likewise $\sin t = \pm\sqrt{1 - \cos^2 t}$. Substituting into equation (3.5) and squaring (so as to remove the radicals), we find that

$$\begin{aligned} s_1 \cos t &= -\sqrt{1 - s_1^2} \cdot \sqrt{1 - \cos^2 t} \\ s_1^2 \cos^2 t &= \left(1 - s_1^2\right)\left(1 - \cos^2 t\right) \\ s_1^2 \cos^2 t &= 1 - \cos^2 t - s_1^2 + s_1^2 \cos^2 t \\ s_1^2 &= 1 - \cos^2 t \\ s_1^2 &= \sin^2 t \\ \therefore s_1 &= \pm \sin t. \end{aligned}$$

Along with equation (3.5), this implies that $s_2 = \mp \cos t$. We already found theses solutions, so we're done.

It can be shown (see Question 3.124) that α satisfies a property called "linear transformation"; that is, for all $P, Q \in \mathbb{R}^2$ and for all $a, b \in \mathbb{R}$, $\alpha(aP + bQ) = a\alpha(P) + b\alpha(Q)$.. Linear algebra tells us that we can describe any linear transformation over a finite-dimensional vector space as a matrix. If $S = (\sin t, -\cos t)$ then

$$\alpha = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix};$$

otherwise

$$\alpha = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

The lemma names the first of these forms φ and the second ρ. □

How do these matrices affect points in the plane?

Example 3·121. Consider the set of points

$$\mathcal{S} = \{(0, 2), (\pm 2, 1), (\pm 1, -2)\};$$

these form the vertices of a (non-regular) pentagon in the plane. Let $t = \pi/4$; then

$$\rho = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \quad \text{and} \quad \varphi = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

If we apply ρ to every point in the plane, then the points of $\mathcal{S}$ move to

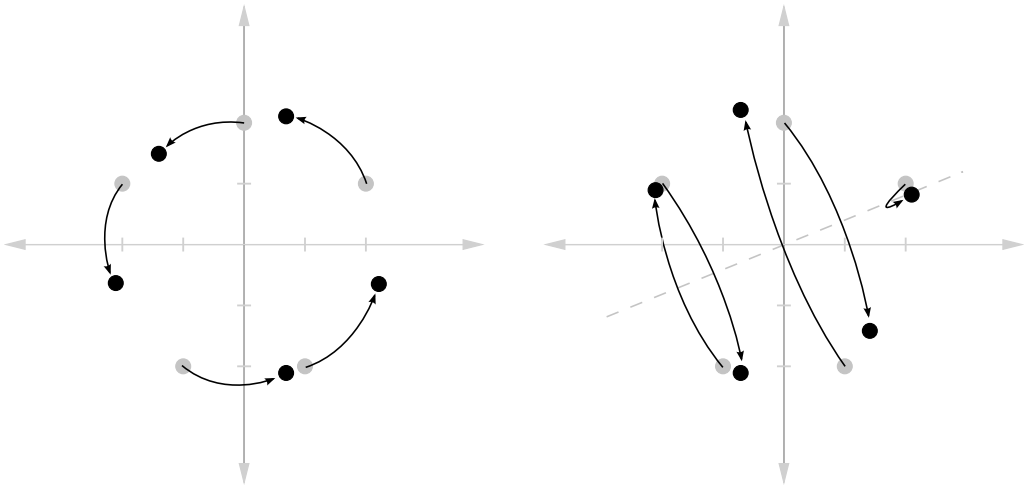
$$\begin{aligned} \rho(\mathcal{S}) &= \{\rho(0, 2), \rho(-2, 1), \rho(2, 1), \rho(-1, -2), \rho(1, -2)\} \\ &= \left\{ \left(-\sqrt{2}, \sqrt{2}\right), \left(-\sqrt{2} - \frac{\sqrt{2}}{2}, -\sqrt{2} + \frac{\sqrt{2}}{2}\right), \right. \\ &\qquad \left(\sqrt{2} - \frac{\sqrt{2}}{2}, \sqrt{2} + \frac{\sqrt{2}}{2}\right), \\ &\qquad \left(-\frac{\sqrt{2}}{2} + \sqrt{2}, -\frac{\sqrt{2}}{2} - \sqrt{2}\right), \\ &\qquad \left. \left(\frac{\sqrt{2}}{2} + \sqrt{2}, \frac{\sqrt{2}}{2} - \sqrt{2}\right) \right\} \\ &\approx \{(-1.4, 1.4), (-2.1, -0.7), (0.7, 2.1), \\ &\qquad (0.7, -2.1), (2.1, -0.7)\}. \end{aligned}$$

This is a 45° ($\pi/4$) counterclockwise rotation in the plane.

If we apply φ to every point in the plane, then the points of $\mathcal{S}$ move to

$$\begin{aligned} \varphi(\mathcal{S}) &= \{\varphi(0, 2), \varphi(-2, 1), \varphi(2, 1), \varphi(-1, -2), \varphi(1, -2)\} \\ &\approx \{(1.4, -1.4), (-0.7, -2.1), (2.1, 0.7), \\ &\qquad (-2.1, 0.7), (-0.7, 2.1)\}. \end{aligned}$$

This is shown in Figure 3·121 . The line of reflection for φ has slope $\left(1 - \cos\frac{\pi}{4}\right) / \sin\frac{\pi}{4}$. (You will show this in Question 3.126.)



ρ φ

Figure 3·4: Actions of ρ and φ on a pentagon, with $t = \pi/4$

The second questions asks which of the matrices described by Lemma 3·120 also preserve the triangle.

- The first solution (ρ) corresponds to a rotation of degree t of the plane. To preserve the triangle, we can only have $t = 0,\ 2\pi/3,\ 4\pi/3$ (0°, 120°, 240°). (See Figure 3·6(a).) Let ι correspond to $t = 0$, the identity rotation, as that gives us

$$\iota = \begin{pmatrix} \cos 0 & -\sin 0 \\ \sin 0 & \cos 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

  which is what we would expect for the identity. Let ρ correspond to a counterclockwise rotation of 120°, or

$$\rho = \begin{pmatrix} \cos \frac{2\pi}{3} & -\sin \frac{2\pi}{3} \\ \sin \frac{2\pi}{3} & \cos \frac{2\pi}{3} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

  A rotation of 240° is the same as rotating 120° twice. We can write that as $\rho \circ \rho$ or ρ^2; matrix multiplication gives us

$$\begin{aligned} \rho^2 &= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

- The second solution (φ) corresponds to a flip along the line whose slope is

$$m = (1 - \cos t) \,/\, \sin t.$$

One way to do this would be to flip across the y-axis (see Figure 3·6(b)). For this we need the slope to be undefined, so the denominator needs to be zero and the numerator needs to be non-zero. One possibility is $t = \pi$. So

$$\varphi = \begin{pmatrix} \cos\pi & \sin\pi \\ \sin\pi & -\cos\pi \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

There are two other flips, but we can actually ignore them, because they are combinations of φ and ρ. (Why? See Question 3.123.)

We can now give more detailed proofs of Theorems 3·112 and 3·115. We'll prove the first here, and you'll prove the second in a moment.

Detailed proof of Theorem 3·112. Compare

$$\varphi\rho = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

and

$$\begin{aligned} \rho^2\varphi &= \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

□

Question 3·122 .

Show explicitly (by matrix multiplication) that $\rho^3 = \varphi^2 = \iota$.

Question 3·123 .

Two other values of t allow us to define flips for the triangle. Find these values of t, and explain why their matrices are equivalent to the matrices $\rho\varphi$ and $\rho^2\varphi$.

Question 3·124 .
Show that any function α satisfying the requirements of Theorem 3·120 is a linear transformation; that is, for all $P, Q \in \mathbb{R}^2$ and for all $a, b \in \mathbb{R}$, $\alpha(aP + bQ) = a\alpha(P) + b\alpha(Q)$. Use the following steps.

(a) Prove that $\alpha(P) \cdot \alpha(Q) = P \cdot Q$, where $\cdot$ denotes the usual dot product (or inner product) on $\mathbb{R}^2$.

(b) Show that $\alpha(1,0) \cdot \alpha(0,1) = 0$.

(c) Show that $\alpha((a,0) + (0,b)) = a\alpha(1,0) + b\alpha(0,1)$.

(d) Show that $\alpha(aP) = a\alpha(P)$.

(e) Show that $\alpha(P + Q) = \alpha(P) + \alpha(Q)$.

Question 3·125 .
Show that the only stationary point in $\mathbb{R}^2$ for the general ρ is the origin. That is, if $\rho(P) = P$, then $P = (0,0)$. (By "general", we mean any ρ, not just the one in D_3.)

Question 3·126 .
Fill in each blank of Figure 3·6 with the appropriate justification.

Question 3·127 .
Let

$$\varphi = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad \Phi = \left\{\varphi, \varphi^2\right\}.$$

(a) Simplify φ^2.

(b) Is Φ a monoid under multiplication? if so, is it commutative?

(c) Is Φ a monoid under addition? if so, is it commutative?

(d) Is Φ a group under addition? if so, is it abelian?

(e) Is Φ a group under multiplication? if so, is it abelian?

(f) Show that Φ has the form

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

by identifying the value of α.

(g) Explain why a matrix φ endowed with the form described in part (f) can serve as the "basis" for a set $\{\varphi, \varphi^2\}$ that satisfies or fails the structures you determined in parts (a)–(e). *Hint:* You should be able to do this using induction and properties of the 'trigonometric functions involved.

Claim: The only stationary points of φ lie along the line whose slope is $(1-\cos t)/\sin t$, where $t \in [0, 2\pi)$ and $t \neq 0, \pi$. If $t = 0$, only the x-axis is stationary, and for $t = \pi$, only the y-axis.
Proof:

1. Let $P \in \mathbb{R}^2$. By ____, there exist $x, y \in \mathbb{R}$ such that $P = (x, y)$.
2. Assume φ leaves P stationary. By ____,
$$\begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$
3. By linear algebra,
$$\begin{pmatrix} \_\_\_\_ \\ \_\_\_\_ \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$
4. By the principle of linear independence, ____$= x$ and ____$= y$.
5. For each equation, collect x on the left hand side, and y on the right, to obtain
$$\begin{cases} x(\_\_\_\_) = -y(\_\_\_\_) \\ x(\_\_\_\_) = \ \ y(\_\_\_\_) \end{cases}.$$
6. If we solve the first equation for y, we find that $y =$____.
   (a) This, of course, requires us to assume that ____$\neq 0$.
   (b) If that *was* in fact zero, then $t =$____, ____ (remembering that $t \in [0, 2\pi)$).
7. Put these values of t aside. If we solve the second equation for y, we find that $y =$____.
   (a) Again, this requires us to assume that ____$\neq 0$.
   (b) If that *was* in fact zero, then $t =$____. We already put this value aside, so ignore it.
8. Let's look at what happens when $t \neq$____ and ____.
   (a) Multiply numerator and denominator of the right hand side of the first solution by the denominator of the second to obtain $y =$____.
   (b) Multiply right hand side of the second with denominator of the first: $y =$____.
   (c) By ____, $\sin^2 t = 1 - \cos^2 t$. Substitution into the second solution gives the first!
   (d) That is, points that lie along the line $y =$____ are left stationary by φ.
9. Now consider the values of t we excluded.
   (a) If $t =$____, then the matrix simplifies to $\varphi =$____.
   (b) To satisfy $\varphi(P) = P$, we must have ____$= 0$, and ____ free. The points that satisfy this are precisely the ____-axis.
   (c) If $t =$____, then the matrix simplifies to $\varphi =$____.
   (d) To satisfy $\varphi(P) = P$, we must have ____$= 0$, and ____ free. The points that satisfy this are precisely the ____-axis.

Figure 3·5: Material for Question 3.126

Question 3·128 .

Let

$$\varrho = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \text{and} \quad P = \left\{\varrho, \varrho^2, \varrho^3\right\}.$$

If you've not seen the symbol that looks like a backwards g, we call it "rho". How does it differ from ρ? It's fancier. (There's no other difference.) Likewise, the symbol that looks like a capital P is actually a capital rho.

(a) Simplify ϱ^2 and ϱ^3.

(b) Is P a monoid under multiplication? if so, is it commutative?

(c) Is P a monoid under addition? if so, is it commutative?

(d) Is P a group under addition? if so, is it abelian?

(e) Is P a group under multiplication? if so, is it abelian?

(f) Show that P has the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

by identifying the value of α.

(g) Explain why a matrix ϱ with the form described in part (f), *and* the condition $\alpha = \pi/n$, can serve as the "basis" for a set $\{\varrho, \varrho^2, \ldots, \varrho^n\}$ that satisfies or fails the structures you determined in parts (a)–(d). *Hint:* First show that for any k ϱ^k has almost the same form as ϱ, but with $a = k\pi/n$. You should be able to do this using induction and properties of the trigonometric functions involved.

Chapter 4

Subgroups and Ideals, Cosets and Quotients

A subset of a group is not necessarily a group; for example, $\{2, 4\} \subsetneq \mathbb{Z}$, but $\{2, 4\}$ doesn't satisfy the same group properties as $\mathbb{Z}$ unless we change the operation. On the other hand, if we do change the operation, it doesn't make sense to call $\{2, 4\}$ a subgroup of $\mathbb{Z}$, because the group property depends not only on the elements, but on the operation, as well.

Some subsets of groups *are* groups, and one key to algebra consists in understanding the relationship between subgroups and groups. We start this chapter by describing the properties that guarantee that a subset is a "subgroup" of a group (Section 4·1). In a ring, we are more interested in a special sort of subgroup called *an ideal*. Ideals are related to roots of polynomial equations (Section 4·2) and generalize a number of ideas you have seen, including the bases of vector spaces (Section 4·3). We then explore how equivalence relations and classes related to $\mathbb{Z}_d$ (Section 4·5) lead to a more general relationship between subgroups and ideals, which generalizes the idea of division and modular arithmetic via *cosets* (Section 4·6). In finite groups and rings, we can count the number of cosets quite easily (Section 4·7). Cosets open the door to a special class of groups called *quotient groups*, (Sections 4·8) which form an important foundation of the second half of these notes.

4·1 Subgroups

Definition 4·1. Let G be a group and $H \subseteq G$ be nonempty. If H is also a group under the same operation as G, then H is a **subgroup** of G. We call H a **proper subgroup** if $\{я\} \subsetneq H \subsetneq G$.

Notation 4·2. If H is a subgroup of G, then we write $H < G$.

Question 4·3 .
Verify the following statements by checking that the properties of a group are satisfied.

(a) $\mathbb{Z}$ is a subgroup of $\mathbb{Q}$.

(b) Let $4\mathbb{Z} := \{4m : \ m \in \mathbb{Z}\} = \{\ldots, -4, 0, 4, 8, \ldots\}$. Then $4\mathbb{Z}$ is a subgroup of $\mathbb{Z}$.

(c) Let $d \in \mathbb{Z}$ and $d\mathbb{Z} := \{dm : \ m \in \mathbb{Z}\}$. Then $d\mathbb{Z}$ is a subgroup of $\mathbb{Z}$.

(d) The set of multiples of the quaternion $\mathbf{i}$ is a subgroup of Q_8.

Checking all four properties of a group is cumbersome. It would be convenient to verify that a set is a subgroup by checking fewer properties. Which properties can we skip when checking whether a subset is a subgroup?

Intuitively, we can skip a property if it is "inheritable." For instance, if the operation is commutative on a set, then it remains commutative any *sub*set; after all, the elements of the subset are elements of the original set.

Lemma 4·4. *Let G be a group and $H \subseteq G$. Then H satisfies the associative property of a group. In addition, if G is abelian, then H satisfies the commutative property of an abelian group. So, we only need to check the closure, identity, and inverse properties to ensure that G is a group.*

Be careful: Lemma 4·4 neither assumes nor concludes that H is a subgroup. The other three properties may not be satisfied: H may not be closed; it may lack an identity; or some element may lack an inverse. The lemma merely states that any subset automatically satisfies two important properties of a group.

Proof. If $H = \varnothing$, then the lemma is true trivially.

Otherwise, $H \neq \varnothing$. Let $a, b, c \in H$. Since $H \subseteq G$, we have $a, b, c \in G$. Since the operation is associative in G, $a\,(bc) = (ab)\,c$; that is, the operation remains associative for H. Likewise, if G is abelian, then $ab = ba$; that is, the operation also remains commutative for H. □

Lemma 4·4 has reduced the number of requirements for a subgroup from four to three. Amazingly, we can simplify this further, to *one criterion alone!*

The Subgroup Theorem. *Let $A \subseteq G$ be nonempty. The following are equivalent:*

(A) $A < G$;

(B) for every $a, b \in A$, we have $a^{-1}b \in A$.

(C) for every $a, b \in A$, we have $ab^{-1} \in A$.

Notation 4·5. If the operation governing G were addition, we would write $-a+b$ or $a-b$ instead of $a^{-1}b$ or ab^{-1}.

Characterization (C) of the Subgroup Theorem gives us a nice, intuitive guideline: "A nonempty subset is a subgroup iff it closed under division (or subtraction)." We will typically go by this characterization.

Proof. Assume (A). Let $a, b \in A$. By the inverse property, $a^{-1} \in A$; by closure, $a^{-1}b \in A$. We chose a and b arbitrarily, so this holds for all $a, b \in A$.

Conversely, assume (B). By Lemma 4·4, we need to show only that A satisfies the closure, identity, and inverse properties. We do this slightly out of order:

identity: Let $a \in A$. By (B), $я = a^{-1}a \in A$.[1]

inverse: Let $a \in A$. We just showed A satisfies the identity property, so $я \in A$. By (B), $a^{-1} = a^{-1} \cdot я \in A$.

closure: Let $a, b \in A$. We just showed A satisfies the inverse property, so $a^{-1} \in H$. By (B), $ab = \left(a^{-1}\right)^{-1} b \in A$.

Since H satisfies the closure, identity, and inverse properties, $A < B$.

We have show that (A) is equivalent to (B). We leave the proof that (A) is equivalent to (C) □

Question 4·6 .

Show that item (C) of the Subgroup Theorem is equivalent to item (A): that is, $A < G$ if and only if A is closed under division (or subtraction).

Let's take a look at the Subgroup Theorem in action.

Example 4·7. Let $d \in \mathbb{Z}$. We claim that $d\mathbb{Z} < \mathbb{Z}$. (Recall that $d\mathbb{Z}$, defined in Example 4.3, is the set of integer multiples of d.) *Why?* Let's use the Subgroup Theorem.

Let $x, y \in d\mathbb{Z}$. If we can show that $x - y \in d\mathbb{Z}$, or in other words, $x - y$ is an integer multiple of d, then we will satisfy part (B) of the Subgroup Theorem. The theorem states that (B) is equivalent to (A); that is, $d\mathbb{Z}$ is a group.

Since x and y are by definition integer multiples of d, we can write $x = dm$ and $y = dn$ for some $m, n \in \mathbb{Z}$. Note that $-y = -(dn) = d(-n)$. Then

$$\begin{aligned} x - y = x + (-y) &= dm + d(-n) \\ &= d(m + (-n)) = d(m - n). \end{aligned}$$

Now, $m - n \in \mathbb{Z}$, so $x - y = d(m - n) \in d\mathbb{Z}$.

We did it! We took two integer multiples of d, and showed that their difference is also an integer multiple of d. By the Subgroup Theorem, $d\mathbb{Z} < \mathbb{Z}$.

Example 4·7 gives us an example of how the Subgroup Theorem verifies subgroups of *abelian* groups. Two interesting examples of subgroups of a nonabelian group appear in D_3.

Example 4·8. Recall D_3 from Section 3·6. Both $H = \{\iota, \varphi\}$ and $K = \{\iota, \rho, \rho^2\}$ are subgroups of D_3. *Why?* Certainly $H, K \subsetneq G$, and Theorem 3·49 on page 85 tells us that H and K are groups, since $H = \langle\varphi\rangle$, and $K = \langle\rho\rangle$.

[1] Notice that here we are replacing the b in (B) with a. This is fine, since nothing in (B) requires a and b to be distinct.

Let G be a group and $A_1, A_2, \ldots, A_m$ subgroups of G. Let

$$B = A_1 \cap A_2 \cap \cdots \cap A_m.$$

Claim: $B < G$.
Proof:

1. Let $x, y \in$_____.
2. By _____, $x, y \in A_i$ for all $i = 1, \ldots, m$.
3. By _____, $xy^{-1} \in A_i$ for all $i = 1, \ldots, m$.
4. By _____, $xy^{-1} \in B$.
5. By _____, $B < G$.

Figure 4·1: Material for Question 4.10

Sometimes we can build new subgroups from old ones. The following questions consider these possibilities.

Question 4·9 .
Will the union of two subgroups form a subgroup? Not usually. Find a group G and subgroups H, K of G such that $A = H \cup K$ *is not* a subgroup of G.

Question 4·10 .
Will the intersection of two subgroups form a subgroup? Yes! To see why, fill each blank of Figure 4·1 with the appropriate justification or expression.

Question 4·11 .
Will a subset formed by applying the operation to elements of two subgroups form a subgroup? We consider two cases.

(a) If G is an *abelian* group and H, K are subgroups of G, let

$$H + K = \{x + y : \ x \in H, y \in K\}.$$

Show that $H + K < G$.

(b) If G is a *nonabelian* group and H, K are subgroups of G, let

$$HK = \{xy : x \in H, y \in K\}.$$

Find G, H, K such that HK is *not* a subgroup of G.

Question 4·12 .
Let $H = \{\iota, \varphi\} < D_3$.

(a) Find a different subgroup K of D_3 with only two elements.

(b) Let $HK = \{xy : x \in H, y \in K\}$. Confirm that $HK \not< D_3$.

The following geometric example gives a visual image of what a subgroup "looks" like.

Example 4·13. Recall that $\mathbb{R}$ is a group under addition, and let G be the direct product $\mathbb{R} \times \mathbb{R}$. Geometrically, this is the set of points in the x-y plane. As is usual with a direct product, we define an addition for elements of G in the natural way: for $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, define

$$P_1 + P_2 = (x_1 + x_2, y_1 + y_2)\,.$$

Let H be the x-axis; a set definition would be, $H = \{x \in G : x = (a, 0)\ \exists a \in \mathbb{R}\}$. We claim that $H < G$. *Why?* Use the Subgroup Theorem! Let $P, Q \in H$. By the definition of H, we can write $P = (p, 0)$ and $Q = (q, 0)$ where $p, q \in \mathbb{R}$. Then

$$P - Q = P + (-Q) = (p, 0) + (-q, 0) = (p - q, 0)\,.$$

Membership in H requires the first ordinate to be real, and the second to be zero. As $P - Q$ satisfies these requirements, $P - Q \in H$. The Subgroup Theorem implies that $H < G$.

Let K be the line $y = 1$; a set definition would be, $K = \{x \in G : x = (a, 1)\ \exists a \in \mathbb{R}\}$. We claim that $K \not< G$. *Why not?* Again, use the Subgroup Theorem! Let $P, Q \in K$. By the definition of K, we can write $P = (p, 1)$ and $Q = (q, 1)$ where $p, q \in \mathbb{R}$. Then

$$P - Q = P + (-Q) = (p, 1) + (-q, -1) = (p - q, 0)\,.$$

Membership in K requires the second ordinate to be one, but the second ordinate of $P - Q$ is zero, not one. Since $P - Q \notin K$, the Subgroup Theorem tells us that K is not a subgroup of G.

There's a more direct explanation as to why K is not a subgroup; it doesn't contain the origin. In a direct product of groups, the identity is formed using the identities of the component groups. In this case, the identity is $(0, 0)$, which is *not* in K.

Figure 4·13 gives a visualization of H and K. You will diagram another subgroup of G in Question 4.14.

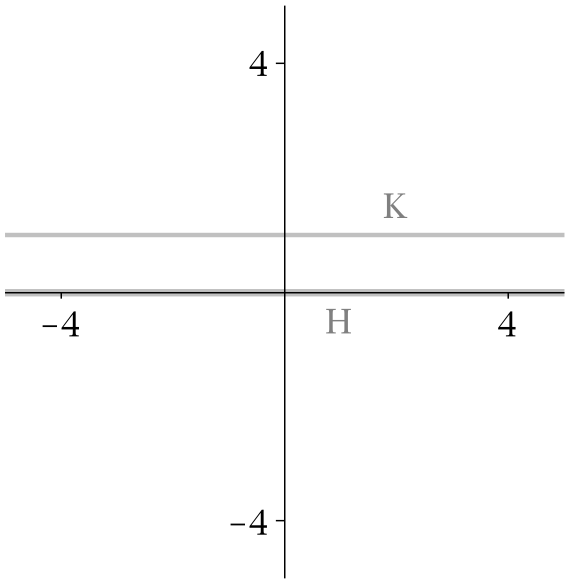


Figure 4·2: H and K from Example 4·13

Question 4·14 .

Let $G = \mathbb{R}^2$, with addition defined as in Example 4·13. Let

$$L = \{x \in G : \ x = (a, a) \ \ \exists a \in \mathbb{R}\}.$$

(a) Describe L geometrically.

(b) Show that $L < G$.

(c) Suppose $\ell \subseteq G$ is any line. Identify the simplest criterion possible that decides whether $\ell < G$. Justify your answer.

(d) Show that any subgroup ℓ you identify in part (c), which includes our original L, is isomorphic to $\mathbb{R}$ as an additive group.
Hint: Use a isomorphism f that maps $\mathbb{R}$ to ℓ, then the symmetry of isomorphism (Question 4.74 on page 144).

Aside from the basic group properties, what other properties can a subgroup inherit from a group? The answer is not always obvious. Cyclic groups are a good example: is every subgroup of a cyclic group also cyclic? The answer relies on the Division Theorem.

Theorem 4·15. *Subgroups of cyclic groups are also cyclic.*

Proof. Let G be a cyclic group, and $H < G$. From the fact that G is cyclic, choose $g \in G$ such that $G = \langle g \rangle$.

First we must find a candidate generator of H. Every element of H is an element of G, and every element of G is a power of g, so we will work strictly in terms of powers of g. If $H = \{я\}$,

then $H = \langle \mathfrak{s} \rangle = \langle g^0 \rangle$, and we are done. So assume there exists $x \in H$ such that $x \neq e$. By inclusion, every element $x \in H$ is also an element of G, which is generated by g, so $x = g^n$ for some $n \in \mathbb{Z}$. Without loss of generality, we may assume that $n \in \mathbb{N}^+$; after all, we just showed that we can choose $x \neq \mathfrak{s}$, so $n \neq 0$, and if $n \notin \mathbb{N}$, then closure of H implies that $x^{-1} = g^{-n} \in H$, so choose x^{-1} instead.

Now, if you were to take all the positive powers of g that appear in H, which would you expect to generate H? Certainly not the larger ones! The ideal candidate for the generator would be the smallest positive power of g in H, if it exists. Let S be the set of positive natural numbers i such that $g^i \in H$; in other words, $S = \{i \in \mathbb{N}^+ : g^i \in H\}$. The Well-Ordering Principle means that S has a smallest element; call it d, and assign $h = g^d$.

We claim that $H = \langle h \rangle$. Let $x \in H$; then $x \in G$. By hypothesis, G is cyclic, so $x = g^a$ for some $a \in \mathbb{Z}$. By the Division Theorem, we know that there exist unique $q, r \in \mathbb{Z}$ such that

- $a = qd + r$, and
- $0 \leqslant r < d$.

Let $y = g^r$; by Question 3.55, we can rewrite this as

$$y = g^r = g^{a-qd} = g^a g^{-(qd)} = x \cdot \left(g^d\right)^{-q} = x \cdot h^{-q}.$$

Now, $x \in H$ by definition, and $h^{-q} \in H$ by closure and the existence of inverses, so by closure $y = x \cdot h^{-q} \in H$ as well. We chose d as the smallest positive power of g in H, and we just showed that $g^r \in H$. Recall that $0 \leqslant r < d$. If $0 < r$; then $g^r \in H$, so $r \in S$. But $r < d$, which contradicts the choice of d as the smallest element of S. Hence r cannot be positive; instead, $r = 0$ and $x = g^a = g^{qd} = h^q \in \langle h \rangle$.

Since x was arbitrary in H, every element of H is in $\langle h \rangle$; that is, $H \subseteq \langle h \rangle$. Since $h \in H$ and H is a group, closure implies that $H \supseteq \langle h \rangle$, so $H = \langle h \rangle$. In other words, H is cyclic. □

We again look to $\mathbb{Z}$ for an example.

Question 4·16 .

Recall from Example 3·43 on page 84 that $\mathbb{Z}$ is cyclic; in fact $\mathbb{Z} = \langle 1 \rangle$. By Theorem 4·15, $d\mathbb{Z}$ is cyclic. In fact, $d\mathbb{Z} = \langle d \rangle$. Can you find another generator of $d\mathbb{Z}$?

Question 4·17 .

Recall that Ω_n, the nth roots of unity, is the cyclic group $\langle \omega \rangle$.

(a) List the elements of Ω_2 and Ω_4, and explain why $\Omega_2 < \Omega_4$.

(b) List the elements of Ω_8, and explain why both $\Omega_2 < \Omega_8$ and $\Omega_4 < \Omega_8$.

(c) Explain why, if $d \mid n$, then $\Omega_d < \Omega_n$.

Question 4·18 .

Show that even though the Klein 4-group is not cyclic, each of its proper subgroups is cyclic (see Definition 2·50 on page 61 and Questions 2.49 on page 61 and 3.45 on page 84).

Let G be any group and $g \in G$.
Claim: $\langle g \rangle < G$.
Proof:

1. Let $x, y \in$____.
2. By definition of ____, there exist $m, n \in \mathbb{Z}$ such that $x = g^m$ and $y = g^n$.
3. By ____, $y^{-1} = g^{-n}$.
4. By ____, $xy^{-1} = g^{m+(-n)} = g^{m-n}$.
5. By ____, $xy^{-1} \in \langle g \rangle$.
6. By ____, $\langle g \rangle < G$.

Figure 4·3: Material for Question 4.19

Question 4·19 .
Fill in each blank of Figure 4.19 with the appropriate justification or expression to show that the set of powers of an element g of a group G forms a subgroup of G.

Question 4·20 .
Explain why $\mathbb{R}$ cannot be cyclic. *Hint:* You already showed that one of its subgroups is not cyclic. Which one, and why does this make the difference?

Question 4·21 .
Recall that the ring of matrices $\mathbb{R}^{n\times n}$ is a ring, and therefore a group under addition, while the general linear group $\mathrm{GL}_n(\mathbb{R})$ is a group under multiplication.

(a) Let $D_n(\mathbb{R}) = \{aI_n : a \in \mathbb{R}\}$; that is, $D_n(\mathbb{R})$ is the set of all diagonal matrices whose values along the diagonal is constant. Show that $D_n(\mathbb{R}) < \mathbb{R}^{n\times n}$.

(b) Let $D_n^*(\mathbb{R}) = \{aI_n : a \in \mathbb{R}\setminus\{0\}\}$; that is, $D_n^*(\mathbb{R})$ is the set of all non-zero diagonal matrices whose values along the diagonal is constant. Show that $D_n^*(\mathbb{R}) < \mathrm{GL}_n(\mathbb{R})$.

Question 4·22 .
Recall the set of orthogonal matrices from Question 3.94.

(a) Show that $\mathrm{O}(n) < \mathrm{GL}(n)$. We call $O(n)$ the **orthogonal group**.

Let $\mathrm{SO}(n)$ be the set of all orthogonal $n \times n$ matrices whose determinant is 1. We call $\mathrm{SO}(n)$ the **special orthogonal group**.

(b) Show that $\mathrm{SO}(n) < \mathrm{O}(n)$.
Hint: The easiest way to show this requires some properties of determinants. Since you may not remember them, or may not even have *seen* them (it could depend on the class, on the teacher, on which universe you existed in the day they were presented…) here are the ones you need: for any matrix that has an inverse, $\det A = \det A^T$, $\det(AB) = (\det A)(\det B)$, and $\det A^{-1} = (\det A)^{-1}$.

In keeping with with the analogy of matrices, we say that the **kernel** of a group homomorphism is the subset of the domain that is sent to the identity of the range. That is, for a group homomorphism $f : G \to H$, we

$$g \in \ker f \qquad \text{iff} \qquad f(g) = я_H.$$

A ring homomorphism is a group homomorphism on the additive group of the ring, so the kernel of a ring homomorphism is the subset of the domain that is sent to the additive identity of the range, 0.

The kernel of a monoid is somewhat more complicated; we omit that for now.

Question 4·23 .
This question builds on Question 4.22. Let $\varphi : \mathrm{O}(n) \to \Omega_2$ by $\varphi(A) = \det A$.

(a) Show that φ is a homomorphism, but not an isomorphism.

(b) Explain why $\ker \varphi = \mathrm{SO}(n)$.

4·2 Ideals

A major reason for the study of rings and fields is to analyze polynomial roots. How do the roots of a polynomial behave with respect to basic arithmetic on the polynomials? Start with a ring R, an element $a \in R$, and two univariate polynomials f and g over R.

Example 4·24. Consider $R = \mathbb{Z}[x]$. Two polynomials with a root at $a = 3$ are $f(x) = x^2 - 9$ and $g(x) = x^2 - 7x + 12$. Their sum is $h(x) = 2x^2 - 7x + 3$, and $h(3) = 2 \times 9 - 7 \times 3 + 3 = 0$.

Adding f and g gave us a new polynomial, h, that also had a root at $a = 3$. This is true in general; if a is a root of two polynomials f and g, then a is also a root of both their sum *and* their difference $h = f - g$, since

$$h(a) = (f - g)(a) = f(a) - g(a) = 0.$$

Closure of subtraction means the Subgroup Theorem applies, giving us the following result.

Fact 4·25. *Let R be a ring, and $a \in R$. The set of polynomials with a root at a forms a subgroup of $R[x]$ under addition.*

We can do better. If a is a root of f, then it is a root of any polynomial multiple of f, such as $h = fp$. After all,

$$h(a) = (fp)(a) = f(a)\,p(a) = 0 \cdot p(a) = 0.$$

Example 4·26. Consider $R = \mathbb{R}[x]$ and $f = x^2 - 1$. The roots of f are ± 1. Let $p = x^4 + x^2 + 1$; the roots of p do not include ± 1; after all, $0 \neq 3 = p(1) = p(-1)$. On the other hand, let $h = fp = x^6 - 1$; we see quickly that ± 1 are indeed roots of h.

Even though p does not have a root at $x = \pm 1$, h does!

Let's put this together. Let $a \in R$ and A be the set of polynomials that have a root at a. Let f and g be any such polynomials; we saw above that their difference $f - g$ is also in A; that makes A a subgroup of R. *In addition,* any multiple of f is also in A, so there's something special about A: its element "absorbs" the products of its polynomials with others.

This property is not true within a group and its usual operation; even within the polynomial ring, adding a polynomial outside a subgroup to one within the subgroup always results in a polynomial outside the subgroup.

Question 4·27 .

Continuing the previous example, show that adding p to f gives you a polynomial that, like p, does *not* have a root at ± 1.

The phenomenon of absorption is quite simple. You will see that it appears in a number of important contexts. Here's an example.

Question 4·28 .

Let A be the set of all integers that are a sum of multiples of 4 and 6; that is,

$$A = \{4m + 6n : m, n \in \mathbb{Z}\}.$$

(a) Show that A is in fact a subgroup of $\mathbb{Z}$.

(b) Show that A absorbs multiplication by nonmembers; that is, $3a \in A$ for all $a \in A$.

Definition and examples

As usual, R is a ring.

Definition 4·29. Let $A \subseteq R$. If A

- is a subgroup of R under addition, and
- satisfies the **absorption property**:

$$\forall r \in R \quad \forall a \in A \qquad ra \in A \quad \text{and} \quad ar \in A,$$

then A is an **ideal** of R, and we write $A \lhd R$. An ideal A is **proper** if it is a proper subgroup under addition; that is, $\{0\} \neq A \neq R$.

Recall that we work in commutative rings unless otherwise specified, so if $ra \in A$ then usually $ar \in A$ is free.

Example 4·30. Recall the subring $2\mathbb{Z}$ of the ring $\mathbb{Z}$. We claim that $2\mathbb{Z} \lhd \mathbb{Z}$. Why? Let $r \in \mathbb{Z}$, and $a \in 2\mathbb{Z}$. By definition of $2\mathbb{Z}$, there exists $q \in \mathbb{Z}$ such that $a = 2q$. Substitution gives us

$$ra = r \cdot 2q = 2\,(rq) \in 2\mathbb{Z},$$

so $2\mathbb{Z}$ "absorbs" multiplication by $\mathbb{Z}$. We know from Example 4·7 that $2\mathbb{Z}$ was a subgroup of $\mathbb{Z}$ (use $d = 2$), so $2\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

We can generalize this example to arbitrary $d \in \mathbb{Z}$, so let's do that. Remember that you already know $d\mathbb{Z}$ is a subgroup of $\mathbb{Z}$; you need merely show that $d\mathbb{Z}$ absorbs multiplication.

Question 4·31 .
Show that for any $d \in \mathbb{N}$, $d\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

Our original example of an ideal came from roots of univariate polynomials. What about multivariate polynomials? If $a_1, \ldots, a_n \in R$, $f \in R[x_1, \ldots, x_n]$, and $f(a_1, \ldots, a_n) = 0$, then we call $(a_1, \ldots, a_n)$ a **root** of f.

Example 4·32. Let $f = x^2 + y^2 - 4$, $g = xy - 1$, and $S = \{hf + kg : h, k \in \mathbb{R}[x, y]\}$. As in the univariate case, *the common roots of f and g are roots of any element of S.* To see this, let (α, β) be a common root of f and g; that is, $f(\alpha, \beta) = g(\alpha, \beta) = 0$. Figure 4·4 depicts the root

$$(\alpha, \beta) = \left(\sqrt{2 + \sqrt{3}}, 2\sqrt{2 + \sqrt{3}} - \sqrt{6 + 3\sqrt{3}}\right).$$

Do all the elements of S have (α, β) as a root? Let $s \in S$; by definition, we can write $s = hf + kg$ for some $h, k \in \mathbb{R}[x, y]$. By substitution,

$$\begin{aligned} s(\alpha, \beta) &= (hf + kg)(\alpha, \beta) \\ &= h(\alpha, \beta) \cdot f(\alpha, \beta) + k(\alpha, \beta) \cdot g(\alpha, \beta) \\ &= h(\alpha, \beta) \cdot 0 + k(\alpha, \beta) \cdot 0 \\ &= 0; \end{aligned}$$

that is, (α, β) is a root of s. In fact, S is an ideal. To show this, we must show that S is a subring of $\mathbb{R}[x, y]$ that absorbs multiplication.

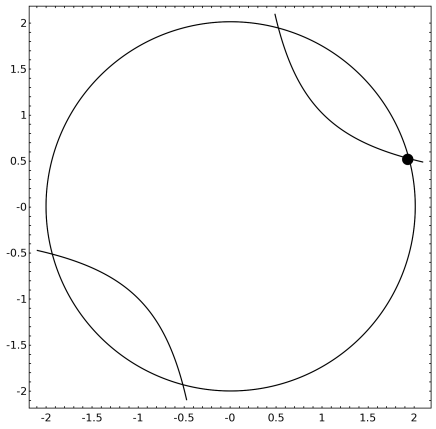


Figure 4·4: A common root of $x^2 + y^2 - 4$ and $xy - 1$

- Is S a subgroup under addition? Let $s, r \in S$. By definition, we can find $h, k, p, q \in \mathbb{R}[x, y]$ such that $s = hf + kg$ and $r = pf + qg$. A little arithmetic gives us

$$\begin{aligned} s - r &= (hf + kg) - (pf + qg) \\ &= (h - p) f + (k - q) g \in S. \end{aligned}$$

  A ring is an abelian group under addition, so the Subgroup Theorem implies S is a subgroup of $\mathbb{C}[x, y]$.

- Does S absorb multiplication? Let $s \in S$, and $p \in \mathbb{R}[x, y]$. As above, we can write $s = hf + kg$. A little arithmetic gives us

$$\begin{aligned} ps &= p(hf + kg) = p(hf) + p(kg) \\ &= (ph) f + (pk) g \in S. \end{aligned}$$

  Let

$$h' = ph \quad \text{and} \quad k' = pk;$$

  then $ps = h'f + k'g$. By closure, $h', k' \in \mathbb{R}[x, y]$, so by definition, $ps \in S$, as well. By definition, S satisfies the absorption property.

We have shown that S satisfies the subgroup and absorption properties; thus, $S \lhd \mathbb{R}[x, y]$.

You will show in Question 4.59 that the ideal of Example 4·32 can be generalized to other rings and larger numbers of variables.

Important properties of ideals

An ideal inherits the associative, commutative, and distributive properties of the ring. It also inherits closure of multiplication, though you might not notice why at first:

Fact 4·33. *An ideal is closed under multiplication.*

Why? Let A be an ideal of a ring R. Let $a, b \in A$. By absorption, $ab \in A$. □

An ideal might not contain the multiplicative identity. Proper ideals *never* contain the multiplicative identity.

Question 4·34 .
Let $A \lhd R$. Show that A is proper if and only if $A \neq \{0\}$ and $1 \notin A$.

Also, proper ideals *never* contain elements with multiplicative inverses.

Question 4·35 .
Let r be any nonzero element of a ring. Show that r has a multiplicative inverse if and only if any ideal that contains r also contains unity, and thus is not proper.

Since an ideal is really a special sort of subgroup, an analog of the Subgroup Theorem determines whether a subset of a ring is an ideal, with only one or two criteria.

The Ideal Theorem. *Let R be a ring and $A \subseteq R$ with A nonempty. The following are equivalent:*

(A) *A is an ideal of R.*

(B) *A is closed under subtraction and absorption. That is,*

- *(I1)* *for all $a, b \in A$, $a - b \in A$; and*
- *(I2)* *for all $a \in A$ and $r \in R$, we have $ar, ra \in A$.*

Question 4·36 .
Prove the Ideal Theorem.

Question 4·37 .
We can take Question 4.31 further. Fill in the blanks of Figure 4·5 to show that every ideal of $\mathbb{Z}$ has the form $d\mathbb{Z}$, for some $d \in \mathbb{N}$.

Question 4·38 .
Suppose A is an ideal of R and B is an ideal of S. Is $A \times B$ an ideal of $R \times S$?

Question 4·39 .
Let R be a ring and A, B two ideals of R. Decide whether the following subsets of R are also ideals, and explain your reasoning:

(a) $A \cap B$

(b) $A \cup B$

(c) $A + B = \{a + b : a \in A, b \in B\}$

(d) $AB = \{\sum_{i=1}^{n} a_i b : n \in \mathbb{N}, a_i \in A, b_i \in B\}$

Claim: Every ideal of $\mathbb{Z}$ has the form $d\mathbb{Z}$, for some $d \in \mathbb{Z}$.
Proof:

1. Let A be an ideal of $\mathbb{Z}$.
2. Let $D = A \cap \mathbb{N}^+$.
3. By ____, we can find a smallest element of D, which we call d.
4. We claim that $A = d\mathbb{Z}$. To see why, first let $b \in d\mathbb{Z}$. By definition of $d\mathbb{Z}$, $b =$____.
   (a) By ____, $b \in A$.
   (b) By ____, $d\mathbb{Z} \subseteq A$.
5. We now claim $A \subseteq d\mathbb{Z}$. To see why, let $a \in$____.
   (a) By ____, we can find $q, r \in \mathbb{Z}$ such that $a = qd + r$ and $0 \leqslant r < d$.
   (b) Rewrite the equation as $r =$____.
   (c) By ____, $qd \in A$.
   (d) By ____, $a - qd \in A$.
   (e) By ____, $r \in A$.
   (f) If $r > 0$, then $r \in D$, since ____.
   (g) However, we cannot have $r > 0$, since ____.
   (h) That forces $r =$____.
   (i) Hence d divides a, since ____.
   (j) By ____, $A \subseteq d\mathbb{Z}$.
6. We have shown $A \subseteq d\mathbb{Z}$ and $d\mathbb{Z} \subseteq A$. Hence ____.
7. ____ means that *every* ideal of $\mathbb{Z}$ has the form $d\mathbb{Z}$, for some $d \in \mathbb{Z}$.

Figure 4·5: Material for Question 4.37

Question 4·40 .
Let A, B be two ideals of a ring R. The definition of AB appears in Question 4.39.

(a) Show that $AB \subseteq A \cap B$.

(b) Show that sometimes $AB \neq A \cap B$; that is, find a ring R and ideals A, B such that $AB \neq A \cap B$. *Hint:* A good example is related to Bézout's Identity. Look at ideals generated by integers with a common divisor.

4·3 The basis of an ideal

The ideals of Questions 4.31 and 4.37 are cyclic subgroups of the additive group of $\mathbb{Z}$, so it makes sense to write

$$\langle d \rangle = d\mathbb{Z},$$

just as we write $\langle d \rangle$ for the cyclic group generated by d. This works in general, too.

Fact 4·41. *Let R be a ring, and $a \in R$. The set*

$$\langle a \rangle = \{ar : r \in R\}$$

is an ideal of R.

(Some authors use (a), and some use aR. We will stick with $\langle a \rangle$, but you are likely to see these other notations from time to time.)

Why? First we check that $\langle a \rangle$ is a subgroup of R under addition. Let $x, y \in \langle a \rangle$; by definition, there exist $r, s \in R$ such that $x = ar$ and $y = as$. Substitution and the distributive property show us that

$$x - y = ar - as = a(r - s) \in \langle a \rangle .$$

Let $r \in R$ and $b \in \langle a \rangle$. By definition, we can find $x \in R$ such that $b = ax$. Then $rb = r(ax) = r(xa) = (rx)a$; that is, rb is also a multiple of a. The arbitrary choice of r and b show that $\langle a \rangle$ absorbs multiplication; $\langle a \rangle$ is indeed an ideal of R. □

We call these ideals **principal ideals**. Principal ideals of the integers have a nice property that we will use in future examples.

Example 4·42. Certainly $3 \mid 6$ since $3 \cdot 2 = 6$. Look at the ideals generated by 3 and 6:

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

Inspection suggests that $\langle 6 \rangle \subseteq \langle 3 \rangle$. Is it? Let $x \in \langle 6 \rangle$. By definition, $x = 6q$ for some $q \in \mathbb{Z}$. By substitution, $x = (3 \cdot 2)q = 3(2 \cdot q) \in \langle 3 \rangle$. Since x was arbitrary in $\langle 6 \rangle$, we have $\langle 6 \rangle \subseteq \langle 3 \rangle$.

This property holds both in the integers and in every ring, using more or less the same reasoning. It will prove useful in subsequent sections.

Lemma 4·43. *Let R be a ring and $a, b \in R$. The following are equivalent:*

(A) $a \mid b$;

(B) $\langle b \rangle \subseteq \langle a \rangle$.

Question 4·44 .
Prove Lemma 4·43.

Ideals generated by more than one element

One way to look at $\langle d \rangle \subsetneq \mathbb{Z}$ is that $\langle d \rangle$ is the smallest ideal that contains d: any other ideal must contain all its multiples. Extending this line of thinking, define the set $\langle a_1, a_2, \ldots, a_m \rangle$ as the intersection of all the ideals of R that contain all of $a_1, a_2, \ldots, a_m$.

Theorem 4·45. *For any choice of $m \in \mathbb{N}^+$ and $a_1, a_2, \ldots, a_m \in R$, $\langle a_1, a_2, \ldots, a_m \rangle$ is an ideal.*

We will not prove this directly, as it follows immediately from:

Lemma 4·46. *For every set S of ideals of a ring R, $\bigcap_{I \in S} I$ is also an ideal.*

Proof. Let $J = \bigcap_{I \in S} I$. We are protected from $J \neq \varnothing$ by the fact that the additive identity 0 is an element of every ideal. Let $a, b \in J$ and $r \in R$. Let $I \in S$. Since J contains only those elements that appear in every element of S, and $a, b \in J$, we know that $a, b \in I$. By the Ideal Theorem, $a - b \in I$, and also $ra \in I$. Since I was an arbitrary ideal in S, *every* element of S contains $a - b$ and ra. Thus $a - b$ and every ra are in the intersection of these sets, which is J; in other words, $a - b, ra \in J$. By the Ideal Theorem, J is an ideal. □

Since $\langle a_1, a_2, \ldots, a_m \rangle$ is defined as the intersection of ideals containing $a_1, a_2, \ldots, a_m$, Theorem 4·46 implies that $\langle a_1, a_2, \ldots, a_m \rangle$ is an ideal. This ideal is closely related to Example 4·32, making it important enough to identify by a special name.

Definition 4·47. We call $\langle a_1, a_2, \ldots, a_m \rangle$ the **ideal generated by** $a_1, a_2, \ldots, a_m$, and $\{a_1, a_2, \ldots, a_m\}$ a **basis** of $\langle a_1, a_2, \ldots, a_m \rangle$.

Theorem 4·48. *For any commutative ring R, $\langle a_1, a_2, \ldots, a_m \rangle$ is precisely the set*

$$A = \{r_1 a_1 + r_2 a_2 + \cdots + r_m a_m : \ r_i \in R\}.$$

Proof. First, we show that $A \subseteq \langle a_1, a_2, \ldots, a_m \rangle$. Let $b \in A$; by definition, there exist $r_1, \ldots, r_m \in R$ such that $b = \sum_{i=1}^{m} r_i a_i$. Let I be any ideal that contains all of $a_1, \ldots, a_m$. By absorption, $r_i a_i \in I$ for each i. By closure, $b = \sum_{i=1}^{m} r_i a_i \in I$. Since I was an arbitrary ideal containing all of $a_1, \ldots, a_m$, we infer that all the ideals containing all of $a_1, \ldots, a_m$ contain b. Since b is an arbitrary element of A, A is a subset of all the ideals containing all of $a_1, \ldots, a_m$. By definition, $A \subseteq \langle a_1, a_2, \ldots, a_m \rangle$.

Now we show that $A \supseteq \langle a_1, a_2, \ldots, a_m \rangle$. We claim that A is (a) an ideal that (b) contains all of $a_1, \ldots, a_m$. If true, the definition of $\langle a_1, a_2, \ldots, a_m \rangle$ does the rest, as it consists of elements common to *every* ideal that contains $a_1, \ldots, a_m$.

(a) But why is A an ideal? Consider the absorption property. By definition of A, we can identify for any $b \in A$ ring elements $r_1, \ldots, r_m \in R$ such that

$$b = r_1 a_1 + \cdots + r_m a_m.$$

Let $p \in R$; by the distributive and associative properties,

$$pb = (pr_1)\, a_1 + \cdots + (pr_m)\, a_m.$$

By closure, $pr_i \in R$ for each $i = 1, \ldots, m$. We have written pb in a form that satisfies the definition of A, so $ps \in A$. We still need subtraction, so let $b, c \in A$, and choose $p_i, q_i \in R$ such that

$$\begin{aligned} b &= p_1 a_1 + \cdots + p_m a_m \text{ and} \\ c &= q_1 a_1 + \cdots + q_m a_m. \end{aligned}$$

Using the associative property, the commutative property of addition, the commutative property of multiplication, distribution, and the closure of subtraction, we see that

$$\begin{aligned} b - c &= (p_1 a_1 + \cdots + p_m a_m) - (q_1 a_1 + \cdots + q_m a_m) \\ &= (p_1 a_1 - q_1 a_1) + \cdots + (p_m a_m - q_m a_m) \\ &= (p_1 - q_1)\, a_1 + \cdots + (p_m - q_m)\, a_m. \end{aligned}$$

By closure, $p_i - q_i \in R$ for each $i = 1, \ldots, m$, so $b - c$ has a form that satisfies the definition of A, so $b - c \in A$. By the Ideal Theorem, A is an ideal.

(b) But, is $a_i \in A$ for each $i = 1, 2, \ldots, m$? Well,

$$a_i = 1 \cdot a_i + \sum_{j \neq i} \left(0 \cdot a_j\right) \quad \in \quad A.$$

Since R has unity, this expression of a_i satisfies the definition of A, so $a_i \in A$.

Hence A is an ideal containing all of $a_1, \ldots, a_m$. By definition of $\langle a_1, a_2, \ldots, a_m \rangle$, $A \supseteq \langle a_1, a_2, \ldots, a_m \rangle$.

We have shown that $A \subseteq \langle a_1, a_2, \ldots, a_m \rangle \subseteq A$. Hence $A = \langle a_1, a_2, \ldots, a_m \rangle$ as claimed. □

Remark 4·49. The structure and properties of ideals should remind you of *vector spaces* from linear algebra. In linear algebra, we analyze systems of *linear* equations. By manipulating a matrix, we obtain a *triangular basis* of a system of linear polynomials, with which we analyze the system's solutions.

Example 4·32 illustrates that ideals are an important analog for non-linear polynomial equations. As with linear systems, a "triangular basis" of a polynomial ideal allows us to analyze its solutions in a systematic method. We take up this task in due course… but not just yet.

Question 4·50 .
Let's explore how $\langle a_1, a_2, \ldots, a_m \rangle$ behaves in $\mathbb{Z}$. Keep in mind that the results do not necessarily generalize to other rings.

(a) For the following values of $a, b \in \mathbb{Z}$, list a few elements of $\langle a, b \rangle$. Then verify that $\langle a, b \rangle = \langle c \rangle$ for a certain $c \in \mathbb{Z}$.

(i) $a = 3, b = 6$

(ii) $a = 4, b = 6$

(iii) $a = 5, b = 6$

(iv) $a = 6, b = 6$

(b) Can you identify a relationship between a, b, and c in part (a)?

(c) Prove your observation in part (b).
Hint: Bézout's Identity can be useful.

Principal ideal domains

The basis of an ideal *need not be unique!*

Example 4·51. Consider the ring $\mathbb{Z}$, and let $I = \langle 6, 8 \rangle$. Proposition 4·48 claims that

$$I = \{6m + 8n : \ m, n \in \mathbb{Z}\}.$$

Choosing concrete values of m and n, we see that

$$\begin{aligned} 6 &= 6 \cdot 1 + 8 \cdot 0 \in I \\ 0 &= 6 \cdot 0 + 8 \cdot 0 \in I \\ -24 &= 6 \cdot (-4) + 8 \cdot 0 \in I \\ -24 &= 6 \cdot 0 + 8 \cdot (-3) \in I. \end{aligned}$$

Notice that for some elements of I, we can provide more than one representation in terms of 6 and 8.

While we're at it, we claim that we can simplify I as $I = 2\mathbb{Z}$. Why? For starters, it's pretty easy to see that $2 = 6 \cdot (-1) + 8 \cdot 1$, so $2 \in I$. Now that we have $2 \in I$, let $x \in 2\mathbb{Z}$; then $x = 2q$ for some $q \in \mathbb{Z}$. By substitution and distribution,

$$x = 2q = [6 \cdot (-1) + 8 \cdot 1] \cdot q = 6 \cdot (-q) + 8 \cdot q \in I.$$

Since x was arbitrary, $I \supseteq 2\mathbb{Z}$. On the other hand, let $x \in I$. By definition, there exist $m, n \in \mathbb{Z}$ such that

$$x = 6m + 8n = 2(3m + 4n) \in 2\mathbb{Z}.$$

Since x was arbitrary, $I \subseteq 2\mathbb{Z}$. We already showed that $I \subseteq 2\mathbb{Z}$, so we conclude that $I = 2\mathbb{Z}$.

So $I = \langle 6, 8\rangle = \langle 2\rangle = 2\mathbb{Z}$. If we think of $r_1, \ldots, r_m$ as a "basis" for $\langle r_1, \ldots, r_m\rangle$, then the example above shows that any given ideal can have bases of different sizes.

You might wonder if every ideal can be written as $\langle a\rangle$, the same way that $I = \langle 4, 6\rangle = \langle 2\rangle$. As you will see in due course, "Not always." However, the statement *is* true for ideals of $\mathbb{Z}$ (as you saw above), as well as a number of other rings. Rings where every ideal is principal, are called **principal ideal rings**. If the ring is an integral domain, we call it a **principal ideal domain**. Alas, not all integral domains are principal ideal domains.

Example 4·52. Let R be a ring, and $R[x, y]$ the ring of polynomials over R. Let $A = \langle x, y\rangle$. Can we find $f \in A$ such that $A = \langle f\rangle$?

We cannot. Suppose to the contrary that we could; in that case, both x and y would be multiples of f. This is not possible, because only 1 divides both x and y. If $f = 1$, then $1 \in A$, and $A = R$. That means A is not principal, and $R[x, y]$ is not a principal domain.

Theorem 4·53. *The following rings* are *principal ideal domains.*

(A) $\mathbb{Z}$ *is a principal ideal domain.*

(B) Any field is a principal ideal domain (so $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, *and finite fields* $\mathbb{F}_n$ *are principal ideal domains).*

(C) Any univariate *polynomial ring over a field is a principal ideal domain.*

Proof. (A) You proved this when you answered Question 4.37, since $\langle d\rangle = d\mathbb{Z}$.

(B) Let A be an ideal in a field. If $A = \{0\}$, then $A = \langle 0\rangle$. Otherwise, let a be a non-zero element of A. As an element of a field, it has a multiplicative inverse a^{-1}; by absorption, $a^{-1}a \in A$. By Question 4.35, A is not proper. Every improper ideal is generated by the multiplicative identity; that is, $A = \langle 1\rangle$.

Question 4·54 .
How do we know that if $A = R$, then $A = \langle 1\rangle$?

Proof of Theorem 4·53 (continued). (C) Let $\mathbb{F}$ be any field, $R = \mathbb{F}[x]$, and A an ideal of R. Let $D = \{\deg f : f \in A\}$; that is, D is the set of all degrees of polynomials in A.

Example 4·55. Suppose that $f = 2x^3 - 3x$, $g = 5x^7 - 12$, and $h = 128x^2 - 2x + 13$ are all elements of A. Then $3, 7, 2 \in D$.

Proof of Theorem 4·53 (continued). Degrees are nonnegative integers, so $D \subseteq \mathbb{N}$. By the Well-Ordering Principle, there is a least element of D; call it d. By definition of D, there exists $f \in A$ such that $\deg f = d$. Let c be the leading coefficient of f, and let $g = c^{-1}f$. By absorption, $g \in A$; by polynomial arithmetic, $\deg g = d$ and the leading coefficient of g is now 1.

Let h be any element of A. Use Polynomial Division to identify $q, r \in R$ such that $h = qg + r$ and $r \neq 0$ or $\deg r < \deg g$. If $r = 0$, then h is a multiple of g, as claimed, and we're done. Otherwise, rewrite the division equation as

$$r = h - qg.$$

By absorption, $qg \in A$. By definition, $h \in A$. By the Ideal Theorem, $h - qg \in A$, so $r \in A$ itself. If $r \neq 0$, then $\deg r < \deg g = d$, contradicting the choice of d as the smallest element of D, the degrees of polynomials in A. Hence $r = 0$, and g divides h. We chose h arbitrarily in A, and found that g *has* to divide h. That makes every element of a a multiple of g, so $A = \langle g \rangle$. We chose A as an arbitrary ideal of R, and found it was principal. That makes every ideal of R principal, as claimed. □

Question 4·56 .

Show that in any field $\mathbb{F}$, the only two distinct ideals are the zero ideal and $\mathbb{F}$ itself.
Hint: Consider Question 4.54.

Question 4·57 .

Let R be any ring and $P = R[x, y]$. Let $A = \langle x + 1, xy \rangle$, $B = \langle x, y \rangle$, and $C = \langle x, y + 1 \rangle$.

(a) Show that $A = P$.

Hint: Use the result of Question 4.34.

(b) Show that $B \neq P$ and $C \neq P$.

Hint: Proceed by contradiction. We need $1 \in B$ (why?) so there must be polynomials $f, g \in P$ such that $xf + yg = 1$. The right side is constant, so x and y must cancel on the left. That forces f and g to have a certain form — what form is it? Following this to its conclusion leads to a contradiction.

Question 4·58 .

Let A and B be ideals of R. Define $A \cdot B = \{ab : a \in A, b \in B\}$. (This is not the same as AB, defined in Question 4.39.)

(a) Show that $A \cdot B$ need not be an ideal.

Hint: Two ideals of Question 4.57 do the trick.

(b) Show that if R is a commutative, principal ideal ring, then $A \cdot B$ is an ideal.

Question 4·59 .

Let R be any commutative ring. Recall the polynomial ring $P = R[x_1, x_2, \ldots, x_n]$, whose ground ring is R. Let

$$\langle f_1, \ldots, f_m \rangle = \{h_1 f_1 + \cdots + h_m f_m : \ h_1, h_2, \ldots, h_m \in P\}.$$

Show that the common roots of $f_1, f_2, \ldots, f_m$ are common roots of all polynomials in this ideal.

Question 4·60 .
Let A be an ideal of a ring R. Define its **radical** to be

$$\sqrt{A} = \left\{ r \in R : r^n \in A \ \exists n \in \mathbb{N}^+ \right\}.$$

(a) Suppose $R = \mathbb{Z}$. Compute $\sqrt{A}$ for

(i) $A = 4\mathbb{Z}$

(ii) $A = 5\mathbb{Z}$

(iii) $A = 12\mathbb{Z}$
Hint: Every element of $12\mathbb{Z}$ is a multiple of 12, so it will help to look at how 12 factors. How could you simplify those factors so that some power of the simplification is a multiple of 12?

(b) Suppose $R = \mathbb{Q}[x]$. Compute $\sqrt{A}$ for

(i) $A = \langle x^2 + 1 \rangle$

(ii) $A = \langle x^2 + 2x + 1 \rangle$

(iii) $A = \langle x^3 + x^2 - x - 1 \rangle$

(c) Show that $\sqrt{A}$ is an ideal.
Hint: You need to show that if $a, b \in \sqrt{A}$, then $ab, a + b \in \sqrt{A}$. The hypothesis implies that you can find m and n such that $a^m \in A$ and $b^n \in A$. Use m and n to build an exponent e such that $(a + b)^e \in A$. As a further hint, a very big e is probably easier than the smallest possible e. As a *final* hint, don't forget that *you already know* elements of A absorb multiplication — you only have to show that this is also true of elements of $\sqrt{A}$.

4·4 Equivalence relations and classes

> *I remember one occasion when I tried to add a little seasoning to a review … The domains of the underlying measures were not sets but elements of more general Boolean algebras, and their range consisted not of positive numbers but of certain abstract equivalence classes. My proposed first sentence was: "The author discusses valueless measures in pointless spaces."*
>
> — Paul Halmos

At this point we can tie together two topics that share a relationship you likely haven't noticed yet. In the following section, we tie it to a third phenomenon. At the end of the chapter, these will come together in a very beautiful relationship.

Throughout this section, $d \in \mathbb{N}^+$. We've written $a \equiv_d b$ using a symbol $\equiv$ that looks like an equals sign, but does it *behave* like an equals sign? Don't rush into an answer; just because I use a symbol that looks like an equals sign, that doesn't mean it is. Three important and useful properties of an equals sign are the *reflexive, symmetric,* and *transitive* properties.

Definition 4·61. An **equivalence relation on** S is a subset R of $S \times S$ that satisfies the properties

reflexive: $a \sim a$ for all $a \in S$;

symmetric: for all $a, b \in S$, if $a \sim b$, then $b \sim a$; and

transitive: for all $a, b, c \in S$, if $a \sim b$ and $b \sim c$, then $a \sim c$.

Does the $\equiv_d$ relationship of clockwork arithmetic satisfy these three properties?

Fact 4·62. *For any integer a, $a \equiv_d a$.*

Why? The statement $a \equiv_d a$ translates to, "a and a have the same remainder after division by d." Even if we divide different ways, the Division Theorem guarantees that remainders are unique! So our clockwork equivalence is "reflexive", in that any integer is equivalent to itself. □

Fact 4·63. *For any integers a and b, $a \equiv_d b$ implies $b \equiv_d a$.*

Why? The statement that "a and b have the same remainder after division by d" surely means (thanks in part to uniqueness of remainder) that "b and a have the same remainder after division by d," so our clockwork equivalence is symmetric. □

Is it also *transitive?* This is a big deal, because *substitution* is a powerful tool, and substitution requires the transitive property; that is,

$$\text{If} \quad a \equiv_d b \quad \text{and} \quad b \equiv_d c, \quad \text{then} \quad \text{is } a \equiv_d c \quad ?$$

What we're asking translates to,

- if a and b have the same remainder after division by d, and
- b and c have the same remainder after division by d, then
- do a and c have the same remainder after division by d?

Fact 4·64. *For any integers a, b, and c, $a \equiv_d b$ and $b \equiv_d c$ imply that $a \equiv_d c$.*

Why? Let r be the remainder of division of a by d. This remainder is unique, so $a \equiv_d b$ means it's the same as the remainder of division of b by d. Likewise, $b \equiv_d c$ tells us that r is the remainder of division of c by d. We have $a \equiv_d c$. □

There are plenty of relations that *aren't* equivalence relations.

Example 4·65. Define a relation $\sim$ on $\mathbb{Z}$ such that $a \sim b$ if $ab \in \mathbb{N}$. Is this an equivalence relation?

Reflexive? Let $a \in \mathbb{Z}$. By properties of arithmetic, $a^2 \in \mathbb{N}$. By definition, $a \sim a$, and the relation is reflexive.

Symmetric? Let $a, b \in \mathbb{Z}$. Assume that $a \sim b$; by definition, $ab \in \mathbb{N}$. By the commutative property of multiplication, $ba \in \mathbb{N}$ also, so $b \sim a$, and the relation is symmetric.

Transitive? Let $a, b, c \in \mathbb{Z}$. Assume that $a \sim b$ and $b \sim c$. By definition, $ab \in \mathbb{N}$ and $bc \in \mathbb{N}$. I could argue that $ac \in \mathbb{N}$ using the trick

$$ac = \frac{(ab)\,(bc)}{b^2},$$

and pointing out that ab, bc, and b^2 are all natural, which suggests that ac is also natural. However, this argument contains a fatal flaw. Do you see it?

It lies in the fact that we don't know whether $b = 0$. If $b \neq 0$, then the argument above works just fine, but if $b = 0$, then we encounter division by 0, which you surely know is not allowed! (If you're not sure *why* it is not allowed, fret not. We explain this in a moment.)

This apparent failure should not discourage you; in fact, it gives us the answer to our original question. We asked if $\sim$ was an equivalence relation. *It is not!* This illustrates an important principle of mathematical study. Failures like this typically suggested an unexpected avenue to answer a question. In this case, the fact that $a \cdot 0 = 0 \in \mathbb{N}$ for any $a \in \mathbb{Z}$ implies that $1 \sim 0$ and $-1 \sim 0$. However, $1 \not\sim -1$! The relation is *not* transitive, so it *cannot* be an equivalence relation on this set!

In the context of an equivalence relation, related elements of a set are considered "equivalent".

Example 4·66. Let $\sim$ be a relation on $\mathbb{Z}$ such that $a \sim b$ if and only if a and b have the same remainder after division by 4. Then $7 \sim 3$ and $7 \sim 19$ but $7 \not\sim 6$.

We will find it *very useful* to group elements that are equivalent under a certain relation.

Definition 4·67. Let $\sim$ be an equivalence relation on a set A, and let $a \in A$. The **equivalence class** of a in A with respect to $\sim$ is $[a] = \{b \in S : a \sim b\}$, the set of all elements equivalent to a.

Example 4·68. Continuing our example above, $3, 19 \in [7]$ but $6 \notin [7]$.

Normally, we think of the division of n by d as dividing a set of n objects into q groups, where each group contains d elements, and r elements are left over. For example, $n = 23$ apples divided among $d = 6$ bags gives $q = 3$ apples per bag and $r = 5$ apples left over.

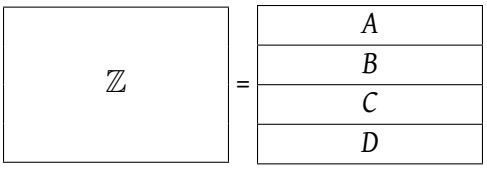
Another way to look at division by d is that it sorts every integer into one of d sets, according to its remainder after division. An illustration using $d = 4$:

$\mathbb{Z}$:	…	-2	-1	0	1	2	3	4	5	6	…
		↓	↓	↓	↓	↓	↓	↓	↓	↓	
division by 4:	…	2	3	0	1	2	3	0	1	2	…

In other words, division by 4 "divides" $\mathbb{Z}$ into the sets

$$\begin{aligned} A &= \{\ldots, -4, 0, 4, 8, \ldots\} = [0] \\ B &= \{\ldots, -3, 1, 5, 9, \ldots\} = [1] \\ C &= \{\ldots, -2, 2, 6, 10, \ldots\} = [2] \\ D &= \{\ldots, -1, 3, 7, 11, \ldots\} = [3]. \end{aligned} \tag{4.1}$$

Observe that



which means to say that

- the sets A, B, C, and D *cover* $\mathbb{Z}$; that is,

$$\mathbb{Z} = A \cup B \cup C \cup D;$$

  and

- the sets A, B, C, and D are *disjoint;* that is,

$$A \cap B = A \cap C = A \cap D = B \cap C = B \cap D = C \cap D = \varnothing.$$

When a collection $\mathcal{B}$ of subsets of a set S form a disjoint cover, we call that collection a **partition**.

Example 4·69. In the example above, $S = \mathbb{Z}$ and the collection $\mathcal{B} = \{A, B, C, D\}$ where A, B, C, and D are defined as in (4.1). Since the elements of $\mathcal{B}$ are disjoint, and they cover $\mathbb{Z}$, we conclude that $\mathcal{B}$ is a partition of $\mathbb{Z}$.

There is nothing special about the number "4" in this discussion; clockwork arithmetic always induces a partition. Is this true of every equivalence relation?

Surprisingly, yes! Actually, it isn't so surprising if you just think about the meaning of an equivalence relation:

- the reflexive property implies that every element is in relation with itself, and
- the symmetric and transitive properties help ensure that no element can be related to two elements that are not themselves related.

Theorem 4·70. *An equivalence relation partitions a set, and any partition of a set defines an equivalence relation.*

Proof. Does any partition of any set define an equivalence relation? Let S be a set, and $\mathcal{B}$ a partition of S. Define a relation $\sim$ on S in the following way: $x \sim y$ if and only if x and y are in the same element of $\mathcal{B}$. That is, if $X \in \mathcal{B}$ is the set such that $x \in X$, then $y \in X$ as well.

We claim that $\sim$ is an equivalence relation. It is reflexive because a partition covers the set; that is, $S = \bigcup_{B \in \mathcal{B}}$, so for any $x \in S$, we can find $B \in \mathcal{B}$ such that $x \in B$, which means the statement that "x is in the same element of $\mathcal{B}$ as itself" ($x \sim x$) actually makes sense. The relation is symmetric because $x \sim y$ means that x and y are in the same element of $\mathcal{B}$, which is equivalent to saying that y and x are in the same element of $\mathcal{B}$; after all, set membership is not affected by which element we list first. So, if $x \sim y$, then $y \sim x$. Finally, the relation is transitive because distinct elements of a partition are disjoint. Let $x, y, z \in S$, and assume $x \sim y$ and $y \sim z$. Choose $X, Z \in \mathcal{B}$ such that $x \in X$ and $z \in Z$. The symmetric property tells us that $z \sim y$, and the definition of the relation implies that $y \in X$ and $y \in Z$. The fact that they

share a common element tells us that X and Z are not disjoint ($X \cap Z \neq \varnothing$). By the definition of a partition, X and Z are not distinct, or, $X = Z$. That shows x and z are in the same element of the partition, so $x \sim z$.

Does an equivalence relation partition a set? Let S be a set, and $\sim$ an equivalence relation on S. If S is empty, the claim is "vacuously true;"that is, nothing about S can make it false. So assume S is non-empty. Let $s \in S$. Notice that $[s] \neq \varnothing$, since the reflexive property of an equivalence relation guarantees that $s \sim s$, which implies that $s \in [s]$.

Let $\mathcal{B}$ be the set of all equivalence classes of elements of S; that is, $\mathcal{B} = \{[s] : s \in S\}$. We have already seen that every $s \in S$ appears in its own equivalence class, so $\mathcal{B}$ covers S. Are distinct equivalence classes also disjoint?

Let $X, Y \in \mathcal{B}$ and assume that assume that $X \cap Y \neq \varnothing$; this means that we can choose $z \in X \cap Y$. By definition, $X = [x]$ and $Y = [y]$ for some $x, y \in S$. By definition of $X = [x]$ and $Y = [y]$, we know that $x \sim z$ and $y \sim z$. Now let $w \in X$ be arbitrary; by definition, $x \sim w$; by the symmetric property of an equivalence relation, $w \sim x$ and $z \sim y$; by the transitive property of an equivalence relation, $w \sim z$, and by the same reasoning, $w \sim y$. Since w was an arbitrary element of X, *every* element of X is related to y; in other words, every element of X is in $[y] = Y$, so $X \subseteq Y$. A similar argument shows that $X \supseteq Y$. By definition of set equality, $X = Y$.

We took two arbitrary equivalence classes of S, and showed that if they were not disjoint, then they were not distinct. The contrapositive states that if they are distinct, then they are disjoint. Since the elements of $\mathcal{B}$ are equivalence classes of S, we conclude that distinct elements of $\mathcal{B}$ are disjoint. They also cover S, so as claimed, $\mathcal{B}$ is a partition of S induced by the equivalence relation. □

Question 4·71 .

(a) Show that divisibility is transitive for the natural numbers; that is, if $a, b, c \in \mathbb{N}$, $a \mid b$, and $b \mid c$, then $a \mid c$.

(b) However, divisibility is not an equivalence relation. Show that it is not symmetric.

(c) In fact, divisibility is a partial ordering for the natural numbers. Show why.

(d) Can a partial ordering ever be an equivalence relation? Explain.

Question 4·72 .

(a) Explain why $2 \cdot 3 \equiv_6 0$.

(b) Integer equations such as

$$(x+1)(x+2) = 0$$

rely on the equivalence relation properties of equality. In this case, we can solve the equation by rewriting it as

$$x + 1 = 0 \quad \text{or} \quad x + 2 = 0.$$

Explain how part (a) shows that we cannot do this for

$$(x+1)(x+2) \equiv_6 0.$$

We observe, then, that integer equations really are a special kind of equivalence relation; that is, they enjoy a property that not all equivalence relations enjoy, even when they look similar.

Question 4·73 .

Define a relation $\bowtie$ on $\mathbb{Q}$, the set of rational numbers, in the following way:

$$a \bowtie b \text{ if and only if } a - b \in \mathbb{Z}.$$

(a) Give some examples of rational numbers that are related. Include examples where a and b are not themselves integers.

(b) Show that that $a \bowtie b$ if they have the same sign and the same *fractional part*. That is, if we write a and b in decimal form, we see exactly the same numbers on the right hand side of the decimal point, in exactly the same order. (You may assume, without proof, that we can write any rational number in decimal form.)

(c) Is $\bowtie$ an equivalence relation?

For any $a \in \mathbb{Q}$, let S_a be the set of all rational numbers b such that $a \bowtie b$. We'll call these new sets **classes**.

(d) Is every $a \in \mathbb{Q}$ an element of some class? If so, which? If not, why not?

(e) Show that if $S_a \neq S_b$, then $S_a \cap S_b = \varnothing$.

(f) Does $\bowtie$ partition $\mathbb{Q}$?

So far, we've restricted ourselves to talking about clockwork groups, but here's the surprise: these are intimately related to isomorphism. We tease you with your first hint here, another hint in the next section, and the full glory later on.

Question 4·74 .
Let $(M, \times)$, $(N, +)$, and $(P, \sqcap)$ be monoids.

(a) Show that the identity function $I(m) = m$ is an isomorphism on M.

(b) Suppose that we know $(M, \times) \cong (N, +)$. That means there is an isomorphism $f : M \to N$. One of the requirements of isomorphism is that f be a bijection. Recall from previous classes that this means f has an inverse *function*, $f^{-1} : N \to M$. Show that f^{-1} is an isomorphism.

Hint: You need to show that $f^{-1}(xy) = f^{-1}(x) f^{-1}(y)$ for every $x, y \in N$. You already know f is an isomorphism, so you can find $a, b \in M$ such that $f(a) = x$ and $f(b) = y$. The fact that f is a homomorphism will help you a *lot* with showing f^{-1} is a homomorphism.

(c) Suppose that we know $(M, \times) \cong (N, +)$ and $(N, +) \cong (P, \sqcap)$. As above, we know there exist isomorphisms $f : M \to N$ and $g : N \to P$. Let $h = g \circ f$; that is, h is the composition of the functions g and f. Explain why $h : M \to P$, and show that h is also an isomorphism.

(d) Explain how (a), (b), and (c) prove that isomorphism is an equivalence relation.

4·5 Clockwork rings and ideals

In this section, we combine our work using remainders to create a consistent "clockwork arithmetic" (Sections 2·1, 4·4, and 3·4) with our observation that the multiples of an integer form an ideal of a ring, and thus a subgroup of a group (Section 4·2). We highlight some relationships between these two phenomena, which the following sections generalize to other situations.

Recall that we defined $\mathbb{Z}_d$ as the set of remainders $\{0, 1, \ldots, d-1\}$ and that this forms a ring under addition and multiplication, modulo d. This congruence relationship (modulo d) is an equivalence relation, and we saw that this means it partitions the integers via the elements of $d\mathbb{Z}$.

Example 4·75. In Section 4·4 we considered the case where $d = 4$. We'll rename those equivalence classes from A, B, C, and D to

$$\begin{aligned} 4\mathbb{Z} &= \{\ldots, -4, 0, 4, 8, \ldots\} \\ 1 + 4\mathbb{Z} &= \{\ldots, -3, 1, 5, 9, \ldots\} \\ 2 + 4\mathbb{Z} &= \{\ldots, -2, 2, 6, 10, \ldots\} \\ 3 + 4\mathbb{Z} &= \{\ldots, -1, 3, 7, 11, \ldots\}. \end{aligned}$$

We will see in a moment that we can write them differently, using *any* element of that equivalence class:

$$\begin{aligned} 4 + 4\mathbb{Z} &= \{\dots, -4, 0, 4, 8, \dots\} \\ -3 + 4\mathbb{Z} &= \{\dots, -3, 1, 5, 9, \dots\} \\ 10 + 4\mathbb{Z} &= \{\dots, -2, 2, 6, 10, \dots\} \\ 7 + 4\mathbb{Z} &= \{\dots, -1, 3, 7, 11, \dots\}. \end{aligned}$$

However, it's typical to use the remainder, and we call that way of writing these equivalence classes the **canonical representation** for each equivalence class.

In general, if X is an equivalence class of the remainder after division by d, we write $X = x + d\mathbb{Z}$ for any $x \in X$. This notation causes no confusion, since the equivalence class is a partition, and forces every element of $\mathbb{Z}$ into a unique class. We can actually make a stronger statement:

Fact 4·76. *Two such equivalence classes X and Y d are equal if and only if any representations $X = x + d\mathbb{Z}$ and $Y = y + d\mathbb{Z}$ satisfy the relationship $d \mid (x - y)$.*

Why? The equivalence classes partition $\mathbb{Z}$, so $X = Y$ if and only if $x \equiv y$ modulo d. By definition, $d \mid (x - y)$. □

For instance, our example above shows that $1 + 4\mathbb{Z} = -3 + 4\mathbb{Z}$. Here we have $x = 1$ and $y = -3$, and indeed $4 \mid (1 - (-3))$.

Henceforth we write $\mathbb{Z}/d\mathbb{Z}$ for the set of equivalence classes of the remainders after division by d. Another observation:

Fact 4·77. *The set $\mathbb{Z}/d\mathbb{Z}$ of equivalence classes of the remainders after division by d forms a ring under the following arithmetic:*

$$(a + d\mathbb{Z}) + (b + d\mathbb{Z}) = (a + b) + d\mathbb{Z} \quad \textit{and} \quad (a + d\mathbb{Z})(b + d\mathbb{Z}) = (ab) + d\mathbb{Z}.$$

In fact, this ring is isomorphic to $\mathbb{Z}_d$.

Example 4·78. Recall that $\mathbb{Z}/4\mathbb{Z} = \{4\mathbb{Z}, 1 + 4\mathbb{Z}, 2 + 4\mathbb{Z}, 3 + 4\mathbb{Z}\}$. Addition in this group will always give us one of those four representations of the classes:

$$\begin{aligned} (2 + 4\mathbb{Z}) + (1 + 4\mathbb{Z}) &= 3 + 4\mathbb{Z}; \\ (1 + 4\mathbb{Z}) + (3 + 4\mathbb{Z}) &= 4 + 4\mathbb{Z} = 4\mathbb{Z}; \\ (2 + 4\mathbb{Z}) + (3 + 4\mathbb{Z}) &= 5 + 4\mathbb{Z} = 1 + 4\mathbb{Z}; \end{aligned}$$

and so forth. Likewise, multiplication will give us one of those four representations of classes:

$$\begin{aligned} (0 + 4\mathbb{Z})(2 + 4\mathbb{Z}) &= 0 + 4\mathbb{Z}; \\ (1 + 4\mathbb{Z})(3 + 4\mathbb{Z}) &= 3 + 4\mathbb{Z}; \\ (2 + 4\mathbb{Z})(3 + 4\mathbb{Z}) &= 6 + 4\mathbb{Z} = 2 + 4\mathbb{Z}; \end{aligned}$$

and so forth.

Why is Fact 4·77 true? Let $f : \mathbb{Z}_d \to (\mathbb{Z}/d\mathbb{Z})$ map a remainder r to the equivalence class $r + d\mathbb{Z}$. We claim that f is one-to-one and onto, and it also preserves addition, multiplication, and multiplicative identity. In this case, $\mathbb{Z}/d\mathbb{Z}$ will be a ring, as claimed. To see why, observer that any sum of classes corresponds to addition of two remainders, their preimages via f. The sum of these remainders gives another remainder, which f maps to a class that corresponds to the defined addition. This shows closure of addition; the remaining properties will follow similarly.

So let $a, b \in \mathbb{Z}_d$; f maps them to $A = a + d\mathbb{Z}$ and $B = b + d\mathbb{Z}$. First we show the homomorphism properties of a ring. For addition, we need to show that $f(a + b)$ is the same class as

$$f(a) + f(b) = (a + d\mathbb{Z}) + (b + d\mathbb{Z}) = (a + b) + d\mathbb{Z}.$$

Let r be the remainder of division of $a + b$ by d; we have have

$$f(a + b) \underset{\text{subst}}{=} f(r) \underset{\text{def of } f}{=} r + d\mathbb{Z}.$$

So we really need to show that

$$(a + b) + d\mathbb{Z} = r + d\mathbb{Z};$$

that is, $a + b$ and r lie in the same equivalence class. By the definition of our equivalence classes, this is equivalent to saying that $a + b \equiv_d r$, but that is true by definition of r (the remainder of $a + b$). Hence $f(a + b) = f(a) + f(b)$. Preservation of multiplication is shown so similarly that we pass over it. As for the multiplicative identity,

$$(1 + d\mathbb{Z})(a + d\mathbb{Z}) = a + d\mathbb{Z} = (a + d\mathbb{Z})(1 + d\mathbb{Z})$$

regardless of the choice of a, making $1 + d\mathbb{Z}$ the identity of $\mathbb{Z}/d\mathbb{Z}$, but $f(1) = 1 + d\mathbb{Z}$, so the identity is preserved. It remains to show that f is one-to-one and onto.

One-to-one? Let $a, b \in \mathbb{Z}_d$, and assume $f(a) = f(b)$. By definition of f, this means $a + d\mathbb{Z} = b + d\mathbb{Z}$; by Fact 4·76, $d \mid (a - b)$. As remainder, however, $0 \leqslant a, b < d$, so $-d < a - b < d$. The only multiple of d between $-d$ and d itself is 0, so $a - b = 0$; in other words, $a = b$.

Onto? For any class $a + d\mathbb{Z}$, let r be the remainder of division of a by d; then $f(r) = r + d\mathbb{Z}$. We need $f(r) = a + d\mathbb{Z}$, but this is no problem; by the Division Theorem, we can find $q \in \mathbb{Z}$ such that $a = qd + r$, or $a - r = qd$, which by Fact 4·76 means $f(r) = r + d\mathbb{Z} = a + d\mathbb{Z}$, as desired. □

It is burdensome to write $a + n\mathbb{Z}$ whenever we want to discuss an element of $\mathbb{Z}/d\mathbb{Z}$, so we adopt the following convention.

Notation 4·79. Let $A \in \mathbb{Z}/d\mathbb{Z}$ and choose $a \in \mathbb{Z}$ such that $A = a + d\mathbb{Z}$.

- If it is clear from context that A is an element of $\mathbb{Z}/d\mathbb{Z}$, then we simply write a instead of $a + d\mathbb{Z}$.
- If we want to emphasize that A is an element of $\mathbb{Z}/d\mathbb{Z}$ (perhaps there are a lot of integers hanging about) then we write $[a]_d$ instead of $a + d\mathbb{Z}$.

- If the value of d is obvious from context, we simply write $[a]$.

To help you grow accustomed to the notation $[a]_d$, we use it for the rest of this chapter, even when d is mindbogglingly obvious.

Definition 4·80. On account of Fact 4·77, we can designate the remainder of division of a by d, whose value is between 0 and $|d| - 1$, inclusive, as the **canonical representation** of $[a]_d$ in $\mathbb{Z}/d\mathbb{Z}$.

Question 4·81 .
Write out the Cayley tables for $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$ (both addition and multiplication).

Question 4·82 .
Write out the Cayley table for $\mathbb{Z}/5\mathbb{Z}$ (both addition and multiplication). Which elements generate $\mathbb{Z}/5\mathbb{Z}$?

Question 4·83 .
Write down the Cayley table for $\mathbb{Z}/6\mathbb{Z}$ (both addition and multiplication). Which elements generate $\mathbb{Z}/6\mathbb{Z}$?

We now present two more properties. Both properties follow immediately from the isomorphism between $\mathbb{Z}/d\mathbb{Z}$ and $\mathbb{Z}_d$, so we do not provide any further proof.

Theorem 4·84. *$\mathbb{Z}/d\mathbb{Z}$ is finite for every nonzero $d \in \mathbb{Z}$. In fact, if $d \neq 0$ then $\mathbb{Z}/d\mathbb{Z}$ has $|d|$ elements corresponding to the remainders from division by d: 0, 1, 2, …, $d - 1$.*

Question 4·85 .
What if $d = 0$? How many elements would $\mathbb{Z}/d\mathbb{Z}$ have? You can't use division here, so you have to rely on the equivalence classes, not the isomorphism. Illustrate a few additions and subtractions, and indicate whether you think that $\mathbb{Z}/0\mathbb{Z}$ is an interesting or useful group.

Question 4·86 .
In the future, we won't consider $\mathbb{Z}/d\mathbb{Z}$ when $d < 0$. Show that this is because $\mathbb{Z}/d\mathbb{Z} = \mathbb{Z}/|d|\mathbb{Z}$. (Notice that this asks for equality, not merely isomorphism.)

Questions 2.47 on page 61 and 2.48 on page 61 tell us that there is only one group of order 2 (up to isomorphism) and only one group of order 3 (up to isomorphism). So the structure of $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$ was determined well before you ever looked at Question 4.81. On the other hand, there are two possible structures for a group of order 4: the Klein 4-group, and a cyclic group. (See Question 2.49 on page 61.) Which of these structures does $\mathbb{Z}/4\mathbb{Z}$ have? Again, isomorphism gives it away.

Theorem 4·87. *$\mathbb{Z}/d\mathbb{Z}$ is cyclic for every $d \in \mathbb{Z}$.*

This theorem has a more general version, which you will prove in the homework.

A natural and interesting followup question is, which non-zero elements *do* generate $\mathbb{Z}/d\mathbb{Z}$? You need a bit more background in number theory before you can answer that question, but you can still formulate a hypothesis.

Question 4·88 .
Write out a Cayley table for $\mathbb{Z}_4$, and compare it to the results of Questions 4.81, 4.82, and 4.83. Formulate a conjecture as to which elements generate $\mathbb{Z}_n$, for arbitrary n.

Question 4·89 .
Use Bézout's Lemma to prove your conjecture in Question 4.88. *Hint:* If $a \in \mathbb{Z}_n$ generates $\mathbb{Z}_n$, then $ab = 1$ for some $b \in \mathbb{Z}$. Bézout's Lemma should help you find b. On the other hand, if 1 is a multiple of a, then so is every other element of $\mathbb{Z}_n$ — why?

The following important lemma gives an "easy" test for whether two integers are in the same class of $\mathbb{Z}/d\mathbb{Z}$, and summarizes what we have done in this section.

Lemma 4·90. *Let $a, b, d \in \mathbb{Z}$ and assume that $d \neq 0$. The following are equivalent.*

(A) $a + d\mathbb{Z} = b + d\mathbb{Z}$.

(B) $[a]_d = [b]_d$.

(C) $d \mid (a - b)$.

Proof. (A) is equivalent to (B) by definition of the notation $[a]_d$ (see above), and (A) is equivalent to (C) by Fact 4·76. □

4·6 Partitioning groups and rings

We saw in Section 4·4 how clockwork arithmetic uses division to partition the integers according to their remainder. We also found that this partition has group and ring structures; for instance, it's pretty clear that $3 + 5 \equiv_6 2$, but a few additions and subtractions show that $3 \equiv -3$, $5 \equiv 11$, and $2 \equiv 62$; the equivalence classes thus tell us that $-3 + 11 \equiv 62$. We also saw in Section 3·1 that working with division of polynomials gave us a way to model roots and build complex numbers.

Can we do this with other groups and rings? Indeed we can, using a tool called *cosets*. Students often have a hard time wrapping their minds around cosets, so we'll start with an introductory example that should give you an idea of how cosets "look" in a group. Then we'll define cosets, and finally look at some of their properties.

The idea

Two aspects of division were critical for making clockwork arithmetic an equivalence relation, and thus a way to partition of $\mathbb{Z}$:

- *existence of a remainder,* which implies that every integer belongs to at least one class, which in turn implies that the union of the classes covers $\mathbb{Z}$; and
- *uniqueness of the remainder,* which implies that every integer ends up in only one set, so that the classes are disjoint.

Using the vocabulary of groups, recall from Section 138 the sets

$$\begin{aligned} A &= \{\ldots, -4, 0, 4, 8, \ldots\} = [0] \\ B &= \{\ldots, -3, 1, 5, 9, \ldots\} = [1] \\ C &= \{\ldots, -2, 2, 6, 10, \ldots\} = [2] \\ D &= \{\ldots, -1, 3, 7, 11, \ldots\} = [3]. \end{aligned}$$

Recall from Section 118 that $A = 4\mathbb{Z} < \mathbb{Z}$, so it is a group under addition. The other sets are *not* groups; after all, they lack the additive identity.

What interests us is how the equivalence classes relate to the subgroup. All elements of B have the form $1 + a$ for some $a \in A$. For example, $-3 = 1 + (-4)$. Likewise, all elements of C have the form $2 + a$ for some $a \in A$, and all elements of D have the form $3 + a$ for some $a \in A$. So if we define

$$1 + A := \{1 + a : \ a \in A\},$$

then

$$\begin{aligned} 1 + A &= \{\ldots, 1 + (-4), 1 + 0, 1 + 4, 1 + 8, \ldots\} \\ &= \{\ldots, -3, 1, 5, 9, \ldots\} \\ &= B. \end{aligned}$$

Likewise, we can write $A = 0 + A$ and $C = 2 + A$, $D = 3 + A$.

Pursuing this further, you can check that

$$\cdots = -3 + A = 1 + A = 5 + A = 9 + A = \cdots$$

and so forth. Interestingly, all the sets in the previous line are the same as B! In addition, $1 + A = 5 + A$, and $1 - 5 = -4 \in A$. The same holds for C: $2 + A = 10 + A$, and $2 - 10 = -8 \in A$. This relationship will prove important at the end of the section.

So the partition by remainders of division by four is related to the subgroup A of multiples of 4. How can we generalize this phenomenon to other groups, even nonabelian ones?

Definition 4·91. Let G be a group and $A < G$. Let $g \in G$. We define the **left coset of** A **with** g as

$$gA = \{ga : \ a \in A\}$$

and the **right coset of** A **with** g as

$$Ag = \{ag : \ a \in A\}.$$

In general, left cosets and right cosets are not equal, partly because the operation might not commute. If we speak of "cosets" without specifying "left" or "right", we mean "left cosets".

Example 4·92. Recall the group D_3 from Section 3·6 and the subgroup $H = \langle \varphi \rangle = \{\iota, \varphi\}$ from Example 4·8. In this case,

$$\rho H = \{\rho, \rho\varphi\} \text{ and } H\rho = \{\rho, \varphi\rho\}.$$

Since $\varphi\rho = \rho^2\varphi \neq \rho\varphi$, we see that $\rho H \neq H\rho$.

Question 4·93 .
In Question 4.17, you showed that $\Omega_2 < \Omega_8$. Compute the left and right cosets of Ω_2 in Ω_8.

Question 4·94 .
Let $\{\text{ϗ}, a, b, a+b\}$ be the Klein 4-group. (See Questions 2.49 on page 61, 3.45 on page 84, and 4.18 on page 124.) Compute the left and right cosets of $\langle a \rangle$.

Question 4·95 .
Compute the left and right cosets of $\langle \mathbf{j} \rangle$ in Q_8.

For some subgroups, left and right cosets are always equal. This is always true in abelian groups, as illustrated by Example 4·97.

Question 4·96 .
Show explicitly why left and right cosets are equal in abelian groups.

If A is an additive subgroup, we write the left and right cosets of A with g as $g + A$ and $A + g$. Rings are abelian groups under addition, with ideals as subgroups, so if R is a ring, $A \lhd R$, and $r \in R$, then we write the coset of A with r is $r + A$. For now we focus on the theory of cosets in the context of groups, as this applies equally to cosets of ideals of rings.

Example 4·97. Consider the subgroup $H = \{(a, 0) : a \in \mathbb{R}\}$ of $\mathbb{R}^2$ from Question 4.14. Let $p = (3, -1) \in \mathbb{R}^2$. The coset of H with p is

$$\begin{aligned} p + H &= \{(3,-1) + q : q \in H\} \\ &= \{(3,-1) + (a,0) : a \in \mathbb{R}\} \\ &= \{(3+a,-1) : a \in \mathbb{R}\}. \end{aligned}$$

Sketch some of the points in $p + H$, and compare them to your sketch of H in Question 4.14. How does the coset compare to the subgroup?

Generalizing this further, every coset of H has the form $p + H$ where $p \in \mathbb{R}^2$. Elements of $\mathbb{R}^2$ are points, so $p = (x, y)$ for some $x, y \in \mathbb{R}$. The coset of H with p is

$$p + H = \{(x + a, y) : a \in \mathbb{R}\}.$$

Sketch several more cosets. How would you describe the set of *all* cosets of H in $\mathbb{R}^2$?

Question 4·98 .
Recall the subgroup L of $\mathbb{R}^2$ from Question 4.14 on page 123.

(a) Give a geometric interpretation of the coset $(3, -1) + L$.

(b) Give an algebraic expression that describes $p + L$, for arbitrary $p \in \mathbb{R}^2$.

(c) Give a geometric interpretation of the cosets of L in $\mathbb{R}^2$.

(d) Use your answers to (a) and (c) give a geometric description of how cosets of L partition $\mathbb{R}^2$.

A group does not *have* to be abelian for the left and right cosets to be equal. When deciding if $gA = Ag$, we are not deciding *whether elements of G commute*, but *whether subsets of G are equal*. Returning to D_3, we can find a subgroup whose left and right cosets are equal even though the group is not abelian and the operation is not commutative.

Example 4·99. Let $K = \{\iota, \rho, \rho^2\}$; certainly $K < D_3$, after all, $K = \langle\rho\rangle$. In this case, $\alpha K = K\alpha$ for all $\alpha \in D_3$:

α	αK	$K\alpha$
ι	K	K
φ	$\{\varphi, \varphi\rho, \varphi\rho^2\}$	$\{\varphi, \rho\varphi, \rho^2\varphi\}$
ρ	K	K
ρ^2	K	K
$\rho\varphi$	$\{\rho\varphi, (\rho\varphi)\,\rho, (\rho\varphi)\,\rho^2\}$	$\{\rho\varphi, \varphi, \rho^2\varphi\}$
$\rho^2\varphi$	$\{\rho^2\varphi, (\rho^2\varphi)\,\rho, (\rho^2\varphi)\,\rho^2\}$	$\{\rho^2\varphi, \rho\varphi, \varphi\}$

In each case, the sets φK and $K\varphi$ are equal, even though φ does not commute with ρ. (You should verify these computations by hand.)

Question 4·100 .
In Question 4.12 on page 122, you found another subgroup K of order 2 in D_3. Does K satisfy the property $\alpha K = K\alpha$ for all $\alpha \in D_3$?

When a subgroup's left and right cosets are always equal, we call it a **normal subgroup** of its group. Normal subgroups play a critical role in later sections, but we won't worry too much about them at the moment.

You might notice a few things. In each case, every element appears in a coset: a subgroup A always contains the identity, so any g appears in "its own" coset gA. On the other hand, g seems to appear *only* in gA, and in no other other coset! After all, φK and $(\rho\varphi)\,K$ differ only superficially; when you consider their contents, you find that they are equal. This sounds an awful lot like the partition we were aiming for. Does it hold true in general? What other properties might cosets contain?

Properties of Cosets

We present some properties of cosets that illustrate further their similarities to division.

Theorem 4·101. *The cosets of a subgroup partition the group.*

Before proving this, we pause to point out that combining Theorems 4·101 and 4·70 implies another nice result.

Corollary 4·102. *Let $A < G$. Define a relation $\sim$ on $x, y \in G$ by*

$$x \sim y \iff x \text{ is in the same coset of } A \text{ as } y.$$

This relation is an equivalence relation.

We will make repeated use of this equivalence relation.

Proof of Theorem 4·101. Let G be a group, and $A < G$. We have to show two things:

(CP1) the cosets of A cover G, and

(CP2) distinct cosets of A are disjoint.

We show (CP1) first. Let $g \in G$. The definition of a group tells us that $g = g\text{я}$. Since $\text{я} \in A$ by definition of subgroup, $g = g\text{я} \in gA$. Since g was arbitrary, every element of G is in some coset of A. Hence the union of all the cosets is G.

For (CP2), let X and Y be arbitrary cosets of A. Assume that X and Y are distinct; that is, $X \neq Y$. We need to show that they are disjoint; that is, $X \cap Y = \varnothing$. We will show the contrapositive instead; that is, we will assume $X \cap Y \neq \varnothing$, and show $X = Y$. A contrapositive is logically equivalent to the original statement, so we will have accomplished our goal.

To prove the contrapositive, assume $X \cap Y \neq \varnothing$. By definition of intersection, we can find $z \in X \cap Y$. By definition of a coset, there exist $x, y \in G$ such that $X = xA$ and $Y = yA$; we can write $z = xa$ and $z = yb$ for some $a, b \in A$. By substitution, $xa = yb$, so $x = (yb)\, a^{-1}$, or

$$x = y\left(ba^{-1}\right). \tag{4.2}$$

We still have to show that $X = Y$. We show this by showing that $X \subseteq Y$ and $X \supseteq Y$. For the former, let $w \in X$; by definition of X, $w = x\text{ą}$ for some $\text{ą} \in A$. Applying our conversion mechanism,

$$w = x\text{ą} = \left[y\left(ba^{-1}\right)\text{ą}\right] = y\left[\left(ba^{-1}\right)a\right] \in yA.$$

We chose w as an arbitrary element of X, so $X \subseteq Y$. The proof that $X \supseteq Y$ is so similar that we omit it. By definition of set equality, $X = Y$. Inasmuch as X and Y were arbitrary, this holds for all cosets of A: if two cosets of A are not disjoint, then they are not distinct.

Having shown (CP2) and (CP1), we have shown that the cosets of A partition G. □

We conclude this section with three facts that allow us to decide when cosets are equal.

Lemma 4·103 (Equality of cosets)**.** *Let G be a group and $A < G$. All of the following hold:*

(CE1) $\text{я}A = A$.

(CE2) For all $g \in G$, $gA = A$ if and only if $g \in A$.

(CE3) For all $g, h \in G$, $gA = hA$ if and only if $g \in hA$.

(CE4) For all $g, h \in G$, $gA = hA$ if and only if $g^{-1}h \in A$.

As usual, you should keep in mind that in additive groups (and thus in rings) the first three conditions translate to

(CE1) $0 + A = A$.

(CE2) For all $g \in G$, $g \in A$ if and only if $g + A = A$.

(CE3) For all $g, h \in G$, $g + A = h + A$ if and only if $g \in h + A$.

(CE4) For all $g, h \in G$, $g + A = h + A$ if and only if $g - h \in A$.

Notice also that characterization (CE4) resembles the third criterion of the Subgroup Theorem. *The resemblance is mostly superficial;* in the Subgroup Theorem, $a^{-1}b$ refers to elements of A, while (CE4) refers to elements of G that are not always in A. That said, if it is the case that $g, h \in A$ then the Subgroup Theorem tells us that $g^{-1}h \in A$, so $gA = hA$ — though we already knew that from (CE2), since $gA = A = hA$.

Proof. (CE1) is "obvious" (but you will fill in the details in Question 4.105.).

We jump to (CE3) for the moment. Let $g, h \in G$. We know that $я \in A$, so $g = gя \in gA$. Corollary 4·102 tells us that membership in a coset is an equivalence relation, where the cosets are the equivalence classes. By substitution, $gA = hA$ if and only if $g \in hA$.

We turn to (CE2). Let $g \in G$. By (CE3), $gA = яA$ if and only if $g \in яA$. By (CE1), $яA = A$, so by substitution, $g \in A$ if and only if $gA = A$.

We finally turn to (CE4). Let $g, h \in G$. By (CE3), $gA = hA$ if and only if $g \in hA$. By definition of a coset, $g \in hA$ if and only if $g = ha$ for some $a \in A$. Applying the inverse property twice, we rewrite this equation first as $я = g^{-1}(ha)$, then (after an associative property) as $a^{-1} = g^{-1}h$. Since $a^{-1} \in A$, we have $g^{-1}h \in A$. Every step used an equivalence, so we can connect the chain into the one equivalence, $gA = hA$ if and only if $g^{-1}h \in A$. □

Property (CE4) does little more than restate the partition property, with the added knowledge that any elements lies in its own coset. However, it emphasizes that, when computing cosets of a subgroup A, you can skip hA whenever h appears in gA.

Question 4·104 .
Consider the ideal $A = \langle x^2 + 1 \rangle$ in $\mathbb{R}[x]$. Why can we write every coset of A as $(ax + b) + A$, where $a, b \in \mathbb{R}$? *Hint:* This is related to the isomorphism of Section 3·1.

Question 4·105 .
Fill in each blank of Figure 4.105 with the appropriate justification or statement.

4·7 Lagrange's Theorem and the order of an element of a group

How many cosets can a subgroup have? This section answers this question, as well as some related questions about the size of a subgroup and the order of an element. Throughout this section, we assume $|G|$ is finite, even if we don't say so explicitly.

Notation 4·106. Let G be a group, and $A < G$. We write G/A for the set of all left cosets of A. That is,

$$G/A = \{gA : g \in G\}.$$

We also write $A\backslash G$ for the set of all right cosets of A:

$$A\backslash G = \{Ag : g \in G\}.$$

Let G be a group and $H < G$.
Claim: $\text{я}H = H$.

1. First we show that _____. Let $x \in \text{я}H$.
   (a) By definition of $\text{я}H$, $x =$_____.
   (b) By the identity property, _____.
   (c) By substitution, $x \in$_____.
   (d) We had chosen an arbitrary element of $\text{я}H$, so by inclusion, _____.
2. Now we show the converse, that $\text{я}H \supseteq H$. Let $x \in$_____.
   (a) By the identity property, _____.
   (b) By definition of $\text{я}H$, _____$\in \text{я}H$.
   (c) We had chosen an arbitrary element, so by inclusion, _____.

Figure 4·6: Material for Question 4.105

Example 4·107. Let $G = \mathbb{Z}$ and $A = 4\mathbb{Z}$. We saw in Example 4·69 that

$$G/A = \mathbb{Z}/4\mathbb{Z} = \{A, 1 + A, 2 + A, 3 + A\}.$$

We actually "waved our hands" in Example 4·69. That means that we did not provide a very detailed argument, so let's show the details here. Recall that $4\mathbb{Z}$ is the set of multiples of $\mathbb{Z}$, so $x \in A$ iff x is a multiple of 4. What about the remaining elements of $\mathbb{Z}$?

Let $x \in \mathbb{Z}$; then

$$x + A = \{x + z : \ z \in A\} = \{x + 4n : \ n \in \mathbb{Z}\}.$$

Use the Division Theorem to write

$$x = 4q + r$$

for unique $q, r \in \mathbb{Z}$, where $0 \leqslant r < 4$. Then

$$x + A = \{(4q + r) + 4n : \ n \in \mathbb{Z}\} = \{r + 4(q + n) : \ n \in \mathbb{Z}\}.$$

By closure, $q + n \in \mathbb{Z}$. If we write m in place of $4(q + n)$, then $m \in 4\mathbb{Z}$. So

$$x + A = \{r + m : \ m \in 4\mathbb{Z}\} = r + 4\mathbb{Z}.$$

The distinct cosets of A are thus determined by the distinct remainders from division by 4. Since the remainders from division by 4 are 0, 1, 2, and 3, we conclude that

$$\mathbb{Z}/A = \{A, 1 + A, 2 + A, 3 + A\}$$

as claimed above.

Example 4·108. Let $G = D_3$ and $K = \{\iota, \rho, \rho^2\}$ as in Example 4·99, then

$$G/K = D_3/\langle\rho\rangle = \{K, \varphi K\}.$$

Example 4·109. Let $H < \mathbb{R}^2$ be as in Example 4·13 on page 122; that is,

$$H = \left\{(a, 0) \in \mathbb{R}^2 : \ a \in \mathbb{R}\right\}.$$

Then

$$\mathbb{R}^2/H = \left\{r + H : \ r \in \mathbb{R}^2\right\}.$$

It is not possible to list all the elements of G/H, but some examples would be

$$(1, 1) + H, \ (4, -2) + H.$$

Question 4·110 .
Speaking *geometrically*, what do the elements of $\mathbb{R}^2/H$ look like? This question is similar to Question 4.98.

Keep in mind that G/A is a set whose elements are also sets. Showing equality of two elements of G/A requires one to show that two sets are equal.

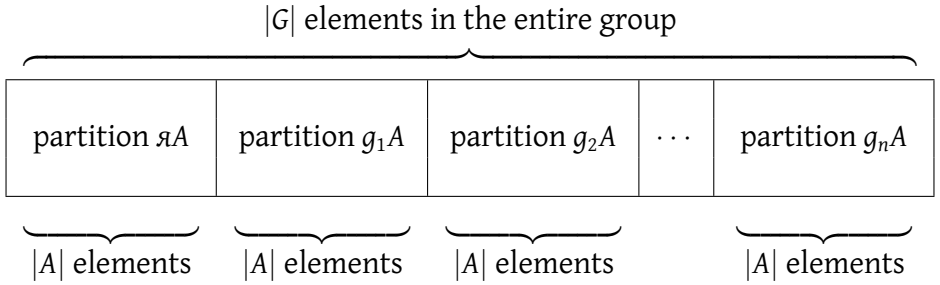
Remember our assumption that G is finite. In this case, a simple formula gives us the size of G/A.

Lagrange's Theorem. *Let $A < G$. The size of G/A is the ratio of the number of elements of G to the number of elements of A. That is,*

$$|G/A| = \frac{|G|}{|A|}.$$

While the notation of cosets is somewhat suggestive of the relationship between cosets and division, Lagrange's Theorem is *not* as obvious as the notation might imply: we can't "divide" the sets G and A. We are not moving the absolute value bars "inside" the fraction; nor can we, as G/A is not a number. Rather, we are partitioning the group G by the cosets of its subgroup A, and counting the number of sets that result.

Proof. We know from Theorem 4·101 that the cosets of A partition G. How many such cosets are there? $|G/A|$, by definition! Each coset has the same size, $|A|$. Suppose there are n cosets; we can visualize the partition in this fashion:



A basic principle of counting tells us that the number of elements of G is thus the product of the number of elements in each coset and the number of cosets. That is, $|G/A| \cdot |A| = |G|$. This implies the theorem. □

The next-to-last sentence of the proof contains the statement $|G/A| \cdot |A| = |G|$. Since $|A|$ is the order of the group A, and $|G/A|$ is an integer, we conclude that:

Corollary 4·111. *The order of any subgroup of G divides the order of the group.*

Example 4·112. Let G be the Klein 4-group (see Questions 2.49 on page 61, 3.45 on page 84, and 4.18 on page 124). Every subgroup of the Klein 4-group has order 1, 2, or 4. As predicted by Corollary 4·111, the orders of the subgroups divide the order of the group.

Likewise, the order of $\{\iota, \varphi\}$ divides the order of D_3.

By contrast, the subset HK of D_3 that you computed in Question 4.12 on page 122 has four elements. Since $4 \nmid 6$, the contrapositive of Lagrange's Theorem implies that HK *cannot* be a subgroup of D_3.

From the fact that every element g generates a cyclic subgroup $\langle g \rangle < G$, Lagrange's Theorem also implies an important consequence about the order of any element of any finite group.

Corollary 4·113. *The order of any element of a group divides the order of a group.*

Proof. You do it! See Question 4.115. □

Question 4·114 .
Recall from Question 4.17 that if $d \mid n$, then $\Omega_d < \Omega_n$. How many cosets of Ω_d are there in Ω_n?

Question 4·115 .
Fill in each blank of Figure 4.115 with the appropriate justification or expression.

Question 4·116 .
Suppose that a group G has order 8, but is not cyclic. Why must $g^4 = \mathfrak{s}$ for all $g \in G$?

Question 4·117 .
Let G be a finite group, and $g \in G$. Why is $g^{|G|} = \mathfrak{s}$?

Question 4·118 .
Suppose that a group has five elements. Why *must* it be abelian?

Question 4·119 .
Find a criterion on the order of a group that guarantees the group is cyclic.

Claim: The order of an element of a group divides the order of a group.
Proof:

1. Let G _____.
2. Let x _____.
3. Let $H = \langle$_____$\rangle$.
4. By _____, every integer power of x is in G.
5. By _____, H is the set of integer powers of x.
6. By _____, $H < G$.
7. By _____, $|H|$ divides $|G|$.
8. By _____, ord (x) divides $|H|$.
9. By definition, there exist $m, n \in$_____ such that $|H| = m\text{ord}\,(x)$ and $|G| = n\,|H|$.
10. By substitution, $|G| =$_____.
11. _____.
    (This last statement must include a justification.)

Figure 4·7: Material for Question 4.115

Question 4·120 .
Let p be an irreducible number, and recall that $\mathbb{Z}_p$ is a field, so that its non-zero elements form a group under multiplication. For instance, in $\mathbb{Z}_7$, the set $\{1, 2, 3, 4, 5, 6\}$ forms a group under multiplication. Explain why, for every $a \in \mathbb{Z}_p$,

(a) $a^{p-1} = 1$, and

(b) $a^p = a$, and

(c) $a^{p-2} = a^{-1}$.

This fact is called **Fermat's Little Theorem**. We explore it in a general context later.

4·8 Quotient Rings and Groups

Consider the polynomial ring $\mathbb{R}[x]$. Looking at remainders from division by $x^2 + 1$ gave us a way to model complex numbers as

$$\mathbb{C} = \{ax + b : a, b \in \mathbb{R}\},$$

where $1x + 0$ stood in for the imaginary number. An isomorphism (Question 3.18) showed that this was equivalent to the traditional model of the complex numbers, with $1x + 0 \mapsto i$.

Since then, we pointed out that every multiple of $x^2 + 1$ has the imaginary number i as a root. Multiples of $x^2 + 1$ have two things in common. First, dividing such polynomials by $x^2 + 1$ gives a remainder of 0. Second, and equivalently, they are in the ideal $A = \langle x^2 + 1 \rangle$. Question 4.104 showed us that the cosets of $\langle x^2 + 1 \rangle$ correspond to remainders from division by $x^2 + 1$. As noted, those remainders formed a field isomorphic to $\mathbb{C}$. In other words, the cosets of $\langle x^2 + 1 \rangle$ give us *another* model of the field $\mathbb{C}$.

Can we do this for cosets of general groups? To make the question precise, let $A < G$. Can we find a natural generalization of the operation(s) of G that makes G/A a group? By a "natural" generalization, we mean something like

$$(gA) * (hA) = (gh)A.$$

Quotient rings

The first order of business it to make sure that the operation even makes sense. The technical word for this is that the operation is **well-defined**. *What does that mean?* A coset can have different representations. An operation must be a function: for every pair of elements, it must produce *exactly one* result. The relation above would not be an operation if different representations of a coset gave us different answers. Example 4·121 shows how it can go wrong.

Example 4·121. Recall $H = \langle\varphi\rangle < D_3$ from Example 4·92. Let $X = \rho H$ and $Y = \rho^2 H$. Notice that $(\rho\varphi)\,H = \{\rho\varphi, \iota\} = \rho H$, so X has two representations, ρH and $(\rho\varphi)\,H$.

Were the operation well-defined, XY would have the same value, *regardless of the representation of X.* That is not the case! When we use the the first representation,

$$XY = (\rho H)\left(\rho^2 H\right) = \left(\rho \circ \rho^2\right) H = \rho^3 H = \iota H = H.$$

When we use the second representation,

$$\begin{aligned} XY = ((\rho\varphi)\,H)\left(\rho^2 H\right) &= \left((\rho\varphi)\,\rho^2\right) H = \left(\rho\left(\varphi\rho^2\right)\right) H \\ &= (\rho\,(\rho\varphi))\,H = \left(\rho^2\varphi\right) H \neq H. \end{aligned}$$

On the other hand, sometimes the operation *is* well-defined.

Example 4·122. Recall the subgroup $A = 4\mathbb{Z}$ of $\mathbb{Z}$. Let $B, C, D \in \mathbb{Z}/A$, so $B = b + 4\mathbb{Z}$, $C = c + 4\mathbb{Z}$, and $D = d + 4\mathbb{Z}$ for some $b, c, d \in \mathbb{Z}$.

We have to make sure that we cannot have $B = D$ and $B + C \neq D + C$. For example, if $B = 1 + 4\mathbb{Z}$ and $D = 5 + 4\mathbb{Z}$, $B = D$. Does it follow that $B + C = D + C$?

From Lemma 4·103, we know that $B = D$ iff $b - d \in A = 4\mathbb{Z}$. That is, $b - d = 4m$ for some $m \in \mathbb{Z}$. Let $x \in B + C$; then $x = (b + c) + 4n$ for some $n \in \mathbb{Z}$. By substitution,

$$x = ((d + 4m) + c) + 4n = (d + c) + 4\,(m + n) \in D + C.$$

Since x was arbitrary in $B + C$, we have $B + C \subseteq D + C$. A similar argument shows that $B + C \supseteq D + C$, so the two are, in fact, equal.

The operation was well-defined in the second example, but not the first. What made for the difference? In the second example, we rewrote

$$((d + 4m) + c) + 4n = (d + c) + 4\,(m + n),$$

but that relies on the fact that *addition commutes in an abelian group*. Without that fact, we could not have swapped c and $4m$.

Right away we see that we can *always* do this for cosets of ideals: after all, ideals are subgroups of rings under addition. Indeed, we can say something more.

Fact 4·123. *Let R be a commutative ring. The cosets of an ideal A of R form a new ring, whose addition and multiplication are natural generalizations of the addition and multiplication of R. That is, for any $r, s \in R$,*

$$(r + A) + (s + A) = (r + s) + A \quad \textit{and} \quad (r + A)\,(s + A) = rs + A.$$

Why? Let $A \lhd R$, and let $X, Y, Z \in R/A$.

Our first task is to show that addition and multiplication are well-defined. To do this, we need to show that the definitions of $X + Y$ and XY give us the same result, *regardless of the representation we choose for X and Y.* To this end, suppose there exist $r, s, x, y \in A$ such that $X = x + A = r + A$ and $Y = y + A = s + A$.

addition? We need to show that $(x+y)+A = (r+s)+A$. The lemma on Coset Equality tells us that this is true if and only if $-(x+y)+(r+s) \in A$, so we'll aim to show this latter expression is true. By hypothesis, $x+A = r+A$, so by Coset Equality, $-x+r \in A$. Similarly, $-y+s \in A$. By closure of addition and properties of ring addition,

$$-(x+y)+(r+s) = (-x+r)+(-y+s) \in A.$$

As we explained earlier, this shows that $(x+y)+A = (r+s)+A$; in short, the addition is well-defined.

multiplication? We need to show that $xy+A = rs+A$. The lemma on Coset Equality tells us that this is true if and only if $-(xy)+rs \in A$, so we'll aim to show this latter expression is true. Recall from the previous paragraph that $-x+r, -y+s \in A$. To get from these to $-(xy)+rs$, we'll use a fairly standard trick of adding zero in "the right form",

$$-(xy)+rs = -(xy)+ry-ry+rs = y(-x+r)+r(-y+s).$$

Absorption implies that $y(-x+r), r(-y+s) \in A$. Closure implies their sum is also in A. By substitution, $-(xy)+rs \in A$. As we explained earlier, this shows that $xy+A = rs+A$; in short, the multiplication is well-defined.

The remaining properties of addition are relatively straightforward. Choose $x, y, z \in R$ such that $X = x+A$, $Y = y+A$, $Z = z+A$.

associative? $(X+Y)+Z = [(x+A)+(y+A)]+(z+A) = [(x+y)+A]+(z+A) = [(x+y)+z]+A$. Apply the associative property of addition in R to obtain $(X+Y)+Z = [x+(y+z)]+A$. Now reverse the simplification to obtain $(X+Y)+Z = (x+A)+[(y+z)+A] = (x+A)+[(y+A)+(z+A)] = X+(Y+Z)$. The ends of this latter chain of equalities show the associative property is satisfied.

identity? We want a coset W such that $X+W = X$ and $W+X = X$. Let $w \in R$ such that $W = w+A$; by substitution, our first desired equation becomes $(x+A)+(w+A) = x+A$, or $(x+w)+A = x+A$. By coset equality, we need $(x+w)-x \in A$; by simplification, $w \in A$. From coset equality (Theorem 4·103(CE2)) that choosing any $a \in A$ gives us A itself, so $W = A$ must be the identity.

inverse? We want an "inverse" coset of $X = x+A$. The natural suspect would be $(-x)+A$; that is, the coset of A with $-x$. Indeed, it works great: $(x+A)+[(-x)+A] = 0+A = A$, and likewise $[(-x)+A]+(x+A) = A$. We just showed A is the identity of R/A, so we have found the inverse of X.

abelian? $X+Y = (x+A)+(y+A) = (x+y)+A$. Apply the commutative property of addition in R to obtain $X+Y = (y+x)+A = (y+A)+(x+A) = Y+X$. The ends of this latter chain of equalities show the addition is abelian.

We've found that R/A is an abelian group under the proposed addition, the first step towards showing it's a ring. We still need to show that multiplication satisfies the properties of a monoid, along with distribution.

We leave the remaining, multiplicative properties of a ring to you, the reader. □

Question 4·124 .
Show the remaining properties of a ring for R/A: closure, associative, identity, and distributive.

"Normal" subgroups

What about the cosets of nonabelian groups? Given the example above, you might be inclined to dismiss them, but that would be too hasty.

The key in Example 4·122 was not that $\mathbb{Z}$ is abelian, but that we could rewrite $(4m + c) + 4n$ as $c + (4m + 4n)$, then simplify $4m + 4n$ to $4(m + n)$. The abelian property makes it easy to do that, but we don't need the *group* G to be abelian; we need the *subgroup* A to satisfy it. If A were not abelian, we could still make it work if, after we move c left, we get *some* element of A to its right, so that it can be combined with the other one. That is, we have to be able to rewrite any ac as ca', where a' is also in A. We *need not have* $a = a'$*!* Let's emphasize that, changing c to g for an arbitrary group G:

The operation defined above is well-defined
iff
for every $g \in G$ *and for every* $a \in A$
there exists $a' \in A$ *such that* $ga = a'g$.

In terms of sets, for every $g \in G$ and every $a \in A$, there exists $a' \in A$ such that $ga = a'g$. Here $ga \in gA$ is arbitrary, so $gA \subseteq Ag$. The other direction must also be true, so $gA \supseteq Ag$. In other words,

The operation defined above is well-defined
iff $gA = Ag$ *for all* $g \in G$.

Definition 4·125. Let $A < G$. If

$$gA = Ag$$

for every $g \in G$, then A is a **normal subgroup** of G.

Since normal subgroups partition a group into a new group, the same way ideals partition a ring into a new ring, let's "promote" them to having the same notation.

Notation 4·126. We write $A \lhd G$ to indicate that A is a normal subgroup of G.

Question 4·127 .
Show that for any group G, $\{я\} \lhd G$ and $G \lhd G$.

Although we have outlined the argument above, we should show explicitly that if A is a normal subgroup, then the operation proposed for G/A is indeed well-defined.

Lemma 4·128. *Let $A < G$. Then (CO1) implies (CO2).*

(CO1) $A \lhd G$.

(CO2) Let $X, Y \in G/A$ and $x, y \in G$ such that $X = xA$ and $Y = yA$. The operation $*$ on G/A defined by

$$X * Y = (xy)\,A$$

is well-defined for all $x, y \in G$.

Proof. Let $W, X, Y, Z \in G/A$ and choose $w, x, y, z \in G$ such that $W = wA$, $X = xA$, $Y = yA$, and $Z = zA$. To show that the operation is well-defined, we must show that if $W = X$ and $Y = Z$, then $WY = XZ$ regardless of the values of w, x, y, or z. Assume therefore that $W = X$ and $Y = Z$. By substitution, $wA = xA$ and $yA = zA$. By Lemma 4·103(CE3), $w^{-1}x \in A$ and $y^{-1}z \in A$.

Since WY and XZ are sets, showing that they are equal requires us to show that each is a subset of the other. First we show that $WY \subseteq XZ$. To do this, let $t \in WY = (wy)\,A$. By definition of a coset, $t = (wy)\,a$ for some $a \in A$. What we will do now is rewrite t by

- using the fact that A is normal to move some element of a left, then right, through the representation of t; and
- using the fact that $W = X$ and $Y = Z$ to rewrite products of the form $w\bar{\alpha}$ as $x\hat{\alpha}$ and $y\dot{\alpha}$ as $z\ddot{\alpha}$, where $\bar{\alpha}, \hat{\alpha}, \dot{\alpha}, \ddot{\alpha} \in A$.

How, precisely? By the associative property, $t = w\,(ya)$. By definition of a coset, $ya \in yA$. By hypothesis, A is normal, so $yA = Ay$; thus, $ya \in Ay$. By definition of a coset, there exists $\bar{\alpha} \in A$ such that $ya = \bar{\alpha}y$. By substitution, $t = w\,(\bar{\alpha}y)$. By the associative property, $t = (w\bar{\alpha})\,y$. By definition of a coset, $w\bar{\alpha} \in wA$. By hypothesis, A is normal, so $wA = Aw$. Thus $w\bar{\alpha} \in Aw$. By hypothesis, $W = X$; that is, $wA = xA$. Thus $w\bar{\alpha} \in xA$, and by definition of a coset, $w\bar{\alpha} = x\hat{\alpha}$ for some $\hat{\alpha} \in A$. By substitution, $t = (x\hat{\alpha})\,y$. The associative property again gives us $t = x\,(\hat{\alpha}y)$; since A is normal we can write $\hat{\alpha}y = y\dot{\alpha}$ for some $\dot{\alpha} \in A$. Hence $t = x\,(y\dot{\alpha})$. Now,

$$y\dot{\alpha} \in yA = Y = Z = zA,$$

so we can write $y\dot{\alpha} = z\ddot{\alpha}$ for some $\ddot{\alpha} \in A$. By substitution and the definition of coset arithmetic,

$$t = x\,(z\ddot{\alpha}) = (xz)\,\ddot{\alpha} \in (xz)\,A = (xA)\,(zA) = XZ.$$

Since t was arbitrary in WY, we have shown that $WY \subseteq XZ$. A similar argument shows that $WY \supseteq XZ$; thus $WY = XZ$ and the operation is well-defined. □

An easy generalization of the argument of Example 4·122 shows the following Theorem.

Theorem 4·129. *Let G be an abelian group, and $H < G$. Then $H \lhd G$.*

Question 4·130 .
Prove Theorem 4·129.

Question 4·131 .
Explain why every subgroup of $D_m(\mathbb{R})$ is normal.

Question 4·132 .
Show that Q_8 is not a normal subgroup of $\mathrm{GL}_m(\mathbb{C})$.

Question 4·133 .
Let G be a group, and $A < G$. Suppose that $|G/A| = 2$; that is, the subgroup A partitions G into precisely two left cosets. Show that:

- $A \lhd G$; and
- G/A is abelian.

We said before that we don't need an abelian group to have a normal subgroup. Here's a *great* example.

Example 4·134. Let

$$A_3 = \left\{\iota, \rho, \rho^2\right\} < D_3.$$

We call A_3 the **alternating group** on three elements. We claim that $A_3 \lhd D_3$. Indeed,

σ	σA_3	$A_3\sigma$
ι	A_3	A_3
ρ	A_3	A_3
ρ^2	A_3	A_3
φ	$\varphi A_3 = \{\varphi, \varphi\rho, \varphi\rho^2\} = A_3\varphi$	$A_3\varphi = \varphi A_3$
$\rho\varphi$	$\{\rho\varphi, (\rho\varphi)\,\rho, (\rho\varphi)\,\rho^2\} = \varphi A_3$	φA_3
$\rho^2\varphi$	$\{\rho^2\varphi, (\rho^2\varphi)\,\rho, (\rho^2\varphi)\,\rho^2\} = \varphi A_3$	φA_3

We have left out some details, though we also computed this table in Example 4·99, calling the subgroup K instead of A_3. Check the computation carefully, using extensively the fact that $\varphi\rho = \rho^2\varphi$.

Quotient groups

The set of cosets of a normal subgroup is, as desired, a group.

Theorem 4·135. *Let G be a group. If $A \lhd G$, then G/A is a group.*

Proof. Assume $A \lhd G$. By Lemma 4·128, the operation is well-defined, so it remains to show that G/A satisfies the properties of a group.

(closure) Closure follows from the fact that multiplication of cosets is well-defined when $A \lhd G$, as shown in Lemma 4·128: Let $X, Y \in G/A$, and choose $g_1, g_2 \in G$ such that $X = g_1A$ and $Y = g_2A$. By definition of coset multiplication, $XY = (g_1A)(g_2A) = (g_1g_2)A \in G/A$. Since X, Y were arbitrary in G/A, coset multiplication is closed.

(associativity) The associative property of G/A follows from the associative property of G. Let $X, Y, Z \in G/A$; choose $g_1, g_2, g_3 \in G$ such that $X = g_1A$, $Y = g_2A$, and $Z = g_3A$. Then

$$(XY)Z = [(g_1A)(g_2A)](g_3A).$$

By definition of coset multiplication,

$$(XY)Z = ((g_1g_2)A)(g_3A).$$

By the definition of coset multiplication,

$$(XY)Z = ((g_1g_2)g_3)A.$$

(Note the parentheses grouping g_1g_2.) Now apply the associative property of G and reverse the previous steps to obtain

$$\begin{aligned}(XY)Z &= (g_1(g_2g_3))A \\ &= (g_1A)((g_2g_3)A) \\ &= (g_1A)[(g_2A)(g_3A)] \\ &= X(YZ).\end{aligned}$$

Since X, Y, Z were arbitrary in G/A, coset multiplication is associative.

(identity) We claim that the identity of G/A is A itself. Let $X \in G/A$, and choose $g \in G$ such that $X = gA$. Since $я \in A$, Lemma 4·103 on page 152 implies that $A = яA$, so

$$XA = (gA)(яA) = (gя)A = gA = X.$$

Since X was arbitrary in G/A and $XA = X$, A is the identity of G/A.

(inverse) Let $X \in G/A$. Choose $g \in G$ such that $X = gA$, and let $Y = g^{-1}A$. We claim that $Y = X^{-1}$. By applying substitution and the operation on cosets,

$$XY = (gA)\left(g^{-1}A\right) = \left(gg^{-1}\right)A = яA = A.$$

Hence X has an inverse in G/A. Since X was arbitrary in G/A, every element of G/A has an inverse.

We have shown that G/A satisfies the properties of a group. □

Definition 4·136. Let G be a group, and $A \lhd G$. Then G/A is **the quotient group of G with respect to A**, also called G **mod** A.

Normally we say "the quotient group" rather than "the quotient group of G with respect to A."

Example 4·137. Since A_3 is a normal subgroup of D_3, D_3/A_3 is a group. By Lagrange's Theorem, it has $6/3 = 2$ elements. The Cayley table is

$\circ$	A_3	φA_3
A_3	A_3	φA_3
φA_3	φA_3	A_3

We meet an important quotient group in Section 4·5.

Question 4·138 .
Prove the following generalization of Theorem 4·87: If G is a cyclic group and $A \lhd G$, then G/A is cyclic.

Question 4·139 .
Recall from Question 4.17 that if $d \mid n$, then $\Omega_d < \Omega_n$.

(a) Explain how we know that, in fact, $\Omega_d \lhd \Omega_n$.

(b) Does the quotient group Ω_8/Ω_2 have the same structure as the Klein 4-group, or as the Cyclic group of order 4?

Question 4·140 .
In Question 4.95, you computed the left and right cosets of $\langle \mathbf{j} \rangle$ in Q_8. Is $\langle \mathbf{j} \rangle$ a normal subgroup of Q_8? If so, compute the Cayley table of $Q_8/\langle \mathbf{j} \rangle$.

Question 4·141 .
Let $H = \langle \mathbf{i} \rangle < Q_8$.

(a) Show that $H \lhd Q_8$ by computing all the cosets of H.

(b) Compute the Cayley table of Q_8/H.

Question 4·142 .
Recall the subgroup L of $\mathbb{R}^2$ from Questions 4.14 on page 123 and 4.98 on page 150.

(a) Explain how we know that $L \lhd \mathbb{R}^2$ *without* checking $p + L = L + p$ for any $p \in \mathbb{R}^2$.

(b) Sketch two elements of $\mathbb{R}^2/L$ and show their sum.

Conjugation

Another way to show a subgroup is normal involves rephrasing the idea of equality between left and right cosets. This is tied into an important operation, called conjugation.

Definition 4·143. Let G be a group, $g \in G$, and $H < G$. Define the **conjugation** of H by g as

$$gHg^{-1} = \left\{ghg^{-1} : h \in H\right\}.$$

Theorem 4·144. *$H \lhd G$ if and only if $H = gHg^{-1}$ for all $g \in G$.*

Let G be a group, and $H < G$. **Claim:** $H \lhd G$ if and only if $H = gHg^{-1}$ for all $g \in G$.
Proof:

1. First, we show that if $H \lhd G$, then _____.
   (a) Assume _____.
   (b) By definition of normal, _____.
   (c) Let g_____.
   (d) We first show that $H \subseteq gHg^{-1}$.
      i. Let h_____.
      ii. By 1b, $hg \in$_____.
      iii. By definition, there exists $h' \in H$ such that $hg =$_____.
      iv. Multiply both sides on the right by g^{-1} to see that $h =$_____.
      v. By _____, $h \in gHg^{-1}$.
      vi. Since h was arbitrary, _____.
   (e) Now we show that $H \supseteq gHg^{-1}$.
      i. Let $x \in$_____.
      ii. By _____, $x = ghg^{-1}$ for some $h \in H$.
      iii. By _____, $gh \in Hg$.
      iv. By _____, there exists $h' \in H$ such that $gh = h'g$.
      v. By _____, $x = (h'g)\, g^{-1}$.
      vi. By _____, $x = h'$.
      vii. By _____, $x \in H$.
      viii. Since x was arbitrary, _____.
   (f) We have shown that $H \subseteq gHg^{-1}$ and $H \supseteq gHg^{-1}$. Thus, _____.

2. Now, we show _____: that is, if $H = gHg^{-1}$ for all $g \in G$, then $H \lhd G$.
   (a) Assume _____.
   (b) First, we show that $gH \subseteq Hg$.
      i. Let $x \in$_____.
      ii. By _____, there exists $h \in H$ such that $x = gh$.
      iii. By _____, $g^{-1}x = h$.
      iv. By _____, there exists $h' \in H$ such that $h = g^{-1}h'g$. (This holds for *all* $g \in G$.)
      v. By _____, $g^{-1}x = g^{-1}h'g$.
      vi. By _____, $x = g\,(g^{-1}h'g)$.
      vii. By _____, $x = h'g$.
      viii. By _____, $x \in Hg$.
      ix. Since x was arbitrary, _____.
   (c) The proof that _____ is similar.
   (d) We have show that _____. Thus, $gH = Hg$.

Question 4·145 .
Prove Theorem 4·144 by filling in each blank of Figure 4·8 with the appropriate justification or statement.[2]

Example 4·146. We posed the question of whether $\mathrm{SO}_n(\mathbb{R}) \lhd \mathrm{O}_n(\mathbb{R})$. We claim that it is. To see why, let $M \in \mathrm{SO}_n(\mathbb{R})$ and $A \in \mathrm{O}_n(\mathbb{R})$. By properties of determinants,

$$\det\left(AMA^{-1}\right) = \det A \cdot \det M \cdot \det A^{-1} = \det A \cdot 1 \cdot (\det A)^{-1} = 1.$$

By definition, $AMA^{-1} \in \mathrm{SO}_n(\mathbb{R})$, regardless of the choice of A and M. Hence, $A \cdot \mathrm{SO}_n(\mathbb{R}) \cdot A^{-1} \subseteq \mathrm{SO}_n(\mathbb{R})$ for all $A \in \mathrm{O}_n(\mathbb{R})$.

Conversely, let $B = A^{-1}MA$; an argument similar to the one above shows that $B \in \mathrm{SO}_n(\mathbb{R})$, and substitution gives us $M = ABA^{-1}$, so that $M \in A \cdot \mathrm{SO}_n(\mathbb{R}) \cdot A^{-1}$, regardless of the choice of A and M. Hence, $A \cdot \mathrm{SO}_n(\mathbb{R}) \cdot A^{-1} \supseteq \mathrm{SO}_n(\mathbb{R})$, and the two are equal. By Theorem 4·144, $\mathrm{SO}_n(\mathbb{R}) \lhd \mathrm{O}_n(\mathbb{R})$.

Example 4·147. On the other hand, we can also use conjugation to show easily that $\mathrm{O}_2(\mathbb{R})$ is not a normal subgroup of $\mathrm{GL}_2(\mathbb{R})$. Why not? Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}) \quad \text{and} \quad M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathrm{O}_2(\mathbb{R}); \quad \text{notice that} \quad A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

If we can show that $AMA^{-1} \notin \mathrm{O}_2(\mathbb{R})$, then we would know that $A \cdot \mathrm{O}_2(\mathbb{R}) \cdot A^{-1} \not\subseteq \mathrm{O}_2(\mathbb{R})$, showing that $\mathrm{O}_2(\mathbb{R})$ is not normal. In fact,

$$AMA^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix},$$

and its inverse is *itself,* not its transpose, so in fact $AMA^{-1} \notin \mathrm{O}_2(\mathbb{R})$.

Question 4·148 .
In Question 4.95, you computed the left cosets of $\langle -\mathbf{1} \rangle$ in Q_8.

(a) Show that $\langle -\mathbf{1} \rangle$ is normal.

(b) Compute the Cayley table of $Q_8/\langle -\mathbf{1} \rangle$.

(c) The quotient group of $Q_8/\langle -\mathbf{1} \rangle$ is isomorphic to a group with which you are familiar. Which one?

Question 4·149 .
Fill in every blank of Figure 4.149 with the appropriate justification or statement.

[2] Certain texts define a normal subgroup this way; that is, a subgroup H is normal if every conjugate of H is precisely H. They then prove that in this case, any left coset equals the corresponding right coset.

Let G be a group. The **centralizer** of G is

$$Z(G) = \{g \in G : \ xg = gx \ \forall x \in G\}.$$

Claim: $Z(G) \lhd G$.
Proof:

1. First, we must show that $Z(G) < G$.
   (a) Let g, h, x____.
   (b) By ____, $xg = gx$ and $xh = hx$.
   (c) By ____, $xh^{-1} = h^{-1}x$.
   (d) By ____, $h^{-1} \in Z(G)$.
   (e) By the associative property and the definition of $Z(G)$, $(gh^{-1})\,x =$____=____=...= $x\,(gh^{-1})$.
   (Fill in more blanks as needed.)
   (f) By ____, $gh^{-1} \in Z(G)$.
   (g) By ____, $Z(G) < G$.
2. Now, we show that $Z(G)$ is normal.
   (a) Let x____.
   (b) First we show that $xZ(G) \subseteq Z(G)\,x$.
      i. Let y____.
      ii. By definition of cosets, there exists $g \in Z(G)$ such that $y =$____.
      iii. By definition of $z(G)$, ____.
      iv. By definition of ____, $y \in Z(G)\,x$.
      v. By ____, $xZ(G) \subseteq Z(G)\,x$.
   (c) A similar argument shows that ____.
   (d) By definition, ____. That is, $Z(G)$ is normal.

Figure 4·9: Material for Question 4.149

Question 4·150 .

Let G be a group, and $H < G$. Define the **normalizer** of H as

$$N_G(H) = \{g \in G : \ gH = Hg\}.$$

Show that $H \lhd N_G(H)$.

Question 4·151 .

Recall from Question 2.30 on page 53 the commutator of two elements of a group. Let $[G, G]$ denote the intersection of all subgroups of G that contain $[x, y]$ for all $x, y \in G$.

(a) Compute $[D_3, D_3]$.

(b) Compute $[Q_8, Q_8]$.

(c) Show that $[G, G] < G$.

(d) Fill in each blank of Figure 4·8 with the appropriate justification or statement.

Definition 4·152. We call $[G, G]$ the **commutator subgroup** of G, and make use of it later.

Claim: For any group G, $[G,G]$ is a normal subgroup of G.
Proof:

1. Let _____.
2. We will use Question 4.145 to show that $[G,G]$ is normal. Let $g \in$_____.
3. First we show that $[G,G] \subseteq g\,[G,G]\,g^{-1}$. Let $h \in [G,G]$.
   (a) We need to show that $h \in g\,[G,G]\,g^{-1}$. It will suffice to show that this is true if h has the simpler form $h = [x,y]$, since _____. Thus, choose $x, y \in G$ such that $h = [x,y]$.
   (b) By _____, $h = x^{-1}y^{-1}xy$.
   (c) By _____, $h = яx^{-1}яy^{-1}яxяyя$.
   (d) By _____, $h = (gg^{-1})\,x^{-1}\,(gg^{-1})\,y^{-1}\,(gg^{-1})\,x\,(gg^{-1})\,y\,(gg^{-1})$.
   (e) By _____, $h = g\,(g^{-1}x^{-1}g)\,(g^{-1}y^{-1}g)\,(g^{-1}xg)\,(g^{-1}yg)\,g^{-1}$.
   (f) By _____, $h = g\,(x^{-1})^{g^{-1}}\,(y^{-1})^{g^{-1}}\left(x^{g^{-1}}\right)\left(y^{g^{-1}}\right)g^{-1}$.
   (g) By Question 2.30 on page 53(c), $h =$_____.
   (h) By definition of the commutator, $h =$_____.
   (i) By _____, $h \in g\,[G,G]\,g^{-1}$.
   (j) Since _____, $[G,G] \subseteq g\,[G,G]\,g^{-1}$.
4. Conversely, we show that $[G,G] \supseteq g\,[G,G]\,g^{-1}$. Let $h \in g\,[G,G]\,g^{-1}$.
   (a) We need to show that $h \in [G,G]$. It will suffice to show this is true if h has the simpler form $h = g\,[x,y]\,g^{-1}$, since _____. Thus, choose $x, y \in G$ such that $h = g\,[x,y]\,g^{-1}$.
   (b) By _____, $h = [x,y]^g$.
   (c) By _____, $h = [x^g, y^g]$.
   (d) By _____, $h \in [G,G]$.
   (e) Since _____, $[G,G] \supseteq g\,[G,G]\,g^{-1}$.
5. We have shown that $[G,G] \subseteq g\,[G,G]\,g^{-1}$ and $[G,G] \supseteq g\,[G,G]\,g^{-1}$. By _____, $[G,G] = g\,[G,G]\,g^{-1}$.

Figure 4·10: Material for Question 4.151

4·9 The Isomorphism Theorem

This section describes an important relationship between a subgroup $A < G$ that has a special relationship to a homomorphism, and the image of the quotient group $f(G/A)$. It builds on an important property of the kernel of a group or ring homomorphism.

Fact 4·153. *(A) Let $f : G \to H$ be a homomorphism of groups. Then* $\ker f$ *is a normal subgroup of G.*

(B) Let $\varphi : R \to S$ be a homomorphism of rings. Then $\ker \varphi$ *is an ideal of R.*

Why? (A) First we show that $\ker f$ is a subgroup of G. Let $x, y \in \ker f$; by definition, $f(x) = \mathfrak{s}_H = f(y)$. Multiply both sides by $f(y)^{-1}$ and we have $f(x) f(y)^{-1} = \mathfrak{s}_H$. Properties of homomorphisms show us that $f(xy^{-1}) = \mathfrak{s}_H$. By definition of the kernel, $xy^{-1} \in \ker f$.

We still have to show that $K = \ker f$ is a *normal* subgroup of G. We do this by conjugation (Theorem 4·144); that is, we show that for any $g \in G$, $gKg^{-1} = K$. To see why, let $x \in gKg^{-1}$; by definition, $x = gkg^{-1}$ for some $k \in K$ and $f(k) = \mathfrak{s}_H$. Apply properties of homomorphisms to see that

$$f(x) = f\left(gkg^{-1}\right) = f(g) f(k) f\left(g^{-1}\right) = f(g)\, \mathfrak{s}_H f(g)^{-1} = \mathfrak{s}_H.$$

So $x \in \ker f = K$; it was arbitrary in gKg^{-1}, so $gKg^{-1} \subseteq K$. We also have to show that $gKg^{-1} \supseteq K$, so let $k \in K$. Let $x = g^{-1}kg$; by an argument similar to that of the previous paragraph, $x \in K$. Hence

$$k = \mathfrak{s}_G k \mathfrak{s}_G = \left(gg^{-1}\right) k \left(gg^{-1}\right) = g\left(g^{-1}kg\right) g^{-1} = gxg \in gKg^{-1},$$

as claimed. Since k was arbitrary in K, $gKg^{-1} \supseteq K$, as claimed. We have shown that each set is a subset of the other, so $gKg^{-1} = K$. Since g was arbitrary in G, Theorem 4·144 tells us $K = \ker f$ is a normal subgroup of G.

(B) To show that $\ker \varphi$ is an ideal, we need only show that it absorbs multiplication, since (A) has already shown that it is a subgroup for the additive group of R. To that end, let $r \in R$ and $k \in \ker \varphi$. By properties of a homomorphism, $\varphi(rk) = \varphi(r)\varphi(k) = \varphi(r) \cdot 0 = 0$, so $rk \in \ker \varphi$. Since r was arbitrary in R, $\ker \varphi$ absorbs multiplication by *all* elements of R; it is thus an ideal. □

First, an example.

Motivating example

Example 4·154. Recall $A_3 = \{\iota, \rho, \rho^2\} \lhd D_3$ from Example 4·134. We saw that D_3/A_3 has only two elements, so it must be isomorphic to any group of two elements. First we show this explicitly: Let $\mu : D_3/A_3 \to \mathbb{Z}_2$ by

$$\mu(X) = \begin{cases} 0, & X = A_3; \\ 1, & \text{otherwise.} \end{cases}$$

Is μ a homomorphism? Recall that A_3 is the identity element of D_3/A_3, so for any $X \in D_3/A_3$

$$\mu(X \cdot A_3) = \mu(X) = \mu(X) + 0 = \mu(X) + \mu(A_3).$$

This verifies the homomorphism property for all products in the Cayley table of D_3/A_3 except $(\varphi A_3) \cdot (\varphi A_3)$, which is easy to check:

$$\mu((\varphi A_3) \cdot (\varphi A_3)) = \mu(A_3) = 0 = 1 + 1 = \mu(\varphi A_3) + \mu(\varphi A_3).$$

Hence μ is a homomorphism. The property of isomorphism follows from the facts that

- $\mu(A_3) \neq \mu(\varphi A_3)$, so μ is one-to-one, and
- both 0 and 1 have preimages, so μ is onto.

Notice further that $\ker \mu = A_3$.

Something subtle is at work here. Let $f : D_3 \to \mathbb{Z}_2$ by

$$f(x) = \begin{cases} 0, & x \in A_3; \\ 1, & \text{otherwise.} \end{cases}$$

Is f a homomorphism? The elements of A_3 are ι, ρ, and ρ^2; f maps these elements to zero, and the other three elements of D_3 to 1. Let $x, y \in D_3$ and consider the various cases:

Case 1. Suppose first that $x, y \in A_3$. Since A_3 is a group, closure implies that $xy \in A_3$. Thus

$$f(xy) = 0 = 0 + 0 = f(x) + f(y).$$

Case 2. Next, suppose that $x \in A_3$ and $y \notin A_3$. Since A_3 is a group, closure implies that $xy \notin A_3$. (Otherwise $xy = z$ for some $z \in A_3$, and multiplication by the inverse implies that $y = x^{-1}z \in A_3$, a contradiction.) Thus

$$f(xy) = 1 = 0 + 1 = f(x) + f(y).$$

Case 3. If $x \notin A_3$ and $y \in A_3$, then a similar argument shows that $f(xy) = f(x) + f(y)$.

Case 4. Finally, suppose $x, y \notin A_3$. Inspection of the Cayley table of D_3 (Question 3.117 on page 109) shows that $xy \in A_3$. Hence

$$f(xy) = 0 = 1 + 1 = f(x) + f(y).$$

We have shown that f is a homomorphism from D_3 to $\mathbb{Z}_2$. Again, $\ker f = A_3$.

In addition, consider the function $\eta : D_3 \to D_3/A_3$ by

$$\eta(x) = \begin{cases} A_3, & x \in A_3; \\ \varphi A_3, & \text{otherwise.} \end{cases}$$

It is easy to show that this is a homomorphism; we do so presently.

Now comes the important observation: Look at the composition function $\eta \circ \mu$ whose domain is D_3 and whose range is $\mathbb{Z}_2$:

$$\begin{aligned}
(\mu \circ \eta)(\iota) &= \mu(\eta(\iota)) = \mu(A_3) = 0;\\
(\mu \circ \eta)(\rho) &= \mu(\eta(\rho)) = \mu(A_3) = 0;\\
(\mu \circ \eta)\left(\rho^2\right) &= \mu\left(\eta\left(\rho^2\right)\right) = \mu(A_3) = 0;\\
(\mu \circ \eta)(\varphi) &= \mu(\eta(\varphi)) = \mu(\varphi A_3) = 1;\\
(\mu \circ \eta)(\rho\varphi) &= \mu(\eta(\rho\varphi)) = \mu(\varphi A_3) = 1;\\
(\mu \circ \eta)\left(\rho^2\varphi\right) &= \mu\left(\eta\left(\rho^2\varphi\right)\right) = \mu(\varphi A_3) = 1.
\end{aligned}$$

We have

$$(\mu \circ \eta)(x) = \begin{cases} 0, & x \in A_3;\\ 1, & \text{otherwise,} \end{cases}$$

or in other words

$$\mu \circ \eta = f.$$

In words, f is the composition of a "natural" mapping between D_3 and D_3/A_3, and the isomorphism from D_3/A_3 to $\mathbb{Z}_2$. But another way of looking at this is that the isomorphism μ is related to f and the "natural" homomorphism.

The Isomorphism Theorem

This remarkable correspondence can make it easier to study quotient groups G/A:

- find a group H that is "easy" to work with; and
- find a homomorphism $f : G \to H$ such that
  - $f(g) = \mathfrak{s}_H$ for all $g \in A$, and
  - $f(g) \neq \mathfrak{s}_H$ for all $g \notin A$.

If we can do this, then $H \cong G/A$ and studying G/A is equivalent to studying H.

The reverse is also true: suppose that a group G and its quotient groups are relatively easy to study, whereas another group H is difficult. The isomorphism theorem helps us identify a quotient group G/A that is isomorphic to H, making it easier to study.

Another advantage, which we use later in the course, is that computation in G can be difficult or even impossible, while computation in G/A can be quite easy. This turns out to be the case with $\mathbb{Z}$ when the coefficients grow too large; we will work in $\mathbb{Z}_p$ for several values of p, and reconstruct the correct answers.

We need to formalize this observation in a theorem, but first we have to confirm something that we claimed earlier:

Lemma 4·155. *Let G be a group and $A \lhd G$. The function $\eta : G \to G/A$ by*

$$\eta(g) = gA$$

is a homomorphism.

Question 4·156 .
Prove Lemma 4·155.

Question 4·157 .
Use Question 4.23 to explain why $\Omega_2 \cong \mathrm{O}(n)/\mathrm{SO}(n)$.

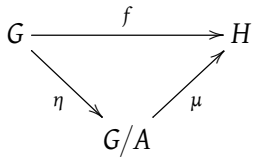
Definition 4·158. We call the homomorphism η of Lemma 4·155 the **natural homomorphism** from G to G/A.

What's special about A_3 in the example that began this section? Of course, A_3 is a normal subgroup of D_3, but something you might not have noticed is that f sent all its elements to the identity of $\mathbb{Z}_2$.

We use this to formalize the observation of Example 4·154.

Theorem 4·159 (The Isomorphism Theorem)**.** *Let G and H be groups, $f : G \to H$ a homomorphism that is onto, and $\ker f = A$. Then $G/A \cong H$, and the isomorphism $\mu : G/A \to H$ satisfies $f = \mu \circ \eta$, where $\eta : G \to G/A$ is the natural homomorphism.*

We can illustrate Theorem 4·159 by the following diagram:



The idea is that "the diagram commutes", or $f = \mu \circ \eta$.

Proof. We are given G, H, f and A. Define $\mu : G/A \to H$ in the following way:

$$\mu(X) = f(g), \text{ where } X = gA.$$

We claim that μ is an isomorphism from G/A to H, and moreover that $f = \mu \circ \eta$.

Since the domain of μ consists of cosets which may have different representations, we must show first that μ is well-defined. Suppose that $X \in G/A$ has two representations $X = gA = g'A$ where $g, g' \in G$ and $g \neq g'$. We need to show that $\mu(gA) = \mu(g'A)$. From Lemma 4·103(CE3), we know that $g^{-1}g' \in A$, so there exists $a \in A$ such that $g^{-1}g' = a$, so $g' = ga$. Applying the definition of μ and the homomorphism property,

$$\mu(g'A) = f(g') = f(ga) = f(g)f(a).$$

Recall that $a \in A = \ker f$, so $f(a) = \mathfrak{R}_H$. Substitution gives

$$\mu(g'A) = f(g) \cdot \mathfrak{R}_H = f(g) = \mu(gA).$$

Hence $\mu(g'A) = \mu(gA)$ and $\mu(X)$ is well-defined.

Is μ a homomorphism? Let $X, Y \in G/A$; we can represent $X = gA$ and $Y = g'A$ for some $g, g' \in G$. We see that

$$\begin{aligned}
\mu(XY) &= \mu((gA)(g'A)) && \text{(substitution)} \\
&= \mu((gg')A) && \text{(coset multiplication)} \\
&= f(gg') && \text{(definition of } \mu) \\
&= f(g) f(g') && \text{(homomorphism)} \\
&= \mu(gA)\mu(g'A). && \text{(definiition of } \mu)
\end{aligned}$$

Thus μ is a homomorphism.

Is μ one-to-one? Let $X, Y \in G/A$ and assume that $\mu(X) = \mu(Y)$. Represent $X = gA$ and $Y = g'A$ for some $g, g' \in G$; we see that

$$\begin{aligned}
f\left(g^{-1}g'\right) &= f\left(g^{-1}\right) f(g') && \text{(homomorphism)} \\
&= f(g)^{-1} f(g') && \text{(homomorphism)} \\
&= \mu(gA)^{-1} \mu(g'A) && \text{(definition of } \mu) \\
&= \mu(X)^{-1} \mu(Y) && \text{(substitution)} \\
&= \mu(Y)^{-1} \mu(Y) && \text{(substitution)} \\
&= \mathfrak{R}_H, && \text{(inverses)}
\end{aligned}$$

so $g^{-1}g' \in \ker f$. By hypothesis, $\ker f = A$, so $g^{-1}g' \in A$. Lemma 4·103(CE3) now tells us that $gA = g'A$, so $X = Y$. Thus μ is one-to-one.

Is μ onto? Let $h \in H$; we need to find an element $X \in G/A$ such that $\mu(X) = h$. By hypothesis, f is onto, so there exists $g \in G$ such that $f(g) = h$. By definition of μ and substitution,

$$\mu(gA) = f(g) = h,$$

so μ is onto.

We have shown that μ is an isomorphism; we still have to show that $f = \mu \circ \eta$, but the definition of μ makes this trivial: for any $g \in G$,

$$(\mu \circ \eta)(g) = \mu(\eta(g)) = \mu(gA) = f(g).$$

□

Question 4·160 .

Recall the normal subgroup L of $\mathbb{R}^2$ from Questions 4.14, 4.98, and 4.142 on pages 123, 150, and 165, respectively. In Question 4.14 on page 123 you found an explicit isomorphism $L \cong \mathbb{R}$.

(a) Use the Isomorphism Theorem to find an isomorphism $\mathbb{R}^2/L \cong \mathbb{R}$.

(b) Argue from this that $\mathbb{R}^2/\mathbb{R} \cong \mathbb{R}$.

(c) Describe geometrically how the cosets of $\mathbb{R}^2/L$ are mapped to elements of $\mathbb{R}$.

Question 4·161 .

Recall the normal subgroup $\langle -\mathbf{1} \rangle$ of Q_8 from Question 4.148 on page 168.

(a) Use Lagrange's Theorem to explain why $Q_8/\langle -\mathbf{1} \rangle$ has order 4.

(b) We know from Question 2.49 on page 61 that there are only two groups of order 4, the Klein 4-group and the cyclic group of order 4, which we can represent by $\mathbb{Z}_4$. Use the Isomorphism Theorem to determine which of these groups is isomorphic to $Q_8/\langle -\mathbf{1} \rangle$.

Question 4·162 .

Recall the kernel of a homomorphism, and that group homomorphisms are also monoid homomorphisms. These two definitions do not look the same, but in fact, one generalizes the other.

(a) Show that if $x \in G$ is in the kernel of a group homomorphism $f : G \to H$ if and only $(x, e) \in \ker f$ when we view f as a monoid homomorphism.

(b) Show that $x \in G$ is in the kernel of a group homomorphism $f : G \to H$ if and only if we can find $y, z \in G$ such that $f(y) = f(z)$ and $y^{-1}z = x$.

Question 4·163 .

Fill in each blank of Figure 4·11 with the appropriate justification or statement.

Let G and H be groups, and $A \lhd G$.
Claim: If $G/A \cong H$, then there exists a homomorphism $\varphi : G \to H$ such that $\ker \varphi = A$.

1. Assume _____.
2. By hypothesis, there exists f_____.
3. Let $\eta : G \to G/A$ be the natural homomorphism. Define $\varphi : G \to H$ by $\varphi(g) =$_____.
4. By _____, φ is a homomorphism.
5. We claim that $A \subseteq \ker \varphi$. To see why,
   (a) By _____, the identity of G/A is A.
   (b) By _____, $f(A) = \mathfrak{R}_H$.
   (c) Let $a \in A$. By definition of the natural homomorphism, $\eta(a) =$_____.
   (d) By _____, $f(\eta(a)) = \mathfrak{R}_H$.
   (e) By _____, $\varphi(a) = \mathfrak{R}_H$.
   (f) Since _____, $A \subseteq \ker \varphi$.
6. We further claim that $A \supseteq \ker \varphi$. To see why,
   (a) Let $g \in G \backslash A$. By definition of the natural homomorphism, $\varphi(g) \neq$_____.
   (b) By _____, $f(\eta(g)) \neq \mathfrak{R}_H$.
   (c) By _____, $\varphi(g) \neq \mathfrak{R}_H$.
   (d) By _____, $g \notin \ker \varphi$.
   (e) Since g was arbitrary in $G \backslash A$, _____.
7. We have shown that $A \subseteq \ker \varphi$ and $A \supseteq \ker \varphi$. By _____, $A = \ker \varphi$.

Figure 4·11: Material for Question 4.163

Chapter 5

Applications to elementary number theory

This text tends to focus on algebra as a study of polynomials, but algebra exhibits an important mark of a profound subject, in that its ideas pop up in many other places. One of these is number theory, which is closely intertwined with algebra; each can explain results and motivate new questions in the other. They also share a common spirit of exploration; it is not uncommon to find them grouped together in departments conferences, or research agencies.

This chapter introduces several of these relationships. Section 5·1 fills some background with two of the most important tools in computational algebra and number theory. The first is a fundamental definition; the second, a fundamental algorithm. Both recur throughout the chapter, and later in the notes. Section 5·2 moves us to our first application of group theory, the *Chinese Remainder Theorem*, used thousands of years ago for the task of counting the number of soldiers who survived a battle. We will use it to explain a neat card trick that you can teach to grade-school children (though they may not understand why it works).

The rest of the chapter moves us toward Section 5·6, the RSA cryptographic scheme, a major component of internet communication and commerce. In Section 4·5 you learned of additive clockwork groups; in Section 5·4 you will learn of multiplicative clockwork groups. These allows us to describe in Section 5·5 the theoretical foundation of RSA, Euler's number and Euler's Theorem.

5·1 The Euclidean Algorithm

Until now, we've focused on division with remainder, extending its notion even to cosets of subgroups. Many problems care about divisibility; that is, division with remainder 0.

Common divisors

Recall that we say the integer a divides the integer b when we can find another integer x such that $ax = b$. Recall that a **common divisor of** m **and** n is an integer d that divides both numbers, and that $d \in \mathbb{N}$ is a **greatest common divisor of** m **and** n if d is a common divisor *and* any other common divisor d' satisfies $d' < d$.

Example 5·1. Common divisors of 36 and -210 are 1, 2, 3, and 6. The greatest common divisor is 6.

Do greatest common divisors always exist? We already know from Bézout's Lemma that they do, but we can prove something a little deeper, too.

Theorem 5·2. *Let $m, n \in \mathbb{Z}$, not both zero. There exists a unique greatest common divisor of m, n.*

Proof. Let D be the set of common divisors of m, n that are also in $\mathbb{N}^+$. Since 1 divides both m and n, we know that $D \neq \varnothing$. We also know that any $d \in D$ must satisfy $d \leqslant \min(|m|, |n|)$; otherwise, the remainder from the Division Algorithm would be nonzero for at least one of m, n. Hence, D is finite. Let d be the largest element of D. By definition of D, d is a common divisor; we claim that it is also the only greatest common divisor. After all, the integers are a linear ordering, so every other common divisor d' of m and n is either

- negative, so that by definition of subtraction, $d - d' \in \mathbb{N}^+$, or (by definition of $<$) $d' < d$; or,
- in D, so that (by definition of d) $d' \leqslant d$, and $d \neq d'$ implies $d' < d$.

□

Question 5·3 .
Show that any common divisor of any two integers divides the integers' greatest common divisor.

How can we compute the greatest common divisor? Common divisors are important enough that they appear in grade school, where you likely learned one way to compute the greatest common divisor of two integers: list all the divisors of each, and pick the largest one in both lists. In practice, this takes a Very Long Time™, so we need a different method. One such method was described by the ancient Greek mathematician, Euclid.

The Euclidean Algorithm

The Euclidean Algorithm. *Let $m, n \in \mathbb{Z}$. We can compute the greatest common divisor of m, n in the following way:*

1. *Let $s = m$ and $t = n$.*
2. *Repeat the following steps until $t = 0$:*
   (a) *Let q be the quotient and r the remainder after dividing s by t.*
   (b) *Assign s the current value of t.*
   (c) *Assign t the current value of r.*

The final value of s is $\gcd(m, n)$.

Algorithm 5.1 The Euclidean algorithm

inputs
  $m, n \in \mathbb{Z}$
outputs
  $\gcd(m, n)$
do
  Let $s = m$
  Let $t = n$
  while $t \neq 0$ **do**
    Let $q, r \in \mathbb{Z}$ be the result of dividing s by t
    Let $s = t$
    Let $t = r$
  return s

It is common to write algorithms in a form called *pseudocode,* and from this point we will make increasing use of this format. Algorithm 5.1 shows the Euclidean Algorithm in pseudocode. If you've seen computer programs, you'll notice that pseudocode is formatted much like most computer programs, in that it specifies inputs, outputs, and indents subtasks. Unlike computer code, pseudocode uses "ordinary" English and mathematical statements to communicate the necessary tasks. This provides two benefits:

- It is usually more intuitive to read and analyze pseudocode than computer code.
- Pseudocode is more easily "translated" into different computer languages.

Pseudocode appears often in texts on mathematical computation, so it's something you need to accustom yourself to reading and thinking about. We will use pseudocode a great deal in the remainder of these notes.

Before proving that the Euclidean algorithm gives us a correct answer, let's do an example.

Example 5·4. We compute $\gcd(36, 210)$. At the outset, let $s = 210$ and $t = 36$. Subsequently:

1. Dividing 210 by 36 gives $q = 5$ and $r = 30$. Let $s = 36$ and $t = 30$.
2. Dividing 36 by 30 gives $q = 1$ and $r = 6$. Let $s = 30$ and $t = 6$.
3. Dividing 30 by 6 gives $q = 5$ and $r = 0$. Let $s = 6$ and $t = 0$.

Now that $t = 0$, we stop, and conclude that $\gcd(36, 210) = s = 6$. This agrees with Example 5·1.

Question 5·5 .
Compute the greatest common divisor of 100 and 140 by (a) listing all divisors, then identifying the largest; and (b) the Euclidean Algorithm.

Question 5·6 .
Compute the greatest common divisor of 100 and 82 by (a) listing all divisors, then identifying the largest; and (b) the Euclidean Algorithm.

Question 5·7 .
Show that $\gcd(n, n-1) = 1$ for any integer n. The argument when $n = 0$ might be a little different.

To prove that the Euclidean algorithm generates a correct answer, we will number each remainder that we compute; so, the first remainder is r_1, the second, r_2, and so forth. We will then show that the remainders give us a chain of equalities,

$$\gcd(m, n) = \gcd(m, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{k-1}, 0),$$

where r_i is the remainder from division of the previous two integers in the chain, and r_{k-1} is the final non-zero remainder from division.

Lemma 5·8. *Let $s, t \in \mathbb{Z}$. Let q and r be the quotient and remainder, respectively, of division of s by t, as per the Division Theorem. Then* $\gcd(s, t) = \gcd(t, r)$.

Example 5·9. We can verify Lemma 5·8 using the numbers from Example 5·4. We know that $\gcd(210, 36) = 6$. The remainder from division of 210 by 36 is $r = 30$. The lemma claims that $\gcd(210, 36) = \gcd(36, 30)$, and indeed $\gcd(36, 30) = 6$.

Question 5·10 .
In Lemma 5·8 we showed that $\gcd(m, n) = \gcd(m, r)$ where r is the remainder after division of m by n. Prove the following more general statement: for all $m, n, q \in \mathbb{Z}$ $\gcd(m, n) = \gcd(n, m - qn)$.

We turn to the proof.

Proof of Lemma 5·8. Let $d = \gcd(s, t)$. First we show that d is a divisor of r. By definition, there exist $a, b \in \mathbb{Z}$ such that $s = ad$ and $t = bd$. By hypothesis, $s = qt + r$ and $0 \leqslant r < |t|$. Substitution gives us $ad = q(bd) + r$; rewriting the equation, we have

$$r = (a - qb)\, d.$$

By definition of divisibility, $d \mid r$.

Since d is a common divisor of s, t, and r, it is a common divisor of t and r. We claim that $d = \gcd(t, r)$. Let $d' = \gcd(t, r)$; since d is also a common divisor of t and r, the definition of *greatest* common divisor implies that $d \leqslant d'$. Since d' is a common divisor of t and r, the definition of divisibility again implies that there exist $x, y \in \mathbb{Z}$ such that $t = d'x$ and $r = d'y$. Substituting into the equation $s = qt + r$, we have $s = q(d'x) + d'y$; rewriting the equation, we have

$$s = (qx + y)\, d'.$$

So $d' \mid s$. We already knew that $d' \mid t$, so d' is a common divisor of s and t.

Recall that $d = \gcd(s,t)$; since d' is also a common divisor of t and r, the definition of *greatest* common divisor implies that $d' \leqslant d$. Earlier, we showed that $d \leqslant d'$. Hence $d \leqslant d' \leqslant d$, which implies that $d = d'$.

Substitution gives the desired conclusion: $\gcd(s,t) = \gcd(t,r)$. □

We can finally prove that the Euclidean algorithm gives us a correct answer. This requires two stages, necessary for any algorithm.

1. **Correctness.** *If* the algorithm terminates, we show it has computed the correct output (result).

2. **Termination.** We show the algorithm concludes its computation in finite time.

If an algorithm has finitely many instructions, how could it go continue running without end? The Euclidean algorithm holds a clue: an instruction asks us to repeat some steps "**while** $t \neq 0$." What if t never attains the value of zero? It's conceivable that its values remain positive at all times, or jump from positive to negative, skipping zero. In that case, the algorithm would continue without end.

In computation, the repetition of tasks is called a **loop**. Loops save us an enormous amount of time, but not all algorithms contain loops.

A proof of termination is needed if and only if an algorithm contains a loop.

These notes use only two kinds of loops: **for** loops and **while** loops.

- A **while** loop repeats *every* subtask as long as the expression that immediately follows it remains true. As soon as it completes a pass through the subtasks and the expression becomes false, the loop ends.

- A **for** loop works exactly like logical quantification: it applies all subtasks to each element of the set specified immediately after the word **for**. The statement "**for** $s \in S$" means to apply the subtasks to each element of the set S, and "**for** $n \in \mathbb{N}$ such that $n < 10$" means to apply the subtasks to each natural number less than 10. You will see examples of **for** loops later.

The proof of the Euclidean algorithm will identify clearly both the Correctness and Termination stages. As it depends on the Division Theorem and the Well-Ordering Principle, you may wish to review those.

Proof of The Euclidean Algorithm. We start with termination. The only repetition in the algorithm occurs in line 8. The first time we compute line 9, we compute the quotient q and remainder r of division of s by t. By the Division Theorem,

$$0 \leqslant r < |t|. \tag{5.1}$$

Denote this value of r by r_1. In the next lines we set s to t, then t to $r_1 = r$. Thanks to equation (5.1), the size of $t_{\text{new}} = r$ is smaller than that of $|s_{\text{new}}| = |t_{\text{old}}|$. (We measure "size" using

absolute value.) If $t \neq 0$, then we return to line 9 and divide s by t, again obtaining a new remainder r. Denote this value of r by r_2; by the Division Theorem, $r_2 = r < t$, so

$$0 \leqslant r_2 < r_1.$$

Proceeding in this fashion, we generate a strictly decreasing sequence of elements,

$$r_1 > r_2 > r_3 > \cdots.$$

By Fact 1·51, this sequence is finite. In other words, the algorithm terminates.

We now show that the algorithm terminates *with the correct answer*. If line 9 of the algorithm repeated a total of k times, then $r_k = 0$. Apply Lemma 5·8 repeatedly to the remainders to obtain the chain of equalities

$$\begin{aligned}
r_{k-1} = \gcd(0, r_{k-1}) &= \gcd(r_k, r_{k-1}) && \text{(definition of gcd, substitution)} \\
&= \gcd(r_{k-1}, r_{k-2}) && \text{(Lemma 5·8)} \\
&= \gcd(r_{k-2}, r_{k-3}) && \text{(Lemma 5·8)} \\
&\vdots \\
&= \gcd(r_2, r_1) && \text{(Lemma 5·8)} \\
&= \gcd(r_1, s) && \text{(substitution)} \\
&= \gcd(t, s) && \text{(substitution)} \\
&= \gcd(m, n). && \text{(substitution)}
\end{aligned}$$

The Euclidean Algorithm terminates with the correct answer. □

The Euclidean Algorithm and Bezout's Lemma

Recall Bézout's Lemma, which tells us that for any integers m and n we can find integers x and y such that

$$\gcd(m, n) = mx + ny.$$

You may have noticed that Bézout's Lemma gives us no advice on *how* to do find this expression; it merely states that we *can* do it. The proof of Bézout's Lemma isn't very helpful, either; it says to look at all the elements of a certain set, and choose the smallest. That set contains infinitely many elements; how would we know when we've found the smallest?

The Euclidean Algorithm turns out to be just the tool for the job.

The Extended Euclidean Algorithm. *Let $m, n \in \mathbb{Z}$. There exist $a, b \in \mathbb{Z}$ such that $am + bn = \gcd(m, n)$. Both a and b can be found by adapting the results from the Euclidean algorithm, using the following steps:*

- *Isolate the remainder of the penultimate division of the Euclidean Algorithm; that is, $r_{k-1} = r_{k-3} - q_{k-1}r_{k-2}$.*
- *The proof of the Euclidean Algorithm tells us that $r_{k-1} = \gcd(m, n)$, so in fact $\gcd(m, n) = r_{k-3} - q_{k-1}r_{k-2}$. We call this the* **working equation.**

Algorithm 5.2 Extended Euclidean Algorithm

inputs
$m, n \in \mathbb{N}$ such that $m > n$
outputs
$\gcd(m, n)$ and $a, b \in \mathbb{Z}$ such that $\gcd(m, n) = am + bn$
do
if $n = 0$ **then**
Let $d = m$, $a = 1$, $b = 0$
else
Let $r_0 = m$ and $r_1 = n$
Let $k = 1$
{First apply the Euclidean Algorithm}
while $r_k \neq 0$ **do**
Increment k by 1
Let q_k, r_k be the quotient and remainder from division of r_{k-2} by r_{k-1}
{Now reverse it}
Let $d = r_{k-1}$ and $p = r_{k-3} - q_{k-1}r_{k-2}$ (do not simplify p)
Decrement k by 2
while $k \geqslant 2$ **do**
Substitute $r_k = r_{k-2} - q_k r_{k-1}$ into p
Decrement k by 1
Let a be the coefficient of r_0 in p, and b be the coefficient of r_1 in p
return d, a, b

- *Working* backwards *from the previous division, until we arrive at the first,*
  - *Isolate the remainder of this division; that is,* $r_\ell = r_{\ell-2} - q_\ell r_{\ell-1}$.
  - *Find* r_ℓ *in the working equation, and replace it by* $r_{\ell-2} - q_\ell r_{\ell-1}$.

Pseudocode appears in Algorithm 5.2.

Example 5·11. Recall from Example 5·4 the computation of $\gcd(210, 36)$. The divisions gave us a series of equations:

$$210 = 5 \cdot 36 + 30 \tag{5.2}$$
$$36 = 1 \cdot 30 + 6 \tag{5.3}$$
$$30 = 5 \cdot 6 + 0.$$

We concluded from the Euclidean Algorithm that $\gcd(210, 36) = 6$. The Extended Euclidean Algorithm gives us a way to find $a, b \in \mathbb{Z}$ such that $6 = 210a + 36b$. Start by rewriting equation (5.3):

$$36 - 1 \cdot 30 = 6. \tag{5.4}$$

This looks a little like what we want, but we need 210 instead of 30. Equation (5.2) allows us to rewrite 30 in terms of 210 and 36:

$$30 = 210 - 5 \cdot 36. \tag{5.5}$$

Substituting this result into equation (5.4), we have

$$36 - 1 \cdot (210 - 5 \cdot 36) = 6 \quad \Longrightarrow \quad 6 \cdot 36 + (-1) \cdot 210 = 6.$$

We have found integers $m = 6$ and $n = -1$ such that for $a = 36$ and $b = 210$, $\gcd(a, b) = 6$.

Question 5·12 .
Compute the greatest common divisor of $m = 4343$ and $n = 4429$ by the Euclidean Algorithm. Use the Extended Euclidean Algorithm to find $a, b \in \mathbb{Z}$ that satisfy Bezout's identity.

The method we applied in Example (5·11) is what we use both to prove correctness of the algorithm, and to find a and b in general.

Proof of the Extended Euclidean Algorithm. Look back at the proof of the Euclidean algorithm to see that it computes a chain of k quotients $\{q_i\}$ and remainders $\{r_i\}$ such that

$$\begin{aligned}
m &= q_1 n + r_1 \\
n &= q_2 r_1 + r_2 \\
r_1 &= q_3 r_2 + r_3 \\
&\vdots \\
r_{k-4} &= q_{k-2} r_{k-3} + r_{k-2} \qquad (5.6) \\
r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1} \qquad (5.7) \\
r_{k-2} &= q_k r_{k-1} + 0 \\
\text{and } r_k &= \gcd(m, n).
\end{aligned}$$

Rewrite equation (5.7) as

$$r_{k-3} = q_{k-1} r_{k-2} + \gcd(m, n).$$

Solving for $\gcd(m, n)$, we have

$$r_{k-3} - q_{k-1} r_{k-2} = \gcd(m, n). \qquad (5.8)$$

Solve for r_{k-2} in equation (5.6) to obtain

$$r_{k-4} - q_{k-2} r_{k-3} = r_{k-2}.$$

Substitute this into equation (5.8) to obtain

$$\begin{aligned}
r_{k-3} - q_{k-1}(r_{k-4} - q_{k-2} r_{k-3}) &= \gcd(m, n) \\
(q_{k-1} q_{k-2} + 1) r_{k-3} - q_{k-1} r_{k-4} &= \gcd(m, n).
\end{aligned}$$

Proceeding in this fashion, we exhaust the list of equations, concluding by rewriting the first equation in the form $am + bn = \gcd(m, n)$ for some integers a, b. □

Let $m, n \in \mathbb{Z}$, $S = \{am + bn : a, b \in \mathbb{Z}\}$, and $M = S \cap \mathbb{N}$. Since M is a subset of $\mathbb{N}$, the Well-Ordering Principle implies that it has a smallest element; call it d.
Claim: $d = \gcd(m, n)$.
Proof:

1. We first claim that $\gcd(m, n)$ divides d.
   (a) By ____, we can find $a, b \in \mathbb{Z}$ such that $d = am + bn$.
   (b) By ____, $\gcd(m, n)$ divides m and n.
   (c) By ____, there exist $x, y \in \mathbb{Z}$ such that $m = x \gcd(m, n)$ and $n = y \gcd(m, n)$.
   (d) By susbtitution, ____.
   (e) Collect the common term to obtain ____.
   (f) By ____, $\gcd(m, n)$ divides d.
2. A similar argument shows that d divides $\gcd(m, n)$.
3. By ____, $d \leqslant \gcd(m, n)$ and $\gcd(m, n) \leqslant d$.
4. By ____, $d = \gcd(m, n)$.

Figure 5·1: Material for Question 5.13

Question 5·13 .
Bezout's Identity states that for any $m, n \in \mathbb{Z}$, we can find $a, b \in \mathbb{Z}$ such that $am + bn = \gcd(m, n)$.

(a) Show that the existence of $a, b, d \in \mathbb{Z}$ such that $am + bn = d$ does *not* imply $d = \gcd(m, n)$.

(b) However, not only does the converse of Bezout's Identity hold, we can specify the relationship more carefully. Fill in each blank of Figure 5·1 with the appropriate justification or statement.

5·2 A card trick

This section describes and explains a card trick based on an old Chinese observation.[1] Recall from Sections 2·1 and 4·5 that for any positive m we can perform clockwork addition in the group $\mathbb{Z}_m$. We often write $[x]$ for the elements of $\mathbb{Z}_m$ to emphasize that its elements are cosets.

The simple Chinese Remainder Theorem

[1] I asked Dr. Ding what the Chinese call this theorem. He looked it up in one of his books, and told me that they call it Sun Tzu's Theorem. This is not the same as the author of *The Art of War*.

The Chinese Remainder Theorem, simple version. *Let $m, n \in \mathbb{Z}$ such that* $\gcd(m, n) = 1$. *Let $\alpha, \beta \in \mathbb{Z}$. There exists a solution $x \in \mathbb{Z}$ to the system of linear congruences*

$$\begin{cases} [x] = [\alpha] \text{ in } \mathbb{Z}_m; \\ [x] = [\beta] \text{ in } \mathbb{Z}_n; \end{cases}$$

and $[x]$ is unique in $\mathbb{Z}_N$ where $N = mn$.

Before giving a proof, let's look at an example of how this works in practice.

Example 5·14. Take twelve cards and ask a friend to pick one, then shuffle them. Do the following:

- Lay the cards out in three columns (from left to right), and ask your friend to identify which column contains the card. Remember the answer as 1, 2, or 3. (Use 1 as leftmost, 3 as rightmost.)
- Collect the cards in such a way that *their order is preserved!*
- Lay the cards out again in four columns (from left to right), and ask your friend to identify which column contains the card. Remember the answer as 1, 2, 3, or 4. (Again, 1 is leftmost, 4 rightmost.)
- If α is the first number and β the second, compute $\gamma = 4\alpha - 3\beta$. If the result is negative, add 12.
- Starting from the first card, *in the same order you laid out the cards,* count to the γ'th card. This is your friend's card.

How does this trick work? Each time, your friend identified the *column* in which the mystery card lay. Laying out the cards in rows of three and four corresponds to division by three and four, so that α and β are the remainders from division by three and by four. This corresponds to a system of linear congruences,

$$\begin{cases} [x] = [\alpha] \text{ in } \mathbb{Z}_3 \\ [x] = [\beta] \text{ in } \mathbb{Z}_4 \end{cases},$$

where x is the location of the mystery card. The simple version of the Chinese Remainder Theorem guarantees that the value of x is unique in $\mathbb{Z}_{12}$. Since there are only twelve cards, the solution is unique in the game: as long as the dealer can compute x, s/he can identify the card infallibly.

"Well, and good," you think, "but knowing only the existence of a solution seems rather pointless. I also need to know *how* to compute x, so that I can pinpoint the location of the card." Bézout's identity is the key to unlocking the Chinese Remainder Theorem. Before doing so, we need an important lemma about numbers whose gcd is 1.

Lemma 5·15. *Let $d, m, n \in \mathbb{Z}$. If $m \mid nd$ and* $\gcd(m, n) = 1$, *then $m \mid d$.*

Proof. Assume that $m \mid nd$ and $\gcd(m,n) = 1$. By definition of divisibility, there exists $q \in \mathbb{Z}$ such that $qm = nd$. Use the Extended Euclidean Algorithm to choose $a, b \in \mathbb{Z}$ such that $am + bn = \gcd(m,n) = 1$. Multiplying both sides of this equation by d, we have

$$\begin{aligned}(am + bn)\, d &= 1 \cdot d \\ amd + b\,(nd) &= d \\ adm + b\,(qm) &= d \\ (ad + bq)\, m &= d.\end{aligned}$$

Hence $m \mid d$. □

Now we prove the Chinese Remainder Theorem. You should study this proof carefully, not only to understand the theorem better, but because the proof tells you how to solve the system.

Proof of the Chinese Remainder Theorem, simple version. Recall that the system is

$$\begin{cases} [x] = [\alpha] \text{ in } \mathbb{Z}_m \\ [x] = [\beta] \text{ in } \mathbb{Z}_n \end{cases}.$$

We have to prove two things: first, that a solution x exists; second, that $[x]$ is unique in $\mathbb{Z}_N$.

Existence: Because $\gcd(m,n) = 1$, the Extended Euclidean Algorithm tells us there exist $a, b \in \mathbb{Z}$ such that $am + bn = 1$. Rewriting this equation two different ways, we have $bn = 1 + (-a)\, m$ and $am = 1 + (-b)\, n$. In terms of cosets of subgroups of $\mathbb{Z}$, these two equations tell us that $bn \in 1 + m\mathbb{Z}$ and $am \in 1 + n\mathbb{Z}$. In the bracket notation, $[bn]_m = [1]_m$ and $[am]_n = [1]_n$. Remember that $[\alpha]_m = \alpha\,[1]_m = \alpha\,[bn]_m = [\alpha bn]_m$ and likewise $[\beta]_n = [\beta am]_n$. Apply similar reasoning to see that $[\alpha bn]_n = [0]_n$ and $[\beta am]_m = [0]_m$ in $\mathbb{Z}_m$. Hence,

$$\begin{cases} [\alpha bn + \beta am]_m = [\alpha]_m \\ [\alpha bn + \beta am]_n = [\beta]_n \end{cases}.$$

If we let $x = \alpha bn + \beta am$, then the equations above show that x is a solution to the system.

Uniqueness: Suppose that there exist $[x], [y] \in \mathbb{Z}_N$ that both satisfy the system. Since $[x] = [\alpha] = [y]$ in $\mathbb{Z}_m$, $[x - y] = [0]$, and by Lemma 4·90 on page 148, $m \mid (x - y)$. A similar argument shows that $n \mid (x - y)$. By definition of divisibility, there exists $q \in \mathbb{Z}$ such that $mq = x - y$. By substitution, $n \mid mq$. By Lemma 5·15, $n \mid q$. By definition of divisibility, there exists $q' \in \mathbb{Z}$ such that $q = nq'$. By substitution,

$$x - y = mq = mnq' = Nq'.$$

Hence $N \mid (x - y)$, and again by Lemma 4·90 $[x]_N = [y]_N$, which means that the solution x is unique in $\mathbb{Z}_N$, as desired. □

Pseudocode to solve the Chinese Remainder Theorem appears as Algorithm 5.3 on the following page.

Algorithm 5.3 Solution to Chinese Remainder Theorem, simple version

inputs
$m, n \in \mathbb{Z}$ such that $\gcd(m, n) = 1$
$\alpha, \beta \in \mathbb{Z}$
outputs
$x \in \mathbb{Z}$ satisfying the Chinese Remainder Theorem
do
Use the Extended Euclidean Algorithm to find $a, b \in \mathbb{Z}$ such that $am + bn = 1$
return $[\alpha bn + \beta am]_N$

Example 5·16. The algorithm of Corollary 5.3 finally explains the method of the card trick. We have $m = 3$, $n = 4$, and $N = 12$. Suppose that the player indicates that his card is in the first column when they are grouped by threes, and in the third column when they are grouped by fours; then $\alpha = 1$ and $\beta = 3$.

Using the Extended Euclidean Algorithm, we find that $a = -1$ and $b = 1$ satisfy $am + bn = 1$; hence $am = -3$ and $bn = 4$. We can therefore find the mystery card by computing

$$x = 1 \cdot 4 + 3 \cdot (-3) = -5.$$

Its canonical representation in $\mathbb{Z}_{12}$ is

$$[x] = [-5 + 12] = [7],$$

which implies that the player chose the 7th card. In fact, $[7] = [1]$ in $\mathbb{Z}_3$, and $[7] = [3]$ in $\mathbb{Z}_4$, which agrees with the information given.

Question 5·17 .
Solve the system of linear congruences

$$\begin{cases} [x] = [2] \text{ in } \mathbb{Z}_4 \\ [x] = [3] \text{ in } \mathbb{Z}_9 \end{cases}.$$

Express your answer so that $0 \leqslant x < 36$.

Question 5·18 .
Explain why you can modify the card trick to use 24 cards by doing everything the same, with one exception: if the γ'th card isn't the one your friend chose, then you can add or subtract 12 to find the right one.

Question 5·19 .
Give directions for a similar card trick on all 52 cards, where the cards are grouped first by 4's, then by 13's. Do you think this would be a practical card trick?

Question 5·20 .
Is it possible to modify the card trick to work with only ten cards instead of 12? If so, how; if not, why not?

The Chinese Remainder Theorem can be generalized to larger systems with more than two equations under certain circumstances.

A generalized Chinese Remainder Theorem

What if you have more than just two ways to arrange the cards? You might like to arrange the cards into rows of 3, 4, and 5, for instance. What about other arrangements?

Chinese Remainder Theorem on $\mathbb{Z}$. *Let $m_1, m_2, \ldots, m_n \in \mathbb{Z}$ and assume $\gcd(m_i, m_j) = 1$ for all $1 \leqslant i < j \leqslant n$. Let $\alpha_1, \alpha_2, \ldots \alpha_n \in \mathbb{Z}$. There exists a solution $x \in \mathbb{Z}$ to the system of linear congruences*

$$\begin{cases} [x] &= [\alpha_1] \text{ in } \mathbb{Z}_{m_1}; \\ [x] &= [\alpha_2] \text{ in } \mathbb{Z}_{m_2}; \\ &\vdots \\ [x] &= [\alpha_n] \text{ in } \mathbb{Z}_{m_n}; \end{cases}$$

and $[x]$ is unique in $\mathbb{Z}_N$ where $N = m_1 m_2 \cdots m_n$.

Before we can prove this version of the Chinese Remainder Theorem, we need to make an observation of $m_1, m_2, \ldots, m_n$.

Lemma 5·21. *Let $m_1, m_2, \ldots, m_n \in \mathbb{Z}$ such that $\gcd(m_i, m_j) = 1$ for all $1 \leqslant i < j \leqslant n$. For each $i = 1, 2, \ldots, n$ define $N_i = N/m_i$ where $N = m_1 m_2 \cdots m_n$; that is, N_i is the product of all the m's except m_i. Then $\gcd(m_i, N_i) = 1$.*

Proof. We show that $\gcd(m_1, N_1) = 1$; for $i = 2, \ldots, n$ the proof is similar.

Use the Extended Euclidean Algorithm to choose $a, b \in \mathbb{Z}$ such that $am_1 + bm_2 = 1$. Use it again to choose $c, d \in \mathbb{Z}$ such that $cm_1 + dm_3 = 1$. Then

$$\begin{aligned} 1 &= (am_1 + bm_2)(cm_1 + dm_3) \\ &= (acm_1 + adm_3 + bcm_2)\, m_1 + (bd)(m_2 m_3). \end{aligned}$$

Let $x = \gcd(m_1, m_2 m_3)$; since x divides both m_1 and $m_2 m_3$, it divides each term of the right hand side above. That right hand side equals 1, so x also divides 1. The only divisors of 1 are ± 1, so $x = 1$. We have shown that $\gcd(m_1, m_2 m_3) = 1$.

Rewrite the equation above as $1 = a' m_1 + b' m_2 m_3$; notice that $a', b' \in \mathbb{Z}$. Use the Extended Euclidean Algorithm to choose $e, f \in \mathbb{Z}$ such that $em_1 + fm_4 = 1$. Then

$$\begin{aligned} 1 &= (a' m_1 + b' m_2 m_3)(em_1 + fm_4) \\ &= (a' em_1 + a' fm_4 + b' em_2 m_e)\, m_1 + (b' f)(m_2 m_3 m_4). \end{aligned}$$

An argument similar to the one above shows that $\gcd(m_1, m_2 m_3 m_4) = 1$.

Repeating this process with each m_i, we obtain $\gcd(m_1, m_2 m_3 \cdots m_n) = 1$. Since $N_1 = m_2 m_3 \cdots m_n$, we have $\gcd(m_1, N_1) = 1$. □

We can now prove the Chinese Remainder Theorem on $\mathbb{Z}$.

Proof of the Chinese Remainder Theorem on $\mathbb{Z}$. Existence: Write $N_i = N/m_i$ for $i = 1, 2, \ldots, n$. By Lemma 5·21, $\gcd(m_i, N_i) = 1$. Use the Extended Euclidean Algorithm to compute appropriate a's and b's satisfying

$$\begin{aligned} a_1 m_1 + b_1 N_1 &= 1 \\ a_2 m_2 + b_2 N_2 &= 1 \\ &\vdots \\ a_n m_n + b_n N_n &= 1. \end{aligned}$$

Put $x = \alpha_1 b_1 N_1 + \alpha_2 b_2 N_2 + \cdots + \alpha_n b_n N_n$. Now, $b_1 N_1 = 1 + (-a_1)\, m_1$, so $[b_1 N_1] = [1]$ in $\mathbb{Z}_{m_1}$, so $[\alpha_1 b_1 N_1] = [\alpha_1]$ in $\mathbb{Z}_{m_1}$. Moreover, for any $i = 2, 3, \ldots, n$, inspection of N_i verifies that $m_1 \mid N_i$, implying that $[\alpha_i b_i N_i]_{m_1} = [0]_{m_1}$ (Lemma 4·90). Hence, in $\mathbb{Z}_{m_1}$ the value of $[x]$ simplifies as

$$\begin{aligned} [x] &= [\alpha_1 b_1 N_1 + \alpha_2 b_2 N_2 + \cdots + \alpha_n b_n N_n] \\ &= [\alpha_1] + [0] + \cdots + [0]. \end{aligned}$$

A similar argument shows that $[x] = [\alpha_i]$ in $\mathbb{Z}_{m_i}$ for $i = 2, 3, \ldots, n$.

Uniqueness: As in the previous case, let $[x], [y]$ be two solutions to the system in $\mathbb{Z}_N$. Then $[x - y] = [0]$ in $\mathbb{Z}_{m_i}$ for $i = 1, 2, \ldots, n$, implying that $m_i \mid (x - y)$ for $i = 1, 2, \ldots, n$. We use the definition of divisibility:

Since $m_1 \mid (x - y)$, there exists $q_1 \in \mathbb{Z}$ such that $x - y = m_1 q_1$.

Since $m_2 \mid (x - y)$, substitution implies $m_2 \mid m_1 q_1$, and Lemma 5·15 implies that $m_2 \mid q_1$. There exists $q_2 \in \mathbb{Z}$ such that $q_1 = m_2 q_2$. Substitution implies that $x - y = m_1 m_2 q_2$.

Since $m_3 \mid (x - y)$, substitution implies $m_3 \mid m_1 m_2 q_2$. By Lemma 5·21, $\gcd(m_1 m_2, m_3) = 1$, and Lemma 5·15 implies that $m_3 \mid q_2$. There exists $q_3 \in \mathbb{Z}$ such that $q_2 = m_3 q_3$. Substitution implies that $x - y = m_1 m_2 m_3 q_3$.

Continuing in this fashion obtains $x - y = m_1 m_2 \cdots m_n q_n$ for some $q_n \in \mathbb{Z}$. By substition, $x - y = N q_n$, so $[x - y] = [0]$ in $\mathbb{Z}_N$, so $[x] = [y]$ in $\mathbb{Z}_n$. That is, the solution to the system is unique in $\mathbb{Z}_N$. □

The algorithm to solve such systems is similar to that given for the simple version, in that it can be obtained from the proof of existence of a solution.

Question 5·22 .

Solve the system of linear congruences

$$\begin{cases} [x] = [2] \text{ in } \mathbb{Z}_5 \\ [x] = [3] \text{ in } \mathbb{Z}_6 \\ [x] = [4] \text{ in } \mathbb{Z}_7 \end{cases}.$$

Question 5·23 .
Solve the system of linear congruences

$$\begin{cases} [x] = [33] \text{ in } \mathbb{Z}_{16} \\ [x] = [-4] \text{ in } \mathbb{Z}_{33} \\ [x] = [17] \text{ in } \mathbb{Z}_{504} \end{cases}.$$

This problem is a little tougher than the previous, since $\gcd(16, 504) \neq 1$ and $\gcd(33, 504) \neq 1$. Since you can't use either of the Chinese Remainder Theorems presented here, you'll have to generalize their approaches to get a method for this one.

Question 5·24 .
Is it possible to modify the card trick to work with only eight cards instead of 12? If so, how; if not, why not?

5·3 The Fundamental Theorem of Arithmetic

In this section, we address a fundamental result of number theory with algebraic implications. Let's recall what Definition 3·20 means in the context of natural numbers.

Definition 5·25. Let $n \in \mathbb{N}^+$ and $n \neq 1$. We say that n is **irreducible** if the only integers that divide n are ± 1 and $\pm n$.

(We may sometimes refer to certain negative numbers as irreducible. While certain negative numbers do satisfy the property of irreducibility, there are reasons that only natural numbers are properly called prime.)

You may be wondering why we call these integers *irreducible* instead of *prime,* the customary term in earlier classes. We'll say more about that in a moment.

Example 5·26. The integer 36 is not irreducible, because $36 = 6 \times 6$. The integer 7 is irreducible, because the only integers that divide 7 are ± 1 and ± 7.

One useful aspect to irreducible integers is that, aside from ± 1, any integer is divisible by at least one irreducible integer.

Theorem 5·27. *Let n be any integer besides ± 1. There exists at least one irreducible integer p such that $p \mid n$.*

Proof. Case 1: If $n = 0$, then 2 is a divisor of n, and we are done.

Case 2: Assume that $n \in \mathbb{N}^+$ and $n \neq 1$. Let $a_0 = n$. If a_0 is not irreducible, then by definition $a_0 = a_1 b_1$ such that $a_1, b_1 \in \mathbb{Z}$ and $a_1, b_1 \neq \pm 1$. Without loss of generality, we may assume that $a_1, b_1 \in \mathbb{N}^+$ (otherwise both are negative and we can replace them with their opposites). Observe further that $a_1 < a_0$ (this is a consequence of Question 1.28 on page 13). If a_1 is irreducible, then we are done; otherwise, we can write $a_1 = a_2 b_2$ where $a_2, b_2 \in \mathbb{N}^+$ and $a_2 < a_1$. Continuing in this fashion, as long as a_i is not irreducible, we can find $a_{i+1}, b_{i+1} \in \mathbb{N}^+$

such that $a_i = a_{i+1}b_{i+1}$, with $a_i > a_{i+1}$ for each i. We have a strictly decreasing sequence of elements,

$$a_0 > a_1 > a_2 > \cdots .$$

By Question 1·51, this sequence *must* be finite. Let a_m be the final element in the sequence. We claim that a_m is irreducible; after all, were it not irreducible, then we could extend the sequence further, which we cannot. By substitution,

$$n = a_1 b_1 = a_2 (b_2 b_1) = \cdots = a_m (b_{m-1} \cdots b_1) .$$

That is, a_m is an irreducible integer that divides n.

Case 3: Assume that n is negative, but not -1. Let $m = -n$. Case 2 implies that there exists an irreducible integer p such that $p \mid m$. By definition, $m = qp$ for some $q \in \mathbb{Z}$. By substitution and properties of arithmetic, $n = -(qp) = (-q)\, p$, so $p \mid n$. □

Question 5·28 .

Show that there are infinitely many irreducible numbers. *Hint:* Proceed by contradiction: suppose there is a finite list of irreducible numbers, then exploit the Division Theorem to construct a remainder whose division by each of those irreducible numbers is nonzero. Theorem 5·27 does the rest.

Let's turn now to the term you might have expected for the definition given above: a *prime* number. We actually associate a different notion with this term.

Definition 5·29. Let R be a ring, and suppose $p \in R$ is not a unit. We say that p is **prime** if, whenever we find $a, b \in R$ such that $p \mid ab$, then $p \mid a$ or $p \mid b$. Consistent with this definition, a natural number p is a **prime number** if $p \neq 1$ and for any two integers a, b we have

$$p \mid ab \quad \Longrightarrow \quad p \mid a \text{ or } p \mid b.$$

(We may sometimes refer to certain negative numbers as prime. While certain negative numbers do satisfy the property of being prime, called **primality**, there are reasons that only natural numbers are properly called prime.)

Example 5·30. Let $a = 68$ and $b = 25$. It is easy to recognize that 10 divides $ab = 1700$. However, 10 divides neither a nor b, so 10 is not a prime number.

It is also easy to recognize that 17 divides $ab = 1700$. Unlike 10, 17 divides one of a or b; in fact, it divides a. Were we to look at every possible product ab divisible by 17, we would find that 17 always divides one of the factors a or b. Thus, 17 is prime.

If the next-to-last sentence in the example, bothers you, *good.* I've claimed something about every product divisible by 17, but haven't explained why that is true. That's cheating! If I'm going to claim that 17 is prime, I need a better explanation than, "look at every possible product ab." After all, there are infinitely many products possible, and we can't do that in finite time. We need a *finite* criterion.

To this end, let's return to the notion of an irreducible number. It's fairly easy to tell if an integer a is irreducible; Question 1.28 tells us to look for factors among natural numbers smaller than $|a|$. If we knew that prime numbers were irreducible, then we could simply test for irreducibility. Could it be that the definitions are *distinctions without a difference?*

Theorem 5·31. *An integer is prime if and only if it is irreducible.*

Proof. This proof has two parts. You will show in Question 5.32 that if an integer is prime, then it is irreducible. Here, we show the converse.

Let $n \in \mathbb{N}^+ \setminus \{1\}$ and assume that n is irreducible. To show that n is prime, we must take arbitrary $a, b \in \mathbb{Z}$ and show that if $n \mid ab$, then $n \mid a$ or $n \mid b$. Therefore, let $a, b \in \mathbb{Z}$ and assume that $n \mid ab$. If $n \mid a$, then we would be done, so assume that $n \nmid a$. We must show that $n \mid b$.

By definition, the common factors of n and a are a subset of the factors of n. Since n is irreducible, its factors are ± 1 and $\pm n$. By hypothesis, $n \nmid a$, so $\pm n$ cannot be common factors of n and a. Thus, the only common factors of n and a are ± 1, which means that $\gcd(n, a) = 1$. By Lemma 5·15, $n \mid b$.

We assumed that if n is irreducible and divides ab, then n must divide one of a or b. By definition, n is prime. □

Question 5·32 .
Show that any prime integer p is irreducible.

If the two definitions are equivalent, why would we give a different definition? It turns out that the concepts are equivalent *for the integers,* but not for other sets; you will see this in detail in Section 6·2.

The following theorem is a cornerstone of Number Theory.

The Fundamental Theorem of Arithmetic. *Let $n \in \mathbb{N}^+$ but $n \neq 1$. We can* **factor n into irreducibles***; that is, we can write*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$$

where $p_1, p_2, \ldots, p_r$ are irreducible and $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{N}$. The representation is unique if we order $p_1 < p_2 < \ldots < p_n$.

Since prime integers are irreducible and vice versa, you can replace "irreducible" by "prime" and obtain the expression of this theorem found more commonly in number theory textbooks.

Proof. The proof has two parts: a proof of existence and a proof of uniqueness.

Existence: We proceed by induction on positive integers.

Inductive base: If $n = 2$, then n is irreducible, and we are finished.

Inductive hypothesis: Assume that the integers $2, 3, \ldots, n-1$ have a factorization into irreducibles.

Inductive step: If n is irreducible, then we are finished. Otherwise, n is not irreducible. By Lemma 5·27, there exists an irreducible integer p_1 such that $p_1 \mid n$. By definition, there exists $q \in \mathbb{N}^+$ such that $n = qp_1$. Since $p_1 \neq 1$, Question 1.58 tells us that $q < n$. By the inductive hypothesis, q has a factorization into irreducibles; say

$$q = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}.$$

Thus $n = qp = p_1^{\alpha_1 + 1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$; that is, n factors into irreducibles.

Uniqueness: Here we use the fact that irreducible numbers are also prime (Lemma 5·31). Assume that $p_1 < p_2 < \cdots < p_r$ and we can factor n as

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r} = p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r}.$$

Without loss of generality, we may assume that $\alpha_1 \leqslant \beta_1$. It follows that

$$p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_r^{\alpha_r} = p_1^{\beta_1 - \alpha_1} p_2^{\beta_2} p_3^{\beta_3} \cdots p_r^{\beta_r}.$$

This equation implies that $p_1^{\beta_1 - \alpha_1}$ divides the expression on the left hand side of the equation. Since p_1 is irreducible, hence prime, $\beta_1 - \alpha_1 \neq 0$ implies that p_1 divides one of $p_2, p_3, \ldots, p_r$. This contradicts the irreducibility of p_2, p_3, …, p_r. Hence $\beta_1 - \alpha_1 = 0$. A similar argument shows that $\beta_i = \alpha_i$ for all $i = 1, 2, \ldots, r$; hence the representation of n as a product of irreducible integers is unique. □

Question 5·33 .
Fill in each blank of Figure 5·2 with the justification.

Question 5·34 .
Let $n \in \mathbb{N}^+$. Modify the proof in Figure 5·2 to show that if p is irreducible, then $\sqrt[n]{p}$ is irrational.

Question 5·35 .
Let $n \in \mathbb{N}^+$. Modify the proof in Figure 5·2 to show that if there exists an irreducible integer p such that $p \mid n$ but $p^2 \nmid n$, then $\sqrt{n}$ is irrational.

5·4 Multiplicative clockwork groups

Throughout this section, $n \in \mathbb{N}^+ \setminus \{1\}$, unless otherwise stated.

Clockwork multiplication

Recall that $\mathbb{Z}_n$ is an additive group, but not multiplicative. In this section we find for each eligible n a subset of $\mathbb{Z}_n$ that we can turn into a multiplicative group.

Example 5·36. Recall that $\mathbb{Z}_5 \cong \mathbb{Z}/\langle 5 \rangle$. We saw that it was a ring; that is, it is an abelian group under addition, a monoid under multiplication, and multiplication distributes over addition.

Can we turn a subset of it into a multiplicative group? We need to identify an identity, and inverses. Certainly $[0]$ won't have a multiplicative inverse, but what about $\mathbb{Z}_5 \setminus \{[0]\}$? This generates a multiplication table that satisfies the properties of an abelian (but non-additive) group:

$\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

That is a group! We'll call it $\mathbb{Z}_5^*$.

In fact, $\mathbb{Z}_5^* \cong \mathbb{Z}_4$; they are both cyclic groups of four elements, and inspection shows that $\mathbb{Z}_5 = \langle 2 \rangle = \langle 3 \rangle = \langle 4 \rangle$. In $\mathbb{Z}_5^*$, however, the nominal operation is multiplication, whereas in $\mathbb{Z}_4$ the nominal operation is addition.

Claim: If p is irreducible, then $\sqrt{p}$ is not rational.
Proof:

1. Assume that p is irreducible.
2. By way of contradiction, assume that $\sqrt{p}$ *is* rational.
3. By _____, there exist $a, b \in \mathbb{N}$ such that $\sqrt{p} = a/b$.
4. Without loss of generality, we may assume that $\gcd(a, b) = 1$.
   (After all, we could otherwise rewrite $\sqrt{p} = (a/d) \,/\, (b/d)$, where $d = \gcd(a, b)$.)
5. By _____, $p = a^2/b^2$.
6. By _____, $pb^2 = a^2$.
7. By _____, $p \mid a^2$.
8. By _____, p is prime.
9. By _____, $p \mid a$.
10. By _____, $a = pq$ for some $q \in \mathbb{Z}$.
11. By _____ and _____, $pb^2 = (pq)^2 = p^2q^2$.
12. By _____, $b^2 = pq^2$.
13. By _____, $p \mid b^2$.
14. By _____, $p \mid b$.
15. This contradicts step _____. Our assumption that $\sqrt{p}$ is rational must have been wrong. Hence, $\sqrt{p}$ is irrational.

Figure 5·2: Material for Question 5.33

You might think that this trick of dropping zero and building a multiplication table always works, *but it doesn't.*

Example 5·37. Recall that $\mathbb{Z}_4 = \mathbb{Z}/\langle 4\rangle = \{[0],[1],[2],[3]\}$. Consider the set $\mathbb{Z}_4\setminus\{[0]\} = \{[1],[2],[3]\}$. The multiplication table for this set *is not closed* because

$$[2]\cdot[2] = [4] = [0] \notin \mathbb{Z}_4\setminus\{[0]\}.$$

We obviously can't fix this by including zero, as well: zero has no inverse. So, we must exclude zero; our mistake seems to have been that we included 2 *Excluding* 2 finally works out:

$\times$	1	3
1	1	3
3	3	1

That is a group! We'll call it $\mathbb{Z}_4^*$.

In fact, $\mathbb{Z}_4^* \cong \mathbb{Z}_2$; they are both the cyclic group of two elements. In $\mathbb{Z}_4^*$, however, the operation is multiplication, whereas in $\mathbb{Z}_2$, the operation is addition.

You can determine for yourself that $\mathbb{Z}_2\setminus\{[0]\} = \{[1]\}$ and $\mathbb{Z}_3\setminus\{[0]\} = \{[1],[2]\}$ are also multiplicative groups. In this case, as in $\mathbb{Z}_5^*$, we need remove only 0. For $\mathbb{Z}_6$, however, we have to remove nearly all the elements! We only get a group from $\mathbb{Z}_6\setminus\{[0],[2],[3],[4]\} = \{[1],[5]\}$.

Why do we need to remove more elements of $\mathbb{Z}_n$ for some values of n than others? Aside from zero, which clearly has no inverse under the operation specified, the elements we've had to remove are those whose multiplication would re-introduce zero. *We're observing zero divisors again.*

Can we find a criterion to detect this? You should have done this in Question 2.44; to be safe, let's flesh it out here.

Lemma 5·38. *Let $x \in \mathbb{Z}_n$ be nonzero. The following are equivalent:*

(A) x is a zero divisor.

(B) x and n have a common divisor besides ± 1.

Proof. That (B) implies (A): Assume that x and n share a common divisor $d \neq 0, 1$. Use the definition of divisibility to choose $t, q \in \mathbb{Z}\setminus\{0\}$ such that $n = qd$ and $x = td$. Let y be the remainder of dividing q by n. Substitution implies that

$$xy \equiv_n xq = (td)\,q = t\,(dq) = tn \equiv 0.$$

Since $d \neq 0, 1$, $-n < q < n$, so $0 \neq q \equiv_n y$. This shows that y is also nonzero, so x is a zero divisor.

Question 5·39 .

You can also prove that (B) implies (A) using Bézout's Lemma. Try it that way.

Proof of Lemma 5·38, continued. That (A) implies (B): Assume that x is a zero divisor. By definition, we can find nonzero $y \in \mathbb{Z}_n$ such that $xy \equiv_n 0$. There are two points to recall here: first, $0 \leqslant y < n$, and second, $n \mid xy$. By definition, we can find $q \in \mathbb{Z}$ such that $nq = xy$. Use the Fundamental Theorem of Arithmetic to factor $n = p_1^{a_1} \cdots p_k^{a_k}$, where the p_i's are distinct irreducibles and the a_i's are natural. By substitution,

$$\left(p_1^{a_1} \cdots p_k^{a_k}\right) q = xy.$$

Not every $p_i^{a_i}$ can appear in y; otherwise, $n \mid y$, and by Question 1.58, we would have $n \leqslant y$, contradicting $y < n$. Hence at least one p_i divides x, so that n and x have a common divisor that is not 1. □

A multiplicative clockwork group

We can thus construct a *multiplicative* clockwork group using the elements of $\mathbb{Z}_n$ that are not zero divisors.

Definition 5·40. Define the set $\mathbb{Z}_n^*$ to be the set of elements of $\mathbb{Z}_n$ that are not zero divisors. In set builder notation,

$$\mathbb{Z}_n^* := \{x \in \mathbb{Z}_n \setminus \{0\} : \ \forall y \in \mathbb{Z}_n \setminus \{0\} \ \ xy \neq 0\}.$$

We claim that $\mathbb{Z}_n^*$ is a group under multiplication. Keep in mind that, while it is a sub*set* of $\mathbb{Z}_n$, it is not a sub*group*, as the operations are different.

Theorem 5·41. *$\mathbb{Z}_n^*$ is an abelian group under its multiplication.*

Proof. We check each requirement of a group, slightly out of order. Let $a, b, c \in \mathbb{Z}_n^*$.

(associative) From Question 2.20, clockwork multiplication is consistent with integer multiplication. Since $(ab)\,c = a\,(bc)$, then, $(ab)\,c \equiv a\,(bc)$. Notice that this applies for elements of $\mathbb{Z}_n$ as well as elements of $\mathbb{Z}_n^*$.

(closed) Assume to the contrary that $ab \notin \mathbb{Z}_n^*$. We have defined ab to give us an element of $\mathbb{Z}_n$, so the only way $ab \notin \mathbb{Z}_n^*$ is if $ab \equiv 0$ or ab is a zero divisor. By definition of $\mathbb{Z}_n^*$, neither a nor b is a zero divisor, so $ab \not\equiv 0$, which forces us to conclude that ab is a zero divisor. Choose $c \in \mathbb{Z}_n$ such that $(ab)\,c \equiv 0$. By the associative property, $a\,(bc) \equiv 0$; that is, a is a zero divisor, contradicting the choice of a! Thus, *ab cannot* be a zero divisor, either; the assumption that $ab \notin \mathbb{Z}_n^*$ must have been wrong.

(identity) We claim that 1 is the identity. Since $\gcd(1, n) = 1$, we have $1 \in \mathbb{Z}_n^*$ by definition. It is then trivial that $1 \cdot a = a = a \cdot 1$.

(inverse) We need to find an inverse of a. By definition, a and n have no common divisors except ± 1; hence $\gcd(a, n) = 1$. Bézout's Lemma tells us we can find $b, m \in \mathbb{Z}$ such that $ab + mn = 1$. We deduce that

$$\begin{aligned} ab - 1 &= n\,(-m) \\ \therefore\ ab - 1 &\in n\mathbb{Z} \\ \therefore\ ab &\equiv 1. \end{aligned}$$

But is $b \in \mathbb{Z}_n^*$? It might not be. To start with, we could have $b \geqslant n$ or $b < 0$. In this case, let q and r be the quotient and remainder of division of b by n; then $ar \equiv 1$. But what if r is a zero divisor? Recall the equation above:

$$ab + mn = 1 \quad \Rightarrow \quad a(nq + r) + mn = 1 \quad \Rightarrow \quad ar + (m + aq)n = 1.$$

This is a form of the identity in Bézout's Lemma not just for a, but also for r! Bézout's Lemma tells us that $\gcd(r, n)$ is the smallest positive number that can be written in that form, so $\gcd(r, n) = 1$, so r is in fact a zero divisor by Lemma 5·38, so $a^{-1} = r \in \mathbb{Z}_n^*$.

(commutative) Use the definition of multiplication in $\mathbb{Z}_n^*$ and the commutative property of integer multiplication to see that $ab = ba$.

□

By removing elements that share non-trivial common divisors with n, we have managed to eliminate those elements that do not satisfy the zero-product rule, and would break closure by trying to re-introduce zero in the multiplication table. We have thereby created a clockwork group for multiplication, $\mathbb{Z}_n^*$.

Example 5·42. Consider $\mathbb{Z}_{10}^*$. To find its elements, collect the elements of $\mathbb{Z}_{10}$ that are not zero divisors. Lemma 5·38 tells us that these are the elements a such that $\gcd(a, n) \neq 1$. Thus

$$\mathbb{Z}_{10}^* = \{1, 3, 7, 9\}.$$

Theorem 5·41 tells us that $\mathbb{Z}_{10}^*$ is a group. Since it has four elements, it must be isomorphic to either the Klein 4-group, or to $\mathbb{Z}_4$. Which is it? In this case, it's probably easiest to decide the question with a glance at its multiplication table:

$\times$	**1**	**3**	**7**	**9**
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

Notice that $3^{-1} \neq 3$. In the Klein 4-group, every element is its own inverse, so $\mathbb{Z}_{10}^*$ cannot be isomorphic to the Klein 4-group. Instead, it must be isomorphic to $\mathbb{Z}_4$.

Question 5·43 .
List the elements of $\mathbb{Z}_7^*$ using their canonical representations, and construct its multiplication table. Use the table to identify the inverse of each element.

Question 5·44 .
List the elements of $\mathbb{Z}_{15}^*$ using their canonical representations, and construct its multiplication table. Use the table to identify the inverse of each element.

5·5 Euler's Theorem and fast exponentiation

In Section 5·4 we defined the group $\mathbb{Z}_n^*$ for all $n \in \mathbb{N}^+$ where $n > 1$. The order of this group is more important than you might think. To begin with, number theorists are very interested in the following function.

Definition 5·45. Euler's φ-function counts the number of positive natural numbers that are both smaller than n and relatively prime to it.

We built the group $\mathbb{Z}_n^*$ using these same integers, so:

Fact 5·46. *For $n > 1$, $\varphi(n) = |\mathbb{Z}_n^*|$.*

To see why this is such a big deal, consider the algebraic ramifications, starting with a corollary to Lagrange's Theorem.

Euler's Theorem for integers. *For all $x \in \mathbb{Z}_n^*$, $x^{\varphi(n)} = 1$.*

Proofs of Euler's Theorem based only on Number Theory are not very easy. They're not particularly difficult, either; they just aren't easy. See for example the proof on pages 18–19 of [2]. Compare this with our algebraic proof of Euler's Theorem: it fits in one line!

Proof. Let $x \in \mathbb{Z}_n^*$. By Question 4.117, $x^{|\mathbb{Z}_n^*|} = 1$. By substitution, $x^{\varphi(n)} = 1$. □

Corollary 5·47. *For all $x \in \mathbb{Z}_n^*$, $x^{-1} = x^{\varphi(n)-1}$.*

Proof. You do it! □

Question 5·48 .
Prove that for all $x \in \mathbb{Z}_n^*$, $x^{\varphi(n)-1} = x^{-1}$.

Question 5·49 .
Prove that for all $x \in \mathbb{N}^+$, if x and n have no common divisors, then $n \mid \left(x^{\varphi(n)} - 1\right)$.

Computing $\varphi(n)$

We see that $\varphi(n)$ is a pretty big deal; and that ain't the half of it; see the next section for a real barn burner. Of course, if we intend to *use* these applications, we first need an efficient way to compute $\varphi(n)$.

Well, then, how do we compute $\varphi(n)$? For an irreducible integer p, this is easy: the only common factors between p and any positive integer less than p are ± 1; there are $p - 1$ of these, so $\varphi(p) = p - 1$.

For integers that factor, it is not so easy. Checking a few examples, no clear pattern emerges:

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\lvert\mathbb{Z}_n^*\rvert$	1	2	2	4	2	6	4	6	4	10	4	12	6	8

Computing $\varphi(n)$ turns out to be hard in practice. It is a major research topic in number theory, and its difficulty makes the RSA algorithm secure (see Section 5·6). One approach, of course, is to factor n and count the integers that do not share any common factors. For example,

$$28 = 2^2 \cdot 7,$$

so to compute $\varphi(28)$, we could look at all the positive integers smaller than 28 that do not have 2 or 7 as factors. Try this on your own, though, and you'll discover how tedious it is. We'd like an *efficient* way to compute $\varphi(n)$.

Another way would be to compute $\varphi(m)$ for each factor m of n, then recombine them. But, how? Lemma 5·50 gives us a first step.

Lemma 5·50. *Let $a, b, n \in \mathbb{N}^+$. If $n = ab$ and $\gcd(a, b) = 1$, then $\varphi(n) = \varphi(a)\varphi(b)$.*

Example 5·51. In the table above, we have $\varphi(15) = 8$. Notice that this satisfies

$$\varphi(15) = \varphi(5 \times 3) = \varphi(5)\varphi(3) = 4 \times 2 = 8.$$

Proof of Lemma 5·50. Assume $n = ab$. Recall that direct products are groups, so that $\mathbb{Z}_a^* \times \mathbb{Z}_b^*$ is a group; the size of this group is $|\mathbb{Z}_a^*| \times |\mathbb{Z}_b^*| = \varphi(a)\varphi(b)$. We claim that $\mathbb{Z}_n^* \cong \mathbb{Z}_a^* \times \mathbb{Z}_b^*$. If true, this would prove the lemma, since

$$\varphi(n) = |\mathbb{Z}_n^*| = |\mathbb{Z}_a^* \times \mathbb{Z}_b^*| = |\mathbb{Z}_a^*| \times |\mathbb{Z}_b^*| = \varphi(a)\varphi(b).$$

To show that they are indeed isomorphic, let $f : \mathbb{Z}_n^* \to \mathbb{Z}_a^* \times \mathbb{Z}_b^*$ by $f([x]_n) = ([x]_a, [x]_b)$. First we show that f is a homomorphism: Let $y, z \in \mathbb{Z}_n^*$; then

$$\begin{aligned}
f([y]_n[z]_n) &= f([yz]_n) && \text{(arithm. in } \mathbb{Z}_n^*) \\
&= ([yz]_a, [yz]_b) && \text{(def. of } f) \\
&= ([y]_a[z]_a, [y]_b[z]_b) && \text{(arithm. in } \mathbb{Z}_a^*, \mathbb{Z}_b^*) \\
&= ([y]_a, [y]_b)([z]_a, [z]_b) && \text{(arithm. in } \mathbb{Z}_a^* \times \mathbb{Z}_b^*) \\
&= f([y]_n) f([z]_n). && \text{(def. of } f)
\end{aligned}$$

It remains to show that f is one-to-one and onto. It is both surprising and delightful that the Chinese Remainder Theorem will do most of the work for us. To show that f is onto, let $([y]_a, [z]_b) \in \mathbb{Z}_a^* \times \mathbb{Z}_b^*$. We need to find $x \in \mathbb{Z}$ such that $f([x]_n) = ([y]_a, [z]_b)$. Consider the system of linear congruences

$$\begin{aligned}
[x] &= [y] \text{ in } \mathbb{Z}_a; \\
[x] &= [z] \text{ in } \mathbb{Z}_b.
\end{aligned}$$

The Chinese Remainder Theorem tells us not only that such x exists in $\mathbb{Z}_n$, but that x is unique in $\mathbb{Z}_n$.

We are not quite done; we have shown that a solution $[x]$ exists in $\mathbb{Z}_n$, but what we really need is that $[x] \in \mathbb{Z}_n^*$. To see that $[x] \in \mathbb{Z}_n^*$, let d be any common divisor of x and n. By way of contradiction, assume $d \neq \pm 1$; by Theorem 5·27, we can find an irreducible divisor r of d; by Question 4.71 on page 142, $r \mid n$ and $r \mid x$. Recall that $n = ab$, so $r \mid ab$. Since r is irreducible,

hence prime, $r \mid a$ or $r \mid b$. Without loss of generality, we may assume that $r \mid a$. Recall that $[x]_a = [y]_a$; Lemma 4·90 on page 148 tells us that $a \mid (x - y)$. Let $w \in \mathbb{Z}$ such that $wa = x - y$. Rewrite this equation as $x - wa = y$. Recall that $r \mid x$ and $r \mid a$; we can factor r from the left-hand side of $x - wa = y$ to see that $r \mid y$.

What have we done? We showed that if x and n have a common factor besides ± 1, then y and a also have a common, irreducible factor r. The definition of irreducible implies that $r \neq 1$.

Do you see the contradiction? We originally chose $[y] \in \mathbb{Z}_a^*$. By definition, $[y]$ cannot be a zero divisor in $\mathbb{Z}_a$, so by Lemma 5·38, $\gcd(y, a) = 1$. But the definition of greatest common divisor means that

$$\gcd(y, a) \geqslant r > 1 = \gcd(y, a),$$

a contradiction! Our assumption that $d \neq 1$ must have been false; we conclude that the only common divisors of x and n are ± 1. Hence, $x \in \mathbb{Z}_n^*$. □

Lemma 5·50 gives us a more efficient way to compute $\varphi(n)$, but it's still not that great, since first you have to find factors a and b of n. This turns out to be quite difficult to do in practice; to see how mathematicians made lemonade of this mathematical lemon, see the next chapter.

Fast exponentiation

Corollary 5·47 gives us an "easy" way to compute the inverse of any $x \in \mathbb{Z}_n^*$. Even supposing we *could* compute $\varphi(n)$ in reasonable time, it can still take a long time to compute $x^{\varphi(n)}$, as it could be a very large number. We take a moment to explain how to compute canonical forms of exponents more quickly. There are two main considerations.

Lemma 5·52. *For any $n \in \mathbb{N}^+$, $[x^a] = [x]^a$ in $\mathbb{Z}_n^*$.*

(In other words, don't compute x^a, and *then* the remainder. Compute the remainder *while* computing x^a.)

Proof. This follows from the fact that multiplication is well-defined, and there are finitely many products. You can prove it by induction if you want more detail than that. □

Example 5·53. In $\mathbb{Z}_{15}^*$ we can determine easily that $[4^{20}] = [4]^{20} = \left([4]^2\right)^{10} = [16]^{10} = [1]^{10} = [1]$. This is a *lot* faster than computing $4^{20} = 1099511627776$, then dividing to find the canonical form.

Do you see what we did? The trick is to break the exponent down into "manageable" powers. How exactly can we do that?

Fast Exponentiation. *Let $a \in \mathbb{N}$ and $x \in \mathbb{Z}$. We can compute x^a in the following way:*

1. *Let b be the largest integer such that $2^b \leqslant a$.*
2. *Let $q_0, q_1, \ldots, q_b$ be the bits of the binary representation of a.*
3. *Let $y = 1$, $z = x$ and $i = 0$.*

4. *Repeat the following until $i > b$:*

   (a) *If $q_i \neq 0$, replace y with the product of y and z.*

   (b) *Replace z with z^2.*

   (c) *Replace i with $i + 1$.*

This ends with $x^a = y$.

Fast Exponentiation effectively computes the *binary representation* of a and uses this to square x repeatedly, multiplying the result only by those powers that matter for the representation. Its algorithm is especially effective on computers, whose mathematics is based on binary arithmetic. Combining it with Lemma 5·52 gives an added bonus in $\mathbb{Z}_n^*$, which is what we care about most.

Example 5·54. Since $10 = 2^3 + 2^1$, we can compute $\left[4^{10}\right]_7$ following the algorithm of Fast Exponentiation:

1. We have $q_3 = 1, q_2 = 0, q_1 = 1, q_0 = 0$.

2. Let $y = 1, z = 4$ and $i = 0$.

3. When $i = 0$:

   (a) We do not change y because $q_0 = 0$.

   (b) Put $z = 4^2 = 16 = 2$. (We're in $\mathbb{Z}_7^*$, remember.)

   (c) Put $i = 1$.

4. When $i = 1$:

   (a) Put $y = 1 \cdot 2 = 2$.

   (b) Put $z = 2^2 = 4$.

   (c) Put $i = 2$.

5. When $i = 2$:

   (a) We do not change y because $q_2 = 0$.

   (b) Put $z = 4^2 = 16 = 2$.

   (c) Put $i = 3$.

6. When $i = 3$:

   (a) Put $y = 2 \cdot 2 = 4$.

   (b) Put $z = 4^2 = 2$.

   (c) Put $i = 4$.

We conclude that $[4^{10}]_7 = [4]_7$. Hand computation the long way, or a half-decent calculator, will verify this.

Proof of Fast Exponentiation.

Termination: Termination is due to the fact that b is a finite number, and the algorithm assigns to i the values $0, 1, \ldots, b+1$ in succession, stopping when $i > b$.

Correctness: First, the theorem claims that $q_b, \ldots, q_0$ are the bits of the binary representation of x^a, but do we actually know that the binary representation of x^a has $b+1$ bits? By hypothesis, b is the largest integer such that $2^b \leqslant a$; if we need one more bit, then the definition of binary representation means that $2^{b+1} \leqslant x^a$, which contradicts the choice of b. Thus, $q_b, \ldots, q_0$ are indeed the bits of the binary representation of x^a. By definition, $q_i \in \{0, 1\}$ for each $i = 0, 1, \ldots, b$. The algorithm multiplies $z = x^{2^i}$ to y only if $q_i \neq 0$, so that the algorithm computes

$$x^{q_b 2^b + q_{b-1} 2^{b-1} + \cdots + q_1 2^1 + q_0 2^0},$$

which is precisely the binary representation of x^a. □

Question 5·55 .

Compute 3^{28} in $\mathbb{Z}$ using fast exponentiation. Show each step.

Question 5·56 .

Compute 24^{28} in $\mathbb{Z}_7^*$ using fast exponentiation. Show each step.

5·6 The RSA encryption algorithm

Whenever you buy a product online, you submit private information: at the very least, a credit card or bank account number, and usually more. There is no guarantee that this information will pass only through servers run by disinterested persons. It is quite possible for the information to pass through a computer run by at least one ill-intentioned hacker, and possibly even organized crime. You probably don't want criminals looking at your credit card number. And not just you – many organizations desire a reliable *and efficient* method to disguise private information so that snoopers cannot understand it.

This problem provides a surprisingly useful application of group theory, via number theory. A number of approaches exist, and a method in common use is the RSA encryption algorithm.[2] First we describe the algorithms for encryption and decryption; then we explain the ideas behind each stage, illustrating with an example; finally we prove that it succesfully encrypts and decrypts messages.

Description and example

The RSA algorithm. *Let M be a list of positive integers. Let p, q be two irreducible integers such that:*

[2] RSA stands for Rivest (of MIT), Shamir (of the Weizmann Institute in Israel), and Adleman (of USC).

- $\gcd(p, q) = 1$; *and*
- $(p-1)(q-1) > \max\{m : m \in M\}$.

Let $N = pq$ and $e \in \mathbb{Z}^*_{\varphi(N)}$. If we apply the following algorithm to M:

1. *Let C be a list of positive integers found by computing the canonical representation of $[m^e]_N$ for each $m \in M$.*

and subsequently apply the following algorithm to C:

1. *Let $d = e^{-1} \in \mathbb{Z}^*_{\varphi(N)}$.*
2. *Let D be a list of positive integers found by computing the canonical representation of $\left[c^d\right]_N$ for each $c \in C$.*

Then $D = M$.

Example 5·57. Consider the text message

ALGEBRA RULZ.

We convert the letters to integers in the fashion that you might expect: A=1, B=2, ..., Z=26. We also assign 0 to the space. This allows us to encode the message as,

$$M = (1, 12, 7, 5, 2, 18, 1, 0, 18, 21, 12, 26).$$

Let $p = 5$ and $q = 11$; then $N = 55$. Let $e = 3$. Is $e \in \mathbb{Z}^*_{\varphi(N)}$? We know that

$$\begin{aligned}\gcd(3, \varphi(N)) &= \gcd(3, \varphi(5) \cdot \varphi(11)) = \gcd(3, 4 \times 10) \\ &= \gcd(3, 40) = 1;\end{aligned}$$

Definition 5·40 and Lemma 5·38 show that, yes, $e \in \mathbb{Z}^*_{\varphi(n)}$.

Encrypt by computing m^e for each $m \in M$:

$$\begin{aligned}C &= \left(1^3, 12^3, 7^3, 5^3, 2^3, 18^3, 1^3, 0^3, 18^3, 21^3, 12^3, 26^3\right) \\ &= (1, 23, 13, 15, 8, 2, 1, 0, 2, 21, 23, 31).\end{aligned}$$

A snooper who intercepts C and tries to read it as a plain message would encounter several difficulties. First, it contains 31, a number that does not fall in the range 0 and 26. If he gave that number the symbol _, he would see

AWMOHBA BUW_

which is not an obvious encryption of ALGEBRA RULZ.

The inverse of $3 \in \mathbb{Z}^*_{\varphi(N)}$ is $d = 27$. (We could compute this using Corollary 5·47, but it's not hard to see that $3 \times 27 = 81$ and $[81]_{40} = [1]_{40}$.) Decrypt by computing c^d for each $c \in C$:

$$\begin{aligned}D &= \left(1^{27}, 23^{27}, 13^{27}, 15^{27}, 8^{27}, 2^{27}, 1^{27}, 0^{27}, 2^{27}, 21^{27}, 23^{27}, 31^{27}\right) \\ &= (1, 12, 7, 5, 2, 18, 1, 0, 18, 21, 12, 26).\end{aligned}$$

Trying to read this as a plain message, we have

ALGEBRA RULZ.

Doesn't it?

Encrypting messages letter-by-letter is absolutely unacceptable for security. For a stronger approach, letters should be grouped together and converted to integers. For example, the first four letters of the secret message above are

ALGE

and we can convert this to a number using any of several methods; for example

$$\text{ALGE} \quad \to \quad 1 \times 26^3 + 12 \times 26^2 + 7 \times 26 + 5 = 25,785.$$

The integers to encrypt here are larger than 55, so we need larger values for p and q. This is too burdensome to compute by hand, so you want a computer to help. We give an example in the exercises.

RSA is an example of a *public-key cryptosystem.* That means that person A broadcasts to the world, "Anyone who wants to send me a secret message can use the RSA algorithm with values $N = \ldots$ and $e = \ldots$." So a snooper knows the method, the modulus, N, and the encryption key, e!

If the snooper knows the method, N, and e, how can RSA be safe? To decrypt, the snooper needs to compute $d = e^{-1} \in \mathbb{Z}^*_{\varphi(N)}$. Corollary 5·47 tells us that computing d is merely a matter of computing $e^{\varphi(N)-1}$, which is easy if you know $\varphi(N)$. The snooper also knows that $N = pq$, where p and q are prime. So, decryption should be a simple matter of factoring $N = pq$ and applying Lemma 5·50 to obtain $\varphi(N) = (p-1)(q-1)$. Right?

Well, yes *and* no. Typical implementations choose *very* large numbers for p and q, many digits long, and there is *no known method* of factoring a large integer "quickly" — *even when you know that it factors as the product of two primes!* In addition, a careful science to choosing p and q makes it hard to determine their values from N and e.

As it is too time-consuming to perform even easy examples by hand, a computer algebra system becomes necessary to work with examples. The end of this section lists programs to help you perform these computations in the Sage and Maple computer algebra systems. The programs are:

- `scramble`, which accepts as input a plaintext message like "ALGEBRA RULZ" and turns it into a list of integers;
- `descramble`, which accepts as input a list of integers and turns it into plaintext;
- `en_de_crypt`, which encrypts or decrypts a message, depending on whether you feed it the encryption or decryption exponent.

Examples of usage:

- in Sage:
  - to determine the list of integers M, type `M = scramble("ALGEBRA RULZ")`

  - to encrypt M, type

```
C = en_de_crypt(M,3,55)
```

  - to decrypt C, type

```
en_de_crypt(C,27,55)
```

- in Maple:

  - to determine the list of integers M, type `M := scramble("ALGEBRA RULZ");`
  - to encrypt M, type

```
C := en_de_crypt(M,3,55);
```

  - to decrypt C, type

```
en_de_crypt(C,27,55);
```

Question 5·58 .

The phrase

$$[574, 1, 144, 1060, 1490, 0, 32, 1001, 574, 243, 533]$$

is the encryption of a message using the RSA algorithm with the numbers $N = 1535$ and $e = 5$. You will decrypt this message.

(a) Factor N.

(b) Compute $\varphi(N)$.

(c) Find the appropriate decryption exponent.

(d) Decrypt the message.

Question 5·59 .

In this exercise, we encrypt a phrase using more than one letter in a number.

(a) Rewrite the phrase GOLDEN EAGLES as a list M of three positive integers, each of which combines four consecutive letters of the phrase.

(b) Find two prime numbers whose product is larger than the largest number you would get from four letters.

(c) Use those two prime numbers to compute an appropriate N and e to encrypt M using RSA.

(d) Find an appropriate d that will decrypt M using RSA.

(e) Decrypt the message to verify that you did this correctly.

Theory

Now, *why* does the RSA algorithm work?

Proof of the RSA algorithm. Let $c \in C$. By definition of C, $c = m^e \in \mathbb{Z}_N^*$ for some $m \in M$. We need to show that $c^d = (m^e)^d = m$.

Since $[e] \in \mathbb{Z}_{\varphi(N)}^*$, which is a group under multiplication, we know that it has an inverse element, $[d]$. That is, $[de] = [d][e] = [1]$. By Lemma 4·90, $\varphi(N) \mid (1 - de)$, so we can find $b \in \mathbb{Z}$ such that $b \cdot \varphi(N) = 1 - de$, or $de = 1 - b\varphi(N)$.

We claim that $[m]^{de} = [m] \in \mathbb{Z}_N$. To do this, we will show two subclaims about the behavior of the exponentiation in $\mathbb{Z}_p$ and $\mathbb{Z}_q$.

Claim 5·1. $[m]^{de} = [m] \in \mathbb{Z}_p$.

If $p \mid m$, then $[m] = [0] \in \mathbb{Z}_p$. Without loss of generality, $d, e \in \mathbb{N}^+$, so

$$[m]^{de} = [0]^{de} = [0] = [m] \in \mathbb{Z}_p.$$

Otherwise, $p \nmid m$. Recall that p is irreducible, so $\gcd(m, p) = 1$. By Euler's Theorem,

$$[m]^{\varphi(p)} = [1] \in \mathbb{Z}_p^*.$$

Recall that $\varphi(N) = \varphi(p)\varphi(q)$; thus,

$$[m]^{\varphi(N)} = [m]^{\varphi(p)\varphi(q)} = \left([m]^{\varphi(p)}\right)^{\varphi(q)} = [1].$$

Thus, in $\mathbb{Z}_p^*$,

$$\begin{aligned}[m]^{de} &= [m]^{1-b\varphi(N)} = [m] \cdot [m]^{-b\varphi(N)} \\ &= [m]\left([m]^{\varphi(N)}\right)^{-b} = [m] \cdot [1]^{-b} = [m].\end{aligned}$$

As p is irreducible, Any element of $\mathbb{Z}_p$ is either zero or in $\mathbb{Z}_p^*$. We have considered both cases; hence,

$$[m]^{de} = [m] \in \mathbb{Z}_p.$$

Claim 5·2. $[m]^{1-b\varphi(N)} = [m] \in \mathbb{Z}_q$.

The argument is similar to that of the first claim.

Since $[m]^{de} = [m]$ in both $\mathbb{Z}_p$ and $\mathbb{Z}_q$, properties of the quotient groups $\mathbb{Z}_p$ and $\mathbb{Z}_q$ tell us that $\left[m^{de} - m\right] = [0]$ in both $\mathbb{Z}_p$ and $\mathbb{Z}_q$ as well. In other words, both p and q divide $m^{de} - m$. You will show in Question 5.3 that this implies that N divides $m^{de} - m$.

From the fact that N divides $m^{de} - m$, we have $[m]_N^{ed} = [m]_N$. Thus, computing $(m^e)^d$ in $\mathbb{Z}_{\varphi(N)}$ gives us m. □

Question 5·3 .

Let $m, p, q \in \mathbb{Z}$ and suppose that $\gcd(p, q) = 1$.

(a) Show that if $p \mid m$ and $q \mid m$, then $pq \mid m$.

(b) Explain why this completes the proof of the RSA algorithm; that is, since p and q both divide $m^{de} - m$, then so does N.

Sage programs

The following programs can be used in Sage to help make the amount of computation involved in the exercises less burdensome:

```
def scramble(s):
  result = []
  for each in s:
    if ord(each) >= ord("A") \
        and ord(each) <= ord("Z"):
      result.append(ord(each)-ord("A")+1)
    else:
      result.append(0)
  return result


def descramble(M):
  result = ""
  for each in M:
    if each == 0:
      result = result + " "
    else:
      result = result + chr(each+ord("A") - 1)
  return result


def en_de_crypt(M,p,N):
  result = []
  for each in M:
    result.append((each^p).mod(N))
  return result
```

Maple programs

The following programs can be used in Maple to help make the amount of computation involved in the exercises less burdensome:

```
scramble := proc(s)
  local result, each, ord;
  ord := StringTools[Ord];
  result := [];
  for each in s do
    if ord(each) >= ord("A")
        and ord(each) <= ord("Z") then
      result := [op(result),
        ord(each) - ord("A") + 1];
    else
      result := [op(result), 0];
    end if;
  end do;
  return result;
end proc:


descramble := proc(M)
  local result, each, char, ord;
  char := StringTools[Char];
  ord := StringTools[Ord];
  result := "";
  for each in M do
    if each = 0 then
      result := cat(result, " ");
    else
      result := cat(result,
        char(each + ord("A") - 1));
    end if;
  end do;
  return result;
end proc:


en_de_crypt := proc(M,p,N)
  local result, each;
  result := [];
  for each in M do
    result := [op(result), (each^p) mod N];
  end do;
```

```
    return result;
  end proc:
```

Chapter 6

Factorization

This chapter builds up some basic algorithms for factoring polynomials. This is actually a tricky subject, so we focus first on some theory before discussing the practice. We will see in Sections 6·1 and 6·2 that factorization is tied to ideals. To keep things simple, we focus on a special kind of ring where factorization is deterministic; Section 6·3 introduces the relevant structure.

The typical trick is to factorize modulo a prime, then reconstruct the integer factorization; this prequires a deeper study of finite fields than the one we had in Section 3·4, which we address in Sections 6·5 and 6·6. That finally gets us to the point where Section 6·7 can describe algorithms for factorization over a field field and Section 6·8 can outline how to approach factorization in $\mathbb{Z}[x]$.

Remark 6·1. In this chapter, every "generic" ring is an integral domain, unless otherwise specified. Thus, it is commutative, has a multiplicative identity, and lacks zero divisors.

Before proceeding, it will be very useful to observe that we make heavy use of Lemma 4·43. Please review that.

6·1 A wrinkle in "prime"

We said earlier that even though the properties of being "prime" and "irreducible" coincide for integers, this is not true in a general ring. This section shows why.

Prime and irreducible: a distinction

Recall Definition 3·20,

> Suppose $r \in R$ is an element of a commutative ring, and r is not a unit. We say that r **factors over** R if we can find $s, t \in R$ such that $r = st$ and neither s nor t is a unit. Otherwise, r is **irreducible**.

Example 6·2. Consider the ring $\mathbb{Q}[x]$.

- The only units are the rational numbers, since no polynomial of degree at least one has a multiplicative inverse that is also a polynomial.

- $x + q$ is irreducible for every $q \in \mathbb{Q}$.
- x^2 is not irreducible, since $x^2 = x \cdot x$.
- $x^2 + q$ is irreducible for every positive $q \in \mathbb{Q}$.

Recall now the definition of "prime" in Definition 5·29,

> A positive integer p is **prime** if $p \neq 1$ and for any two integers a, b we have $p \mid ab \implies p \mid a$ or $p \mid b$.

Fact 5·31 told us that

> An integer is prime if and only if it is irreducible.

This coincidence is because the integers are a *special* sort of ring. In this section we explore rings where the two definitions do not coincide. We start by generalizing the definition of prime:

Definition 6·3. Suppose $p \in R$ is not a unit. We say that p is **prime** if, whenever we find $a, b \in R$ such that $p \mid ab$, then $p \mid a$ or $p \mid b$.

Prime and irreducible: a difference

Unexpected things happen when you look at rings that involve i. For instance, the set of **Gaussian integers** is

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}.$$

Question 6·4 .
Show that $\mathbb{Z}[i]$ is a ring and an integral domain, but not a field.

Question 6·5 .
Show that $\mathbb{Z}[i]$ is isomorphic to the lattice structure of Section 1·5. Explain why this means we can divide with quotient and remainder in $\mathbb{Z}[i]$, so it makes sense to speak of divisibility, irreducible elements, and so forth in $\mathbb{Z}[i]$.

The number 2 is no longer irreducible in $\mathbb{Z}[i]$:

$$2 = (1 + i)(1 - i).$$

Let's see if it will factor further. Suppose $1 + i$ factors as $(a + bi)(c + di)$. Expand the product to obtain the equation

$$1 + i = (ac - bd) + i(ad + bc).$$

The real and complex parts must be equal, giving us the system of equations

$$\begin{aligned} ac - bd &= 1 \\ ad + bc &= 1. \end{aligned}$$

Let's refine this relationship between a, b, c, d. Eliminate b by multiplying the first equation by c and the second equation by d, then subtracting:

$$\left.\begin{array}{l} ac^2 - bcd = c \\ ad^2 + bcd = d \end{array}\right\} \implies a\left(c^2 + d^2\right) = c + d \implies a = \frac{c+d}{c^2+d^2}.$$

By definition, a is an integer, so $c^2 + d^2$ must divide $c + d$, so either $c + d = 0$ or $c^2 + d^2 \leqslant |c + d|$.

If $c + d = 0$, then $c = -d$. Reconsider the first equation in the original system: substitution gives $ac + bc = 1$, so $c(a + b) = 1$. These are *integers,* so $c = \pm 1$ and $d = \mp 1$, giving us the factorization we already had.

On the other hand, suppose $c + d \neq 0$; then $c^2 + d^2 \leqslant |c + d|$. As c and d are also integers, which are less than their squares, we have $|c + d| \leqslant |c^2 + d^2| = c^2 + d^2$. These two inequalities imply $c + d = c^2 + d^2$, which is possible only if $c, d \in \{0, \pm 1\}$; any other integers give $c^2 > c$ or $d^2 > d$.

Consider the following cases.

- We cannot have $c = d = 0$, as that would make the original equation false: $1 + i = (a + bi)(c + di) = 0$.
- Suppose $c = \pm 1$.
  - If $d = 0$, then $c + di = \pm 1$, so $1 + i = (a + bi) \cdot \pm 1$. This factorization of $1 + i$ involves a unit, called a "trivial factorization". Those don't count against the definition of a prime element. (If you doubt me, reread the definition.)
  - If $d = 1$, then either $c + di = 1 + i$ and $a + bi = 1$, a trivial factorization, or $c + di = -1 + i$ and $a + bi = -i$. This only *looks* non-trivial, since $-i$ has a multiplicative inverse in $\mathbb{Z}[i]$. (See Question 6.6.)
  - If $d = -1$, then either $c + di = -1 - i = -(1 + i)$ and $a + bi = -1$, a trivial factorization, or $c + di = 1 - i$ and $a + bi = i$. This only *looks* non-trivial, since i has a multiplicative inverse in $\mathbb{Z}[i]$.

Question 6·6 .
What are the inverses of i and $-i$ in $\mathbb{Z}[i]$?

Recall what we wrote after Definition 3·20: units don't count in factorization, because everything factors with units. We don't consider $2 = (-1) \times (-2)$ to be different factorizations, because, after all, $-1 \times -1 = 1$. In the same way, we won't consider $1 + i = i(1 - i) = -i(-1 + i)$ to be different factorizations, because after all $i \times (-i) = 1$. To call to mind this point, we add a new term to our growing vocabulary:

Definition 6·7. Let R be a commutative ring with unity, not necessarily an integral domain, and $a, b \in R \setminus \{0\}$. We say that a and b are **associates** if $a \mid b$ and $b \mid a$.

Example 6·8. In $\mathbb{Q}[x]$, $4x^2 + 6$ and $6x^2 + 9$ are associates, since $4x^2 + 6 = \frac{2}{3}(6x^2 + 9)$, and $\frac{2}{3}$ is a unit. They are *not* associates in $\mathbb{Z}[x]$.

Question 6·9 .

(a) Explain why 2 and 7 are not associates in $\mathbb{Z}$.

(b) Explain why the only associate of 7 in $\mathbb{Z}$ is -7.

(c) Show that, in an integral domain, a and b are associates if and only if $a = bu$, where u is a unit.

(d) Explain why 2 and 4 are not associates in $\mathbb{Z}$, but they are in $\mathbb{Z}_6$.

Remember that in this chapter, a generic ring is an integral domain, so we will typically treat the characterization of Question 6.9 as if it were the definition of an associate.

Question 6·10 .

(a) Show that $2/3$ and $11/17$ are associates in $\mathbb{Q}$.

(b) Show that 2 and 5 are associates in $\mathbb{Z}_7$.

(c) Show that a ring R is a field if and only if every non-zero element is an associate of every other non-zero element.

In the Gaussian integers, i is a unit, so $1 + i$ and $1 - i = i(1 + i)$ are associates. The only factorizations of $1 + i$ involve associates, so $1 + i$ is irreducible.

Question 6·11 .
Show that $1 - i$ is also irreducible.

On the other hand, consider the ring

$$\mathbb{Z}\left[i\sqrt{5}\right] = \left\{a + bi\sqrt{5} : a, b \in \mathbb{Z}\right\}.$$

It isn't hard to verify that $\mathbb{Z}\left[i\sqrt{5}\right]$ is also a ring, and additionally that

$$6 = 2 \times 3 = \left(1 + i\sqrt{5}\right)\left(1 - i\sqrt{5}\right).$$

Question 6·12 .
Verify that $\mathbb{Z}\left[i\sqrt{5}\right]$ is a ring and an integral domain.

Question 6·13 .
Show that 2, 3, $1 + i\sqrt{5}$, and $1 - i\sqrt{5}$ are irreducible in $\mathbb{Z}\left[i\sqrt{5}\right]$.

This has an amazing consequence:

Integers factor uniquely into irreducibles in $\mathbb{Z}$, but not in $\mathbb{Z}\left[i\sqrt{5}\right]$!

Why is factorization unique in $\mathbb{Z}$, but not in $\mathbb{Z}\left[i\sqrt{5}\right]$? If you look back at the proof of unique factorization of integers, you'll notice that we used the equivlance of "irreducible" and "prime" to infer that the irreducible p_1 divided q_1. In the equation above,

$$2 \times 3 = \left(1 + i\sqrt{5}\right)\left(1 - i\sqrt{5}\right),$$

all four factors are irreducible, *but clearly not prime!* After all, if $2 \mid \left(1 + i\sqrt{5}\right)$, we could find $a + bi\sqrt{5}$ such that

$$2\left(a + bi\sqrt{5}\right) = 1 + i\sqrt{5},$$

or,

$$2a = 1 \quad \text{and} \quad 2b = 1,$$

neither of which is possible if a and b are integers, which they must be in $\mathbb{Z}\left[i\sqrt{5}\right]$. So the property of prime ring elements must be distinguished from that of irreducible ring elements.

So irreducible elements of integral domains need not be prime. On the other hand, prime elements of integral domains *are* irreducible.

Question 6·14 .

Let R be an integral domain. Show that if $p \in R$ is prime, then it is also irreducible.

Question 6·15 .

Show that:

- For any $n \in \mathbb{N}^+$, the ring $\mathbb{Z}_n$ has no irreducible or prime elements unless n is a power of a prime.
- If n is a power of a prime p, then multiples of p that are not multiples of p^2 are both irreducible and prime; moreover, they are associates.

Definition 6·16. The **norm** of a Gaussian integer $a + bi$ is $a^2 + b^2$.

Question 6·17 .

Show that:

(a) irreducible elements of $\mathbb{Z}[i]$ are prime;

(b) if $z = xy$ in $\mathbb{Z}[i]$ is a nontrivial factorization of z, then the norms of x and y are each smaller than the norm of z;

(c) every element of $\mathbb{Z}[i]$ factors into irreducibles; and

(d) these factorizations are unique up to units.

Hint: For (a), you will need a Bézout-like identity, and then you can imitate the proof for integers. You are helped in your quest for a Bézout-like identity by the fact that Question 6.5 gives you division of Gaussian integers. For (b), show also that the norm of z is the product of the norm of x and the norm of y. For (c), use (b) and the Well-Ordering Principle. For (d), imitate the proof for uniqueness of factorization of integers.

Factors are divisors, and greatest common divisors will prove useful in our search for factors. However, we have to define this term a little differently, since not all rings have a linear ordering.

Definition 6·18. Let R be a ring, and $a, b \in R$. Suppose we can find $d \in R$ such that d divides both a and b, *and* for any $r \in R$ that divides both a and b, we also have $r \mid d$. We call d a **greatest common divisor** of a and b.

What makes d "greatest" is that it sits at the top of a tree of divisibilities. Don't get the wrong idea; d might not be alone! At the very least, its associates will sit next to it at the top of the tree.

Example 6·19. To see how you need to be careful with these ideas, consider $\mathbb{Z}_{14}$. Certainly $2 \mid 6$ and $2 \mid 8$, so 2 is a common divisor of 6 and 8. Is it the *greatest* such? Looking just at 6, by congruence we know that $6 \equiv 20 \equiv 34 \equiv \cdots$. Notice that $5 \mid 20$, so $5 \mid 6$. We can likewise show $5 \mid 8$. Is 5 a "greater" common divisor than 2? No, 5 is actually a *unit:* $5 \times 3 \equiv 1$. Because of that, we automatically get $5 \mid 2$; for instance,

$$2 \equiv 1 \times 2 \equiv (5 \times 3) \times 2 = 5 \times (3 \times 2) = 5 \times 6.$$

So 6 actually divides 2, as well… which means 6 divides 8! Likewise, $8 \times 2 = 16 \equiv 2$, so $8 \mid 2$.

Question 6·20 .
Show that in a principal ideal domain R:

(a) $\langle d \rangle = \langle a, b \rangle$, making d a greatest common divisor of a and b;

(b) there exist $r, s \in R$ such that $d = ra + sb$; and

(c) if both c and d are greatest common divisors of a and b, then c and d are associates.

6·2 The ideals of factoring

The link between divisibility and principal ideals in Lemma 4·43 implies that we can rewrite Definition 6·7 in terms of ideals. We start with the facts that (a) it's trivial to obtain the identity from a unit, and hence obtain the entire ring; and (b) since associates differ only by a unit, their ideals shouldn't differ at all.

Theorem 6·21. *Let R be an integral domain, and let $a, b \in R \setminus \{0\}$.*

(A) a is a unit if and only if $\langle a \rangle = R$.

(B) a and b are associates if and only if $\langle a \rangle = \langle b \rangle$.

Example 6·22. This theorem gives us an alternate route to showing that some ring elements are units or associates (or not). In the Gaussian integers, $3 \notin \langle 1 + i \rangle$, so $1 + i$ is not a unit.

It likewise allows us to decide when two ideals are equal. Since $-i(1 + i) = (1 - i)$, and $-i$ is a unit, $\langle 1 + i \rangle = \langle 1 - i \rangle$.

Proof of Theorem 6·21 on the preceding page. (A) This is a straightforward chain: a is a unit if and only if there exists $b \in R$ such that $ab = 1_R$, which is true if and only if $1_R \in \langle a \rangle$, which is true if and only if $R = \langle a \rangle$ (Questions 4.34 and 4.35).

(B) Assume that a and b are associates. Let $c \in R$ be a unit such that $a = bc$. By definition, $a \in \langle b \rangle$. Since any $x \in \langle a \rangle$ satisfies $x = ar = (bc)\, r = b\,(cr) \in \langle b \rangle$, we see that $\langle a \rangle \subseteq \langle b \rangle$. In addition, we can rewrite $a = bc$ as $ac^{-1} = b$, so a similar argument yields $\langle b \rangle \subseteq \langle a \rangle$.

Conversely, assume $\langle a \rangle = \langle b \rangle$. By definition, $a \in \langle b \rangle$, so there exists $c \in R$ such that $a = bc$. Likewise, $b \in \langle a \rangle$, so there exists $d \in R$ such that $b = ad$. By substitution, $a = bc = (ad)\, c$. Use the associative and distributive properties to rewrite this as $a\,(1 - dc) = 0$. By hypothesis, $a \neq 0$; since we are in an integral domain, $1 - dc = 0$. Rewrite this as $1 = dc$; we see that c and d are units, which implies that a and b are associates. □

Remark. The proof requires R to be an integral domain in order to show (B). For a counterexample, consider $R = \mathbb{Z}_6$; we have $\langle 2 \rangle = \langle 4 \rangle$, but $2 \cdot 2 = 4$ and $4 \cdot 2 = 2$. Neither 2 nor 4 is a unit, so 2 and 4 are not associates. Strange things happen with zero divisors!

Question 6·23 .
Show that in an integral domain, factorization terminates iff every ascending sequence of principal ideals $\langle a_1 \rangle \subseteq \langle a_2 \rangle \subseteq \cdots$ is eventually stationary; that is, for some $n \in \mathbb{N}^+$, $\langle a_i \rangle = \langle a_{i+1} \rangle$ for all $i \geqslant n$.

Ideals of irreducible and prime elements

What about prime or irreducible elements of a ring? We'll preface the result with an example that leads to two new definitions.

Start with an irreducible element; for instance, $2 \in \mathbb{Z}$. Let $A = \langle 2 \rangle$. What can we say about it? No other integer divides it, so Lemma 4·43 suggests that no other ideal can contain it — aside from $\mathbb{Z}$ itself, naturally. By definition, $\langle 2 \rangle$ is the smallest ideal that contains 2, but it is also the largest *proper* ideal that contains 2.

Definition 6·24. Let I be an ideal in an integral domain R. If $I \subsetneq R$ and no other ideal of R contains I, we call I a **maximal ideal**.

For prime elements, it might be more instructive to consider first an integer that is *not* prime, $6 \in \mathbb{Z}$. The fact that it is not prime means we can find two integers a and b such that $6 \mid ab$ but $6 \nmid a$ and $6 \nmid b$. For instance, if $a = 3$ and $b = 4$, we see that $6 \mid (3 \times 4)$ but $6 \nmid 3$ and $6 \nmid 4$. Applying Lemma 4·43 again, we see that $\langle 3 \times 4 \rangle \subseteq \langle 6 \rangle$, while $\langle 3 \rangle \nsubseteq \langle 6 \rangle$ and $\langle 4 \rangle \nsubseteq \langle 6 \rangle$. On the other hand, when an integer p *is* prime, we know that if $p \mid ab$, then $p \mid a$ or $p \mid b$; in terms of Lemma 4·43, we would say that if $\langle ab \rangle \subseteq \langle p \rangle$, then $\langle a \rangle \subseteq \langle p \rangle$ or $\langle b \rangle \subseteq \langle p \rangle$.

This is not especially remarkable, but **we can say something stronger!** Recall from Question 4.39 that if A and B are ideals, then

$$AB = \left\{ \sum_{i=1}^{n} a_i b_i : n \in \mathbb{N}^+, a_i \in A, b_i \in B \right\}$$

is also an ideal. A moment ago, we looked at $\langle ab \rangle$ when referring to a prime element p. What of $\langle a \rangle \langle b \rangle$? This is actually a *larger* ideal; for instance, you could have solved Question 4.40 by looking at

$$\langle 6 \rangle \langle 9 \rangle = \langle 3 \rangle ;$$

after all, $\langle 6 \rangle \subseteq \langle 3 \rangle$ and $\langle 9 \rangle \subseteq \langle 3 \rangle$ just by using Lemma 4·43, which easily gives us $\langle 6 \rangle \langle 9 \rangle \subseteq \langle 3 \rangle$, whereas

$$3 = -1 \times 6 + 1 \times 9 \in \langle 6 \rangle \langle 9 \rangle ,$$

which easily gives us $\langle 6 \rangle \langle 9 \rangle \supseteq \langle 3 \rangle$. So $\langle 6 \rangle \langle 9 \rangle = \langle 3 \rangle$, but $\langle 6 \times 9 \rangle = \langle 54 \rangle$, period, full stop, etc. In fact, in the integers we can say that $\langle a \rangle \langle b \rangle = \langle \gcd(a, b) \rangle$.

Question 6·25 .

Why can we say that:

(a) $\langle 6 \rangle \subseteq \langle 3 \rangle$ and $\langle 9 \rangle \subseteq \langle 3 \rangle$ gives us $\langle 6 \rangle \langle 9 \rangle \subseteq \langle 3 \rangle$? (perhaps not as "easily" as I claim above)

(b) In the integers, $\langle a \rangle \langle b \rangle = \langle \gcd(a, b) \rangle$?
Hint: As with Question 4.40, think about Bézout's Identity.

We will carry this *stronger* property of primes with us from $\mathbb{Z}$ to any integral domain.

Definition 6·26. Let P be a proper ideal of an integral domain R. If, for any two ideals A and B of R, we find that $AB \subseteq P$ implies $A \subseteq P$ or $B \subseteq P$, we call P a **prime ideal**.

Here's another example.

Example 6·27. Let $R = \mathbb{Z}[x, y]$ and $P = \langle x \rangle$. Let $A, B \lhd R$ such that $A \nsubseteq P$ but $AB \subseteq P$. By definition, any $f \in AB$ has the form $f = \sum_{i=1}^{n} g_i h_i$ where each $g_i \in A$ and $h_i \in B$. By inclusion, $f \in P$, as well, so $x \mid f$. This means *x divides every term of f, as well.* (That's true for monomials, but not for arbitrary polynomials; for instance, this would not be true of $x + 1$.)

Let $g \in A \backslash P$ and $h \in B$. By definition, $x \nmid g$, so g has at least one term t such that $x \nmid t$. Let u be any term of h; by polynomial multiplication, a ctu is a term of $f = gh$ for some $c \in \mathbb{Z}$. By definition of AB, $f \in AB$, so, as pointed out a moment ago, $x \mid tu$. By hypothesis, $x \nmid t$, so $\deg_x t = 0$, so

$$\deg_x u = \deg_x tu - \deg_x t = \deg_x tu \geqslant 1.$$

By definition of divisibility, $x \mid u$. Now, u was an arbitrary term of h, so x divides every term of h, which means $x \mid h$, so $h \in P$. We chose h arbitrarily from B, so every polynomial of B is also in P. By definition, $B \subseteq P$.

We have shown that if $A, B \lhd R$, $AB \subseteq P$, and $A \nsubseteq P$, then $B \subseteq P$. Hence $P = \langle x \rangle$ is a prime ideal.

Theorem 6·28. *Let R be an integral domain, and let $a, b \in R \backslash \{0\}$.*

(A) In a principal ideal domain, a is irreducible if and only if $\langle a \rangle$ is maximal.

(B) In a principal ideal domain, a is prime if and only if $\langle a \rangle$ is prime.

Proof. (A) Assume that R is a principal ideal domain, and suppose first that a is irreducible. Let B be an ideal of R such that $\langle a\rangle \subseteq B \subseteq R$. Since R is a principal ideal domain, $B = \langle b\rangle$ for some $b \in R$. Since $a \in B = \langle b\rangle$, $a = rb$ for some $r \in R$. By definition of irreducible, r or b is a unit. If r is a unit, then by definition, a and b are associates, and by part (B) of Theorem 6·21, $\langle a\rangle = \langle b\rangle = B$. Otherwise, b is a unit, and by part (A) of the same Theorem, $B = \langle b\rangle = R$. Since $\langle a\rangle \subseteq B \subseteq R$ implies $\langle a\rangle = B$ or $B = R$, we can conclude that $\langle a\rangle$ is maximal.

For the converse, we show the contrapositive. Assume that a is not irreducible; then there exist $r, b \in R$ such that $a = rb$ and neither r nor b is a unit. Thus $a \in \langle b\rangle$ and by Lemma 4·43 and part (B) of Theorem 6·21, $\langle a\rangle \subsetneq \langle b\rangle \subsetneq R$. In other words, $\langle a\rangle$ is not maximal. By the contrapositive, then, if $\langle a\rangle$ is maximal, then a is irreducible. □

Question 6·29 .
Show part (B) of the theorem.

The discussion above did *not* require that R be a principal ideal domain to show that if $\langle a\rangle$ is maximal, then a is irreducible. This remains true even when R is not a principal ideal domain.

On the other hand, it can happen that a is irreducible when R is not a principal ideal domain, but $\langle a\rangle$ is not maximal. To see why, consider *any* ring R, and its bivariate polynomial ring $R[x, y]$. Example 4·52 on page 136 shows that this is not a principal ideal domain, *even if R is!* The element x is irreducible, but $\langle x\rangle \subsetneq \langle x, y\rangle \subsetneq R[x, y]$, so $\langle x\rangle$ is not maximal.

In a similar way, your proof of part (B) should have shown that if $\langle a\rangle$ is prime, then a is prime even if R is not a principal ideal domain. The converse, however, need not be true.

In any case, we have the following result.

Theorem 6·30. *Let R be an integral domain, and let $p \in R$. If $\langle p\rangle$ is maximal, then p is irreducible, and if $\langle p\rangle$ is prime, then p is prime.*

We can take this a little further.

Theorem 6·31. *Maximal ideals are always prime, even if you are not in an integral domain.*

Proof. Let R be a ring, M a maximal ideal of R, and A, B ideals of R. Suppose that $AB \subseteq M$.

By way of contradiction, suppose further that $A, B \nsubseteq M$. That means we can choose $a \in A\backslash M$ and $b \in B\backslash M$; do so. Note that $ab \in AB \subseteq M$.

For any $r \in R$ we can write $\langle r, M\rangle$ for the smallest ideal that contains both r and the elements of M. As a and b are not elements of M, we infer $M \subsetneq \langle a, M\rangle, \langle b, M\rangle$. By maximality of M, we infer $\langle a, M\rangle = \langle b, M\rangle = R$.

Let C be the product ideal of $\langle a, M\rangle$ and $\langle b, M\rangle$. Let $x \in C$; by definition of a product ideal, $x = \sum_{i=1}^{n} y_i z_i$ where $n \in \mathbb{N}^+$, $y_i \in \langle a + M\rangle$, and $z_i \in \langle b + M\rangle$. For any term $y_i z_i$ of this sum, choose $t, u \in M$ such that $y_i = a + t$ and $z_i = b + u$. We have

$$y_i z_i = (a + t)(b + u) = ab + au + bt + tu.$$

We noted above that $ab \in M$, and absorption guarantees $au, bt, tu \in M$. So $y_i z_i \in M$. Closure of ideals under addition means $x \in M$. As x was arbitrary in C, $C \subseteq M$.

On the other hand, recall that $\langle a, M\rangle, \langle b, M\rangle = R$. Let $r \in R$; by inclusion, $r \in \langle a, M\rangle$; by definition of a product ideal, $r = 1 \cdot r \in \langle a, M\rangle \langle b, M\rangle = C$. The arbitrary choice of $r \in R$ implies that $R \subseteq C \subseteq M \subsetneq R$, a contradiction!

The only assumption not forced by the hypotheses was the one that both $A, B \nsubseteq M$. We are forced to conclude that $A \subseteq M$ or $B \subseteq M$. By definition, M is a prime ideal. □

Amazingly, this is true even when the ring is not an integral domain! Combined with Theorem 6·30, we have the following diagram for integral domains:

$$\begin{array}{ccc} \langle p\rangle \text{ maximal} & \Longrightarrow & p \text{ irreducible} \\ \Downarrow & & \\ \langle p\rangle \text{ prime} & \Longrightarrow & p \text{ prime} \end{array}$$

Do the arrows also point in the other directions? We can make a start on answering that.

Question 6·32 .

Show that if p is prime, then $\langle p\rangle$ is also prime. *Hint:* Suppose A, B are ideals of a ring and $AB \subseteq \langle p\rangle$, but $A \nsubseteq \langle p\rangle$. Use an element of $A \setminus \langle p\rangle$ to show that every element of B lies in $\langle p\rangle$.

How are prime and irreducible elements related?

The relationships we have discussed have many useful consequences. Ideals are a powerful enough tool that we can prove quite a few properties about both elements and rings *through their ideals.*

One question that comes to mind is, what is so special about $\mathbb{Z}$, that irreducible elements are prime? After all, it was *not* true about $\mathbb{Z}\left[i\sqrt{5}\right]$! The answer is not obvious if we think only about the properties of the elements per se, but it becomes easier if we think about their ideals. Your eyes should dart immediately to the different hypotheses in the theorems above: to prove one direction, we needed only an integral domain; to prove the other, we needed a *principal ideal domain.*

We have already shown that $\mathbb{Z}$ is a principal ideal domain (Theorem 4·53). Could it be that $\mathbb{Z}\left[i\sqrt{5}\right]$ is not? In the case we studied before, we had $2 \times 3 = \left(1 + i\sqrt{5}\right)\left(1 - i\sqrt{5}\right)$. These elements are all irreducible. In $\mathbb{Z}$, joining two elements in a ring gives us an ideal generated by *one* element, their gcd (see Question 4.37, where you hopefully used Bézout's Identity); since $\gcd(2, 3) = 1$, we have $\langle 2, 3\rangle = \langle 1\rangle$ in $\mathbb{Z}$, so $\langle 2, 3\rangle = \mathbb{Z}$.

It's another story entirely in $\mathbb{Z}\left[i\sqrt{5}\right]$. Consider the ideal $I = \left\langle 2, 1 + i\sqrt{5}\right\rangle$. Both generators are irreducible, and they are not associates, but $1 \notin I$! Hence $I \neq \mathbb{Z}\left[i\sqrt{5}\right]$, and we now have a chain

$$2 \subsetneq I \subsetneq \mathbb{Z}\left[i\sqrt{5}\right].$$

Interestingly, 2 is irreducible, but its ideal is not maximal! On the other hand, the fact that 2 and $1 + i\sqrt{5}$ are irreducible but not associates means *no one element* can generate I. So $\mathbb{Z}\left[i\sqrt{5}\right]$ is not a principal ideal domain!

Question 6·33 .

What are the units of $\mathbb{Z}\left[i\sqrt{5}\right]$? Explain how this shows 2 and $1 + i\sqrt{5}$ are not associates.

Question 6·34 .
We wrote in the discussion above that $1 \notin I$. How do we know this?

We have our criterion for an irreducible element to be prime! Prove it for the general case.

Question 6·35 .
Suppose that R is a principal ideal domain, and $r \in R$. Show that if r is irreducible, then it is prime.

The converse is true even if we are not in a principal ideal domain; see Question 6.14.

You may be wondering why we worked with prime and irreducible elements in the context of integral domains. It may seem intuitive that zero divisors would throw off the properties we expect, but even if not, an ideal you already know provides a direct answer.

Example 6·36. Consider the ring $\mathbb{Z}_6$. This is not an integral domain, so our definition of a "prime" element doesn't apply, but it is not hard to verify that 2 satisfies the requirements of a prime element of $\mathbb{Z}_6$, if such a thing existed:

$$\langle 2 \rangle = \{0, 2, 4\},$$

and if $2 \nmid a$, $2 \nmid b$ then $2 \nmid ab$:

$$1 \times 1,\ 1 \times 3,\ 1 \times 5,\ 3 \times 3,\ 3 \times 5,\ 5 \times 5 \notin \langle 2 \rangle .$$

Alas, 2 is not irreducible; after all, $2 = 8 = 2 \times 2 \times 2$, so 2 factors itself, *even though it isn't a unit!*

We have now answered one question posed at the beginning of the chapter:

- If R is an integral domain, then prime elements are irreducible.
- If R is a principal ideal domain, then irreducible elements are prime.

Because we are generally interested in factoring only for integral domains, many authors restrict the definition of *prime* so that it is defined only in an integral domain. In this case, a prime element is always irreducible, although the converse might not be true, since not all integral domains are principal ideal domains. We went beyond this in order to show, as we did above, *why* it is defined in this way. Since we maintain throughout most of this chapter the assumption that all rings are integral domains, one could shorten this to,

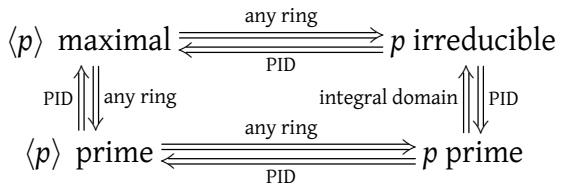
Fact 6·37 (Prime and irreducible elements in integral domains.)**.** A prime element is always irreducible, but an irreducible element is not always prime.

To sum up what we have found:

- We saw in Example 6·27 that in $\mathbb{Z}[x, y]$, the ideal $\langle x \rangle$ is prime, but $\langle x \rangle \subsetneq \langle x, y \rangle \subsetneq \langle 1 \rangle$, so a prime ideal need not be maximal, *even in an integral domain.*

- Likewise, in $\mathbb{Z}[x,y]$, the element x is irreducible, but $\langle x\rangle \subsetneq \langle x,y\rangle \subsetneq \langle 1\rangle$, so an irreducible element need not generate a maximal ideal, *even in an integral domain.*
- We saw in $\mathbb{Z}\left[i\sqrt{5}\right]$ that an irreducible element need not be prime, *even in an integral domain.*
- You will see below that if p is prime, then $\langle p\rangle$ is also prime.

We can, therefore, revise our diagram for integral domains as follows:



Question 6·38 .
Use the theory developed in this section to describe how prime ideals and maximal ideals are related in:

- an integral domain, and
- a principal ideal domain.

6·3 Time to expand our domains

This section considers two ideas essential to factorization: unique factorization and division. You might think that the ability to divide would give you a unique factorization, but on the other hand, the distinction between prime and irreducible elements might also give you pause. Indeed, the ability to divide and the ability to obtain a unique factorization are not quite identical, a fact reflected in the structures of rings with these properties.

Unique factorization domains

The Fundamental Theorem of Arithmetic tells us that every integer factors *uniquely* into a product of irreducible elements. This is not true in every ring; in $\mathbb{Z}\left[-\sqrt{5}\right]$, we factored $6 = 2\cdot 3$ *and* $6 = \left(1+\sqrt{-5}\right)\left(1-\sqrt{-5}\right)$. Since 2, 3, $1+\sqrt{-5}$, and $1-\sqrt{-5}$ are all irreducible in $\mathbb{Z}\sqrt{-5}$, 6 factors two different ways as a product of irreducibles.

Definition 6·39. A ring is a **unique factorization domain** if every nonzero, non-unit $r \in R$ factors into irreducibles $r = p_1^{a_1}p_2^{a_2}\cdots p_m^{a_m}$, and if this factorization is unique up to order and associates.

Another way of saying this is that if r *also* factors into irreducibles $r = q_1^{b_1}q_2^{b_2}\cdots q_n^{b_n}$, then $m = n$ and each p corresponds to a unique q via an associate c, according to the relationship $p = cq$, and the corresponding exponents are also the same, with $a_i = b_j$.

Aside from $\mathbb{Z}$, what are some other unique factorization domains?

Example 6·40. You showed in Question 6.17 that $\mathbb{Z}[i]$ is a unique factorization domain.

Example 6·41. $\mathbb{Z}[x]$ is a unique factorization domain. To see this requires two major steps.

(Existence) Let $f \in \mathbb{Z}[x]$. If the coefficients of f have a common factor, we can factor that out easily; for example, $2x^2+4x = 2(x^2+2x)$. We know that integers have a unique factorization, so we may assume, without loss of generality, that the terms of f have no common factor.

If f is irreducible, then we are done; it has a factorization into irreducibles. Otherwise, we claim it factors into two polynomials *of smaller degree*. After all, if f factors as ag where $\deg g = \deg f$, then we must have $\deg a = 0$. That implies $a \in \mathbb{Z}$, so a is a common factor of f's *coefficients,* a possibility we excluded! So *if* f factors, it factors as $f = gh$, where $\deg g, \deg h < \deg f$. Degrees are natural numbers, and they decrease each time we factor a polynomial further, so Fact 1·51 tells us this process must eventually end with polynomials that do not factor; that is, with irreducibles. Hence f factors into irreducibles; say $f = p_1 \cdots p_m$. Without loss of generality, we may assume that none of the p's are associates.

Of course, having *a* factorization into irreducibles doesn't exclude the possibility of having *more than one* factorization into irreducibles, so we turn our attention to…

(Uniqueness) Suppose we can also factor f into irreducibles as $f = q_1 \cdots q_n$. The coefficients of f are integers, and any integer a corresponds to a rational number $a/1$, so we can consider f as an element of $\mathbb{Q}[x]$. Why would we do this? By Theorem 4·53(C) we know that $\mathbb{Q}[x]$ is a principal ideal domain. You showed in Question 6.35 that irreducible elements of a principal ideal domain are prime. Hence p_1 divides q_j for some $j = 1, \ldots, n$. Without loss of generality, $p_1 \mid q_1$. Since q_1 is also irreducible, p_1 and q_1 are associates; say $p_1 = a_1 q_1$ for some unit a_1. The units of $\mathbb{Q}[x]$ are the nonzero elements of $\mathbb{Q}$, so $a_1 \in \mathbb{Q}\setminus\{0\}$. And so forth; each p_i is an associate of a unique q_j in the product. Without loss of generality, we may assume that p_i is an associate of q_i. This forces $m = n$.

Right now we have p_i and q_i as associates in $\mathbb{Q}[x]$. If we can show that each $a_i = \pm 1$, then we will have shown that the corresponding p_i and q_j are associates in $\mathbb{Z}[x]$ as well, so that $\mathbb{Z}[x]$ is a unique factorization domain. Write $a_1 = b/c$ where $\gcd(b,c) = 1$; we have $p_1 = \frac{b}{c} \cdot q_1$. Rewrite this as $cp_1 = bq_1$. Remember that p_1 and q_1 are *integer* polynomials. What's more, the fact that $\gcd(b,c) = 1$ means we can infer $b \mid p_1$ and $c \mid q_1$ (see below). However, p_1 and q_1 are *irreducible, integer* polynomials, so b and c must be integer units. The only integer units are ± 1, so p_1 and q_1 are associates.

The same argument can be applied to the remaining irreducible factors. Thus, the factorization of f is unique up to order and associates.

This result generalizes to an important class of rings.

Theorem 6·42. *Every principal ideal domain is a unique factorization domain.*

Proof. Let R be a principal ideal domain, and $f \in R$.

(Existence) First we show that f has a factorization. Suppose f is not irreducible; then there exist $r_1, r_2 \in R$ such that $f = r_1 r_2$ and f is an associate of neither. By Theorem 6·21, $\langle f\rangle \subsetneq \langle r_1\rangle$ and $\langle f\rangle \subsetneq \langle r_2\rangle$. If r_1 is not irreducible, then there exist $r_3, r_4 \in R$ such that $r_1 = r_3 r_4$ and r_1 is an associate of neither. Again, $\langle r_1\rangle \subsetneq \langle r_3\rangle$ and $\langle r_1\rangle \subsetneq \langle r_4\rangle$. Continuing in this fashion, we obtain an ascending chain of ideals

$$\langle f\rangle \subsetneq \langle r_1\rangle \subsetneq \langle r_3\rangle \subsetneq \cdots.$$

We step out of this proof a moment to show that such a chain cannot continue indefinitely:

Lemma 6·43. *In any principal ideal domain R, an ascending chain of ideals $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ eventually stabilizes at an ideal B.*

Proof of Lemma 6·43. Let $B = A_1 \cup A_2 \cup A_3 \cup \cdots$. We claim B is an ideal of R. For any $b \in B$ and $r \in R$, we know $b \in A_i$ for some $i \in \mathbb{N}^+$, and since A_i is an ideal of R, $br \in A_i$; by inclusion, $br \in B$. On the other hand, let $c \in B$; we know $c \in A_j$ for some $j \in \mathbb{N}^+$. Let $k = \max(i,j)$; by inclusion, $b, c \in A_k$, which is an ideal, so $b - c \in A_k$, and by inclusion $b - c \in B$. We have shown that B is closed under subtraction, and that it absorbs multiplication from R.

We have established that B is an ideal. By hypothesis, R is a principal ideal domain, so $B = \langle b \rangle$ for some $b \in B$. By definition, $b \in A_i$ for some $i \in \{1, 2, \dots\}$. Every element in B is a multiple of b, so every element in B is also in A_i; that is, $B \subseteq A_i$. But $A_i \subseteq B$ by definition of B. The two sets are therefore equal. Likewise, $A_j = B$ for every $j = i+1, i+12, \dots$. The chain has become

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_{i-1} \subseteq A_i = B = A_{i+1} = A_{i+1} = \cdots.$$

As claimed, the ascending chain of ideals stabilized at B. □

This property of ascending chains of ideals is similar to the Noetherian behavior we observed in $\mathbb{Z}$ and other rings. Indeed, an ascending chain of ideals in $\mathbb{Z}$ corresponds to divisibility and factorization (Lemma 4·43). The Well-Ordering Principle means that integer divisors must eventually end with irreducible factors; thus, an ascending chain of integer ideals must eventually end with a maximal ideal.

Question 6·44 .

Consider the ideal $\langle 180 \rangle \subset \mathbb{Z}$. Use unique factorization to build a chain of ideals $\langle 180 \rangle = \langle a_1 \rangle \subsetneq \langle a_2 \rangle \subsetneq \cdots \subsetneq \langle a_n \rangle = \mathbb{Z}$ such that there are no ideals between $\langle a_i \rangle$ and $\langle a_{i+1} \rangle$. Identify $a_1, a_2, \dots$ clearly.

This property is sufficiently important that we give it a special name. Any ring where an ascending chain of ideals eventually stabilizes is said to satisfy the **ascending chain condition**. We can also say it is a **Noetherian ring**.

Proof of Theorem 6·42, continued: (Still on existence) By Theorem 6·43, a principal ideal domain satisfies the ascending chain condition; thus, the chain

$$\langle f \rangle \subsetneq \langle r_1 \rangle \subsetneq \langle r_3 \rangle \subsetneq \cdots$$

must stabilize eventually. We have already explained that if r_i factors, the chain continues further, so it can stabilize only if we reach an irreducible polynomial. This holds for *each* chain, regardless of whether it starts with $r_1, r_2, r_3, r_4, \dots$. All must terminate with irreducible elements of the ring, which gives us $f = p_1 \cdots p_m$ where each p_i is irreducible.

(Uniqueness) Now we show the factorization is unique. Suppose f also factors as $f = q_1 \cdots q_n$ where each q_j is irreducible. Without loss of generality, $m \leqslant n$. Recall that irreducible elements are prime in a principal ideal domain (Corollary 6.35). Hence p_1 divides one of the q_i; without loss of generality, $p_1 \mid q_1$. However, q_1 is irreducible, so p_1 and q_1 must be

associates; say $c_1 p_1 = q_1$ for some unit $c_1 \in R$. Since we are in an integral domain, we can cancel p_1 and q_1 from $f = f$, obtaining

$$c_1 p_2 \cdots p_m = q_2 \cdots q_n.$$

Since p_2 is irreducible, hence prime, we can continue this process until we conclude with $c_1 c_2 \cdots c_m = q_{m+1} \cdots q_n$. Now, the left hand side is a unit. By definition, irreducible elements are not units, so the right hand side must also be a unit, but that is possible only if there are no more irreducibles on the right hand side; that is, $m = n$. Thus the factorization is unique up to ordering and associates.

We chose an arbitrary element of an arbitrary principal ideal domain R, and showed that it had only one factorization into irreducibles. Thus every principal ideal domain is a unique factorization domain. □

We can likewise extend a result from a previous section.

Question 6·45 .
Show that in a unique factorization domain, irreducible elements are prime.

Corollary 6·46. *In a unique factorization domain:*

- *an element is irreducible iff it is prime; and*
- *an ideal is maximal iff it is prime.*

Euclidean domains

We'd like to define a **Euclidean domain** as a ring with a valid division with quotient and remainder. Once we have a precise notion of such division, we can use the **Euclidean algorithm** to find greatest common divisors — so long as the remainder "shrinks." But how can we decide that the remainder "shrinks", when not all rings have natural orderings?

What we will do is define a **valuation function** v from the nonzero elements of a ring to the naturals, $\mathbb{N}$, satisfying the desirable property

$$v(rs) = v(r)\, v(s) \qquad \text{for all } r, s \in R.$$

Definition 6·47. If R is an integral domain and v is a valuation function on R such that for any $a \in R$ and any $d \in R \setminus \{0\}$, we can find $q, r \in R$ such that $a = qd + r$ and either $r = 0$ or $v(r) < v(d)$, then we call R a **Euclidean domain**.

Example 6·48. In $\mathbb{Z}$, the valuation function is the absolute value: $v(r) = |r|$. Observe that $v(rs) = |rs| = |r| \cdot |s| = v(r)\, v(s)$, and for any $a \in \mathbb{Z}$ and any $d \in \mathbb{Z} \setminus \{0\}$ we can find $q, r \in \mathbb{Z}$ such that $a = qd + r$ and $r = 0$ or $v(r) = |r| < |d| = v(d)$.

Example 6·49. We can also see this in $\mathbb{Z}[i]$. We defined division using the lattice; the valuation function is effectively the norm. We have $a = qd + r$ with $r = 0$ or the norm of r less than the norm of d.

Question 6·50 .
We can adapt the Euclidean algorithm (Theorem) to any Euclidean domain by making just one change: in step 1, replace

1. Let $s = \max(m, n)$ and $t = \min(m, n)$.

by

1. If $v(m) \geqslant v(n)$, let $s = m$ and $t = n$; otherwise, let $s = n$ and $t = m$.

Adapt the original proof of the Euclidean Algorithm to show that this one change does indeed give us an algorithm that terminates correctly in any Euclidean domain.

Question 6·51 .
Building on Question 6.50, explain why Bézout's Lemma also applies in Euclidean domains: if R is a Euclidean domain with valuation function v, $r, s \in R$, and d is a gcd of r and s, then we can find $a, b \in R$ such that $ar + bs = d$, and $v(d) \geqslant v(d')$ for any other common divisor $d' \in R$ of x and y.

Question 6·52 .
Use Bézout's Lemma to show that if $\gcd(b, c) = 1$ and $b \mid ac$, then $b \mid a$. This argument should apply in any Euclidean domain.

Polynomials in one variable also have a division. What is the valuation function when dividing these polynomials? That is, what aspect of polynomials is guaranteed to decrease when you divide them correctly? We use $v(r) = \deg r$.

Question 6·53 .
Use $x^3 + x + 1$ and $x^2 - 1$ as an example of polynomial division: find a quotient and a remainder r with $\deg r < \deg\ (x^2 - 1) = 2$.

However, $\mathbb{Z}[x]$ is *not* a Euclidean domain if the valuation function is $v(r) = \deg r$. After all, if $f = 2$ and $g = x$, we cannot find $q, r \in \mathbb{Z}[x]$ such that $g = qf + r$ and $\deg r < \deg f$. The best we can do is $x = 0 \cdot 2 + x$, but $\deg x > \deg 2$.

Question 6·54 .
Use $\mathbb{Z}[x]$ to show that even if R is a unique factorization domain but not a principal ideal domain, then we cannot always find $r, s \in R$ such that $\gcd(a, b) = ra + sb$ for every $a, b \in R$.

Over a ground *field* $\mathbb{F}$, however, it's another matter.

Fact 6·55. *If $\mathbb{F}$ is a field, then $\mathbb{F}[x]$ is a Euclidean domain with valuation function $v(f) = \deg f$ for all nonzero $f \in \mathbb{F}[x]$.*

Why? The difference in the success of $\mathbb{F}[x]$ and the failure $\mathbb{Z}[x]$ is precisely in that fields contain their multiplicative inverses, whereas in the example above $\mathbb{Z}$ was unable to provide a multiplicative inverse for 2.

To be precise, let $f, g \in \mathbb{F}[x]$. We claim that we can divide f by g using degree for the valuation. If $\deg g > \deg f$, let $q = 0$ and $r = f$, and we have $f = qg + r$ with $v(r) < v(g)$, as claimed. Suppose, then, that $\deg g \leqslant \deg f$. Let a_1 be the leading coefficient of f, b the leading coefficient of g, $m = \deg f$, and $n = \deg g$. Let $c_1 = a_1 b^{-1}$, $\ell = m - n$, and $q_1 = c_1 x^\ell$. Define

$$r_1 = f - q_1 g.$$

By construction, the leading term of $q_1 g$ is

$$\left(a_1 b^{-1}\right) x^\ell \cdot b x^m = a_1 x^{(m-n)+n} = a_1 x^m,$$

the same as the leading term of f. So the leading terms cancel, and $\deg r_1 < \deg f$.

If $\deg r_1 < \deg g$, then let $q = q_1$ and $r = r_1$, and we are done. Otherwise, for $i = 1, 2, \ldots$ let a_{i+1} be the leading coefficient of r_i, $c_{i+1} = a_{i+1} b^{-1}$, $\ell_i = \deg r_i - \deg g$, and $q_{i+1} = c_{i+1} x^\ell$. Define $r_{i+1} = r_i - q_{i+1} g$, and in each case the leading terms will cancel, as above. We obtain a sequence of polynomials f, r_1, r_2, ... whose degrees constitute a decreasing sequence of nonnegative integers. By Fact 1·51, this sequence must eventually stabilize, but the only way it stabilizes is if we can no longer divide by g. That happens only if the remainder eventually is either zero or has a degree smaller than that of g. □

This fact plugs a hole at which we've mostly hinted in the past, without explaining rigorously. In the past, we've pointed out that the Factor Theorem allows us to associate every root α of a polynomial f with a factor $x - \alpha$ of f. That *strongly suggests* that a polynomial has at most n roots (and exactly n roots, if you count multiplicities) but it doesn't guarantee it; after all, irreducibles need not be prime.

Question 6·56 .

Let $f \in \mathbb{F}[x]$ of degree n. Use Fact 6·55 and the Factor Theorem to show that f has at most n roots in $\mathbb{F}$.

Fact 6·57. *If R is a Euclidean domain with valuation function v, r and s are nonzero elements of R, and $r \mid s$, then $v(r) \leqslant v(s)$.*

Proof. Given the hypotheses, we can find $q \in R$ such that $s = qr$. By substitution, $v(s) = v(qr) = v(q)\, v(r)$. These are all positive integers, so $v(r) \leqslant v(s)$. □

Theorem 6·58. *Every Euclidean domain is a principal ideal domain.*

Proof. Let R be a Euclidean domain with respect to v, and let A be any non-zero ideal of R. Let $a_1 \in A$. As long as $A \neq \langle a_i \rangle$, do the following:

- find $b_i \in A \setminus \langle a_i \rangle$;
- let r_i be the remainder of dividing b_i by a_i;
  - notice $v(r_i) < v(a_i)$;
- use the Euclidean algorithm to compute a gcd a_{i+1} of a_i and r_i;

  - notice $v(a_{i+1}) \leqslant v(r_i) < v(a_i)$;

- this means $\langle a_i \rangle \subsetneq \langle a_{i+1} \rangle$; after all,

  - as a gcd, $a_{i+1} \mid a_i$, but
  - $a_i \nmid a_{i+1}$, lest $a_i \mid a_{i+1}$ imply $v(a_i) \leqslant v(a_{i+1}) < v(a_i)$;

- hence, $\langle a_i \rangle \subsetneq \langle a_{i+1} \rangle$ and $v(a_{i+1}) < v(a_i)$.

By Fact 1·51, the sequence $v(a_1) > v(a_2) > \cdots$ cannot continue indefinitely, which means that we cannot compute a_i's indefinitely. Let d be the final a_i computed. If $A \neq \langle d \rangle$, we could certainly compute another a_i, so it must be that $A = \langle d \rangle$. □

Corollary 6·59. *Every Euclidean domain is a unique factorization domain.*

Proof. This is a consequence of Theorem 6·42 and Theorem 6·58. □

The converse is false: $\mathbb{Z}[x]$ is a unique factorization domain, but we saw above that it is not a Euclidean domain. On the other hand, its deficiencies do not extend to $\mathbb{Q}[x]$, or polynomial rings over other fields.

Corollary 6·60. *If $\mathbb{F}$ is a field, then $\mathbb{F}[x]$ is both a principal ideal domain and a unique factorization domain.*

However, the definition of a greatest common divisor that we introduced with Euclidean domains certainly generalizes to unique factorization domains.

Theorem 6·61. *In a unique factorization domain, greatest common divisors are unique up to associates.*

Proof. Let R be a unique factorization domain, and let $f, g \in R$. Let $d, \widehat{d}$ be two gcds of f, g. Let $d = p_1^{a_1} \cdots p_m^{a_m}$ be an irreducible factorization of d, and $\widehat{d} = q_1^{b_1} \cdots q_n^{b_n}$ be an irreducible factorization of $\widehat{d}$. Since d and $\widehat{d}$ are both gcds, $d \mid \widehat{d}$ and $\widehat{d} \mid d$. So $p_1 \mid \widehat{d}$. By Theorem 6.45, irreducible elements are prime in a unique factorization domain, so $p_1 \mid q_i$ for some $i = 1, \ldots, n$. Without loss of generality, $p_1 \mid q_1$. Since q_1 is irreducible, p_1 and q_1 must be associates.

We can continue this argument with $\frac{d}{p_1}$ and $\frac{\widehat{d}}{p_1}$, so that $d = a\widehat{d}$ for some unit $a \in R$. Since d and $\widehat{d}$ are unique up to associates, greatest common divisors are unique up to associates. □

Question 6·62 .

Theorem 6·61 says that gcd's are unique up to associate in every unique factorization domain. Suppose that $P = \mathbb{F}[x]$ for some field $\mathbb{F}$. Since P is a Euclidean domain (Question 6·60), it is a unique factorization domain, and gcd's are unique up to associates (Theorem 6·61). The fact that the base ring is a field allows us some leeway that we do not have in an ordinary unique factorization domain. For any two $f, g \in P$, use the properties of a field to describe a method to define a "canonical" gcd of f and g, and show that this canonical gcd is unique.

Question 6·63 .
Generalize the argument of Example 6·41 to show that for any unique factorization domain R, the polynomial ring $R[x]$ is a unique factorization domain. Explain why this shows that for any unique factorization domain R, the polynomial ring $R[x_1, \ldots, x_n]$ is a unique factorization domain. On the other hand, give an example that shows that if R is not a unique factorization domain, then neither is $R[x]$.

6·4 Field extensions

This section explores the relationship between polynomials, roots, and fields. Let $\mathbb{F}$ be *any* field.

Extending a ring

Let R and S be rings, with $R \subseteq S$ and $s \in S$.

Question 6·64 .
We define $R[s]$ as the smallest ring containing both R and s. Show that:

(a) $R \subseteq R[s] \subseteq S$;

(b) $R = R[s]$ if and only if $s \in R$; and

(c) $R[s] = \left\{ \sum_{i=0}^{k} r_i s^i : k \in \mathbb{N}, r_i \in R \right\}$.

We call $R[s]$ a **ring extension** of R. Sometimes, this is isomorphic to a polynomial ring over R; in this case, s is **transcendental** over R. We won't prove it, but it is fairly well known that e and π are transcendental over $\mathbb{Q}$, so $\mathbb{Q}[e] \cong \mathbb{Q}[\pi] \cong \mathbb{Q}[x]$.

We are not interestined in transcendental extensions. We are interested in the case where $R[s]$ is not isomorphic to $R[x]$; in that case, we call s **algebraic**, as it is the root of a polynomial over R. (This may not be obvious, but we prove it in Theorem 6·66 below.)

Example 6·65. Let $R = \mathbb{R}$, $S = \mathbb{C}$, and $s = i = \sqrt{-1}$. Then $\mathbb{R}[i]$ is a ring extension of $\mathbb{C}$. Moreover, $\mathbb{R}[i]$ is not isomorphic to a polynomial ring over $\mathbb{R}$, since $i^2 + 1 = 0$, but $x^2 + 1 \neq 0$ in $\mathbb{R}[x]$. Every element of $\mathbb{R}[i]$ has the form $a + bi$ for some $a, b \in \mathbb{R}$, so we can view $\mathbb{R}[i]$ as a vector space of dimension 2 over $\mathbb{R}$! The basis elements are $\mathbf{u} = 1$ and $\mathbf{v} = i$, and $a + bi = a\mathbf{u} + b\mathbf{v}$.

We made a rather bold claim here about the isomorphism, so let's pause to verify it before proceeding.

Theorem 6·66. *With R, S, and s defined as above, $R[s] \cong R[x]$ if and only if s is not the root of a polynomial over $R[x]$.*

Proof. Let $\varphi : R[x] \to R[s]$ by $\varphi\left(\sum r_i x^i\right) = \sum r_i s^i$. We claim this is a homomorphism of rings: addition is fairly obvious, and multiplication is harder only because it's a notational disgrace:

$$\varphi\left(\left(\sum r_i x^i\right)\left(\sum a_i x^i\right)\right) = \varphi\left(\sum_{i+j=k}\left(r_i + a_j\right)x^k\right) = \sum_{i+j=k}\left(r_i + a_j\right)s^k = \left(\sum r_i s^i\right)\left(\sum a_i s^i\right).$$

That leaves the question of isomorphism. Suppose that φ is an isomorphism. An isomorphism is one-to-one, so we have $\varphi(0) = 0 = \sum r_i s^i = \varphi\left(r_i x^i\right)$ only if $r_i = 0$ for each i. By definition, s is not a root of a polynomial over $R[x]$. On the other hand, suppose s *is* the root of a nonzero polynomial over $R[x]$; call it $f(x)$, and suppose $f(x) = \sum r_i x^i$. By definition of a root,

$$\varphi(f) = \varphi\left(\sum r_i x^i\right) = \sum r_i s^i = f(s) = 0 = \varphi(0),$$

which shows that φ is not one-to-one. □

Let's see if this result generalizes, at least for fields. For the rest of this section, we let $\mathbb{F}$ and $\mathbb{E}$ be fields, with $\alpha \in \mathbb{E}$. It's helpful to look at polynomials whose leading coefficient is 1.

Extending a field to include a root

Notation 6·67. We write $\mathbb{F}(\alpha)$ for the smallest field containing both $\mathbb{F}$ and α.

Example 6·68. We prove later, in the Fundamental Theorem of Algebra, that $\mathbb{R}[i] = \mathbb{C}$, which is a field. So, $\mathbb{R}[i] = \mathbb{R}(i) = \mathbb{C}$. On the other hand, $\mathbb{Q}\left[\sqrt{2}\right] \subsetneq \mathbb{Q}\left(\sqrt{2}\right) \subsetneq \mathbb{C}$.

Question 6·69 .
Explain why $\mathbb{Q}\left[\sqrt{2}\right] \subsetneq \mathbb{Q}\left(\sqrt{2}\right) \subsetneq \mathbb{C}$.

Theorem 6·72 below generalizes the construction of complex numbers in Section 3·1. When we built the complex numbers, we worked modulo the polynomial $x^2 + 1 \in \mathbb{R}[x]$. That polynomial is irreducible over $\mathbb{R}$, and by Theorem 6·28 you now know that irreducible ring elements generate maximal ideals. Really, then, you were building the quotient ring $\mathbb{R}[x] / \langle x^2 + 1\rangle$, which is modulo a maximal ideal.

Do quotient rings formed by maximal ideals always result in a field? Indeed they do.

Fact 6·70. *Let R be a ring with unity, and M a maximal ideal of R. Then R/M is a field.*

Why? Let $X \in R/M$ be any nonzero coset; choose $x \in R$ such that $X = x + I$. As a *nonzero* coset, $X \neq M$, so $x \notin M$ (Lemma 4·103). We claim that we can find $Y \in R/M$ such that $XY \equiv 1 + M$. Since Y has the form $y + M$, that means we can find $y \in R$ such that $(x + M)(y + M) = 1 + M$; also by coset equality, $xy - 1 \in M$. Written another way, we claim that we can find $y \in R$ and $m \in M$ such that $xy - 1 = m$, or $xy - m = 1$.

To see why the claim is true, observe that $xy \in \langle x\rangle$, so $xy - m \in \langle x\rangle + M$. This is the sum of two ideals, which is also an ideal (Question 4.39). Let's call it $N = \langle x\rangle + M$. Now, $x \in N$ and $x \notin M$ implies that $M \subsetneq N$; by hypothesis, M is maximal, giving us $N = R$. As R is a ring with unity, $1 \in N$. The definition of a sum of ideals tells us that $1 = a + m$ for some $a \in \langle x\rangle$ and some $m \in M$. By definition, we can find $y \in R$ such that $a = xy$. Rewrite the equation $1 = a + m$ as $xy - 1 = -m$, and we have $xy - 1 \in M$, as desired. We finish the proof by reversing the first paragraph: let $Y = y + M$, and $XY \equiv 1 + M$ in R/M. □

Question 6·71 .
The converse is also true: if R is a ring with unity, M is an ideal, and R/M is a field, then M is a maximal ideal. Show why. *Hint:* You should just be able to reverse the main ideas of the explanation above.

Theorem 6·72. *Suppose $f \in \mathbb{F}[x]$ is irreducible.*

(A) $\mathbb{E} = \mathbb{F}[x] / \langle f \rangle$ *is a field.*

(B) $\mathbb{F}$ *is isomorphic to a subfield* $\mathbb{F}'$ *of* $\mathbb{E}$.

(C) Let $\widehat{f} \in \mathbb{E}[y]$ *such that the coefficient of* y^i *is* $a_i + \langle f \rangle$, *where* a_i *is the coefficient of* x^i *in* f. *There exists* $\alpha \in \mathbb{E}$ *such that* $\widehat{f}(\alpha) = 0$.

In other words, $\mathbb{E}$ *contains a root of* $\widehat{f}$.

We call $\mathbb{E}$ an **extension field** of $\mathbb{F}$. The isomorphism between $\mathbb{F}$ and $\mathbb{F}'$ implies that we can *always* assume that an irreducible polynomial over a field $\mathbb{F}$ has a root in another field containing $\mathbb{F}$. We will, in the future, think of $\mathbb{E}$ as a field containing $\mathbb{F}$, rather than containing a field isomorphic to $\mathbb{F}$.

Proof. Denote $I = \langle f \rangle$.

(A) Let $\mathbb{E} = \mathbb{F}[x]/I$. By Corollary 4·53, $\mathbb{F}[x]$ is a principal ideal domain. Theorem 6·28 states that if f is irreducible in $\mathbb{F}[x]$, then I is maximal in $\mathbb{F}[x]$. Fact 6·70 states that $\mathbb{E} = \mathbb{F}[x]/I$ is a field.

(B) To see that $\mathbb{F}$ is isomorphic to

$$\mathbb{F}' = \{a + I : a \in \mathbb{F}\} \subseteq \mathbb{E},$$

use the function $\varphi : \mathbb{F} \to \mathbb{F}'$ by $\varphi(a) = a + I$. You will show in Question 6.73 that φ is a ring isomorphism.

(C) Let $\alpha = y + I$. Let $a_0, a_1, \ldots, a_n \in \mathbb{F}$ such that

$$f = a_0 + a_1 x + \cdots + a_n x^n.$$

As defined in this Theorem,

$$\widehat{f} = (a_0 + I) + (a_1 + I)\, y + \cdots + (a_n + I)\, y^n.$$

By substitution and the arithmetic of ideals,

$$\begin{aligned} \widehat{f}(\alpha) &= (a_0 + I) + (a_1 + I)(x + I) + \cdots + (a_n + I)(x + I)^n \\ &= (a_0 + I) + (a_1 x + I) + \cdots + (a_n x^n + I) \\ &= (a_0 + a_1 x + \cdots + a_n x^n) + I \\ &= f + I. \end{aligned}$$

By Theorem 4·103, $f + I = I$, so $\widehat{f}(\alpha) = I$. Recall that $\mathbb{E} = \mathbb{F}[x]/I$; it follows that $\widehat{f}(\alpha) = 0_{\mathbb{E}}$. □

Question 6·73 .
Show that the function φ defined in part (B) of the proof of Theorem 6·72 is an isomorphism between $\mathbb{F}$ and $\mathbb{F}'$.

The result of this is that, given any irreducible polynomial over a field, we can factor it *symbolically* as follows:

- let $f_0 = f$, $\mathbb{E}_0 = \mathbb{F}$, and $i = 0$;
- repeat while $f_i \neq 1$:
  - let $\mathbb{E}_{i+1} = \mathbb{E}_i[x]/I_i$;
  - let $\alpha_i = x + I_i \in \mathbb{E}_{i+1}$, where $I_i = \langle f_i \rangle$;
  - by Theorem 6·72, $f_i(\alpha_i) = 0$, so by the Factor Theorem, $x - \alpha_i$ is a factor of f_i;
  - let $f_{i+1} \in \mathbb{E}_{i+1}[x]$ such that $f_i = (x - \alpha_i) f_{i+1}$;
  - increment i.

Each pass through the loop generates a new root α_i, and a new polynomial f_i whose degree satisfies the equation

$$\deg f_i = \deg f_{i+1} - 1.$$

Since we have a strictly decreasing sequence of natural numbers, the algorithm terminates after $\deg f$ steps (Question 1·51). We have thus described a way to factor irreducible polynomials.

Corollary 6·74 (Kronecker's Theorem)**.** *Let $f \in \mathbb{F}[x]$ and $n = \deg f$. There exists a field $\mathbb{E}$ such that $\mathbb{F} \subseteq \mathbb{E}$, and f factors into linear polynomials over $\mathbb{E}$.*

Proof. We proceed by induction on $\deg f$.

Inductive base: If $\deg f = 1$, then $f = ax + b$ for some $a, b \in \mathbb{F}$ with $a \neq 0$. In this case, let $\mathbb{E} = \mathbb{F}$; then $-a^{-1}b \in \mathbb{E}$ is a root of f.

Inductive hypothesis: Assume that for any polynomial of degree n, there exists a field $\mathbb{E}$ such that $\mathbb{F} \subseteq \mathbb{E}$, and f factors into linear polynomials in $\mathbb{E}$.

Inductive step: Assume $\deg f = n + 1$. By Question 6.63, $\mathbb{F}[x]$ is a unique factorization domain, so let p be an irreducible factor of f. Let $g \in \mathbb{F}[x]$ such that $f = pg$. By Theorem 6·72, there exists a field $\mathbb{D}$ such that $\mathbb{F} \subsetneq \mathbb{D}$ and $\mathbb{D}$ contains a root α of p. Of course, if α is a root of p, then it is a root of f: $f(\alpha) = p(\alpha) g(\alpha) = 0 \cdot g(\alpha) = 0$. By the Factor Theorem, we can write $f = (x - \alpha) q(x) \in \mathbb{D}[x]$. We now have $\deg q = \deg f - 1 = n$. By the inductive hypothesis, there exists a field $\mathbb{E}$ such that $\mathbb{D} \subseteq \mathbb{E}$, and q factors into linear polynomials in $\mathbb{E}$. But then $\mathbb{F} \subsetneq \mathbb{D} \subseteq \mathbb{E}$, and f factors into linear polynomials over $\mathbb{F}$. □

Remember that we can write elements of $\mathbb{R}[i]$ and $\mathbb{Q}\left[\sqrt{2}\right]$ in the form $ax + b$, where a and b come from the underlying field. We have two free variables a and b to choose *any ring element we want* as coefficients. There are infinitely many such polynomials, but with respect to the field elements they behave a little like vectors. Is that always the case?

Theorem 6·75. *Let f be an irreducible polynomial over the field $\mathbb{F}$, and $\mathbb{E} = \mathbb{F}[x]/\langle f\rangle$. Then $\mathbb{E}$ is a vector space over $\mathbb{F}$ of dimension $d = \deg f$.*

Proof. Let $I = \langle f\rangle$. By Theorem 6·72, we can consider $\mathbb{F} \subseteq \mathbb{E}$. Since f is irreducible, $\langle f\rangle$ is maximal, and $\mathbb{E}$ is a field. Any element of $\mathbb{E}$ has the form $g + I$ where $g \in \mathbb{F}[x]$; we can use the fact that $\mathbb{F}[x]$ is a Euclidean Domain to write

$$g = qf + r$$

where $q, r \in \mathbb{F}[x]$ and $\deg r < \deg f = d$. Notice $g - r \in \langle f\rangle = I$, so coset equality assures us that $g + I = r + I$. In other words, every element of $\mathbb{E}$ has the form

$$\left(a_{d-1}x^{d-1} + \cdots + a_1x^1 + a_0x^0\right) + I$$

where $a_{d-1}, \ldots, a_1, a_0 \in \mathbb{F}$.

Finally, view each coset written in this form as the vector $(a_{d-1}, \ldots, a_1, a_0)$ over the field $\mathbb{F}$; define vector addition as the component-wise addition of the coset representatives and scalar multiplication by $b \in \mathbb{F}$ as$(ba_{d-1}, \ldots, ba_1, ba_0)$. The vector and scalar properties of a vector space are fairly straightforward; we leave them to you as an exercise. Once done, we have proved that $\mathbb{E}$ is a vector space over $\mathbb{F}$ with basis

$$B = \left\{x^0 + I, x^1 + I, \ldots, x^{d-1} + I\right\}.$$

□

Question 6·76 .
Show the remaining details that $\mathbb{E}$ is indeed a vector space over $\mathbb{F}$.

Definition 6·77. Let f, α, and $\mathbb{E}$ be as in Theorem 6·72. Theorem 6·75 tells us that $\deg f = \dim \mathbb{E}$. We call this number the **degree of** α.

It is sensible to say that $\deg f = \deg \alpha$ since we showed in Theorem 6·75 that $\deg f = \dim(\mathbb{F}[x]/\langle f\rangle)$.

Example 6·78. Let $f(x) = x^4 + 1 \in \mathbb{Q}[x]$. We can construct a field $\mathbb{D}$ with a root α of f; using the proofs above,

$$\mathbb{D} = \mathbb{Q}[x]/\langle f\rangle \qquad \text{and} \qquad \alpha = x + \langle f\rangle.$$

Notice that $-\alpha$ is also a root of f, so in fact, $\mathbb{D}$ contains two roots of f. If we repeat the procedure, we obtain two more roots of f in a field $\mathbb{E}$.

What if we extend a field more than once?

Theorem 6·79. *Suppose $\mathbb{F}$ is a field, $\mathbb{E} = \mathbb{F}(\alpha)$, and $\mathbb{D} = \mathbb{E}(\beta)$. Then $\mathbb{E}$ is a vector space over $\mathbb{F}$ of dimension $\deg\alpha \cdot \deg\beta$, and in fact $\mathbb{D} = \mathbb{F}(\gamma)$ for some root γ of a polynomial over $\mathbb{F}$.*

Proof. By Theorem 6·75, $B_1 = \left\{\alpha^0, \ldots, \alpha^{d_1-1}\right\}$ and $B_2 = \left\{\beta^0, \ldots, \beta^{d_2-1}\right\}$ are bases of $\mathbb{E}$ over $\mathbb{F}$ and $\mathbb{D}$ over $\mathbb{E}$, respectively, where d_1 and d_2 are the respective degrees of the irreducible polynomials of which α and β are roots. We claim that $B_3 = \left\{\alpha^{(i)}\beta^{(j)} : 0 \leqslant i < d_1, 0 \leqslant j < d_2\right\}$ is a basis of $\mathbb{D}$ over $\mathbb{F}$. To see this, we must show that it is both a spanning set — that is, every element of $\mathbb{D}$ can be written as a linear combination of elements of B_3 over $\mathbb{F}$ — and that its elements are linearly independent.

To show that B_3 is a spanning set, let $\gamma \in \mathbb{D}$. By definition of basis, there exist b_0, …, $b_{d_2-1} \in \mathbb{E}$ such that

$$\gamma = b_0\beta^0 + \cdots + b_{d_2-1}\beta^{d_2-1}.$$

Likewise, for each $j = 0, \ldots, d_2 - 1$ there exist $a_0^{(j)}, \ldots, a_{d_1-1}^{(j)} \in \mathbb{F}$ such that

$$b_j = a_0^{(j)}\alpha^0 + \cdots + a_{d_1-1}^{(j)}\alpha^{d_1-1}.$$

By substitution,

$$\begin{aligned}\gamma &= \sum_{j=0}^{d_2-1} b_j\beta^j \\ &= \sum_{j=0}^{d_2-1}\left(\sum_{i=0}^{d_1-1} a_i^{(j)}\alpha^i\right)\beta^j \\ &= \sum_{i=0}^{d-1}\sum_{j=0}^{d-1} a_i^{(j)}\left(\alpha^i\beta^j\right).\end{aligned}$$

Hence, B_3 is a spanning set of $\mathbb{D}$ over $\mathbb{F}$.

To show that it is a basis, we must show that its elements are linearly independent. For that, assume we can find $c_i^{(j)} \in \mathbb{F}$ such that

$$\sum_{i=0}^{d_1-1}\sum_{j=0}^{d_2-1} c_i^{(j)}\left(\alpha^i\beta^j\right) = 0.$$

We can rewrite this as an element of $\mathbb{D}$ over $\mathbb{F}$ by rearranging the sum:

$$\sum_{j=0}^{d_2-1}\left(\sum_{i=0}^{d_1-1} c_i^{(j)}\alpha^i\right)\beta^j = 0.$$

Since B_2 is a basis, its elements are linearly independent, so the coefficient of each β^j must be zero. In other words, for each j, we have

$$\sum_{i=0}^{d_1-1} c_i^{(j)}\alpha^i = 0.$$

Of course, B_1 is also a basis, so its elements are also linearly independent, so the coefficient of each α^i must be zero. In other words, for each j and each i,

$$c_i^{(j)} = 0.$$

We took an arbitrary linear combination of elements of B_3 over $\mathbb{F}$, and showed that it is zero only if each of the coefficients are zero. Thus, the elements of B_3 are linearly independent.

Since the elements of B_3 are a linearly independent spanning set, B_3 is a basis of $\mathbb{D}$ over $\mathbb{F}$. If we count the number of elements of B_3, we find that there are $d_1 \cdot d_2$ elements of the basis. Hence,

$$\dim_{\mathbb{F}} \mathbb{D} = |B_3| = d_1 \cdot d_2 = \deg \alpha \cdot \deg \beta.$$

We still have to show that $\mathbb{D} = \mathbb{F}(\gamma)$ for some root $\gamma \in \mathbb{D}$ of a polynomial $f \in \mathbb{F}[x]$. If $\alpha \in \mathbb{F}(\beta)$, then $\mathbb{D} = \mathbb{F}(\beta)$, and we are done. Likewise if $\beta \in \mathbb{F}(\alpha)$, then $\mathbb{D} = \mathbb{F}(\alpha)$. Otherwise, we claim that $\gamma = \alpha + \beta$ does the job. Why? By closure of addition, $\gamma = \alpha + \beta \in \mathbb{D}$, so $\mathbb{D} \supseteq \mathbb{F}(\gamma)$ rather easily. For the reverse inclusion, we point out that if the the sequence

$$\gamma, \quad \gamma^2, \quad \gamma^3, \quad \dots$$

consists entirely of elements that are linearly independent over $\mathbb{F}$, then $\mathbb{D}$ cannot be finite dimensional over $\mathbb{F}$, as closure of multiplication means $\mathbb{D}$ must contain all these powers of γ. We have just seen that $\mathbb{D}$ *is* finite dimensional, so there must be elements $c_0, c_1, \dots, c_{d_1 d_2 - 1} \in \mathbb{F}$ satisfying

$$c_0 + c_1\gamma + \cdots + c_{d_1 d_2 - 1}\gamma^{d_1 d_2 - 1} = 0.$$

Let $f(x) = c_0 + c_1 x + \cdots + c_{d_1 d_2 - 1} x^{d_1 d_2 - 1}$; we have $f \in \mathbb{F}[x]$ and $f(\gamma) = 0$. We claim that $c_{d_1 d_2 - 1} \neq 0$; if it were, then we could substitute $x = \alpha + y$ and obtain a polynomial $g(y) = f(\alpha + y) \in [\mathbb{F}(\alpha)][y] = \mathbb{E}[y]$ with β as a root, but of smaller degree than d_2. This is a contradiction; hence, the smallest degree possible for f is $d_1 d_2 - 1$, which means (using what we just proved) that

$$\dim \mathbb{F}(\alpha + \beta) = d_1 d_2 = (\dim_{\mathbb{F}} \mathbb{E})(\dim_{\mathbb{E}} \mathbb{D}) = \dim_{\mathbb{F}} \mathbb{D}.$$

We have two vector spaces of identical dimension over $\mathbb{F}$, one contained in the other; this is possible if and only if the two are identical. Hence $\mathbb{D} = \mathbb{F}(\alpha + \beta)$. □

Question 6·80 .

Let $\mathbb{F} = \mathbb{Q}\left(\sqrt{2}\right)\left(\sqrt{3}\right)$.

(a) Find an polynomial $f \in \mathbb{Q}[x]$ that is irreducible over $\mathbb{Q}$ but factors over $\mathbb{F}$.

(b) What is $\dim_{\mathbb{Q}} \mathbb{F}$?

Question 6·81 .

Factor $x^3 + 2$ over $\mathbb{Q}$ using the techniques described in this section. You may use the fact that if $a = b^n$, then $x^n + a = (x + b)(x^{n-1} - bx^{n-2} + \cdots + b^{n-1})$.

6·5 Finite Fields I

We saw in Section 3·4 that the characteristic of a finite ring or field tells us a great deal; for instance, $\mathbb{Z}_n$ is a field when n is irreducible. The finite fields that we have worked with so far are of the form $\mathbb{Z}_p$, where p is irreducible.

Don't jump to the conclusion that the size of a finite field is the same as the number of elements! After all, in Example 3·60 on page 90 we encountered a finite field, generated by polynomials, that had characteristic 3 but 9 elements.

You might notice that 9 is a power of 3. This is no mere coincidence; the goal of this section is to establish that every finite field has p^n elements where $p, n \in \mathbb{N}$ and p is irreducible.

Quick review

In a ring R without zero divisors, $cr \neq 0$ for every $c \in \mathbb{N}^+$ and every $r = 0$. Not all rings satisfy this property; the **characteristic** of a ring is therefore 0 when the first property holds, otherwise the smallest integer c satisfying $c \cdot 1 = 0$, and c was the smallest positive integer satisfying this property.

Example 6·82. The rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ have characteristic zero.

The ring $\mathbb{Z}_8$ has characteristic 8. Why? Certainly $8 \cdot [1] = [8] = [0]$, and no smaller integer n gives us $n \cdot [1] = [0]$. In fact, the characteristic of $\mathbb{Z}_n$ is n for *any* $n \in \mathbb{N}^+$.

Let $p \in \mathbb{Z}$ be irreducible. We know from Fact 3·65 that $\mathbb{Z}_p$ is a field. The same argument we used in Example 6·82 shows that the characteristic of $\mathbb{Z}_p$ is p.

In the previous example, the characteristic of a finite ring turned out to be the number of elements in the ring. This is not always the case.

Example 6·83. Let $R = \mathbb{Z}_2 \times \mathbb{Z}_4 = \{(a, b) : a \in \mathbb{Z}_2, b \in \mathbb{Z}_4\}$, with addition and multiplication defined in the natural way:

$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b) \cdot (c, d) = (ac, bd) .$$

From Fact 2·64 on page 67, R is a ring. It has eight elements,

$$R = \{([0]_2 , [0]_4) , ([0]_2 , [1]_4) , ([0]_2 , [2]_4) , ([0]_2 , [3]_4) ,$$
$$([1]_2 , [0]_4) , ([1]_2 , [1]_4) , ([1]_2 , [2]_4) , ([1]_2 , [3]_4)\} .$$

However, the characteristic of R is not eight, but four:

- for any $a \in \mathbb{Z}_2$, we know that $2a = [0]_2$, so $4a = 2\,[0]_2 = [0]_2$; and
- for any $b \in \mathbb{Z}_4$, we know that $4b = [0]_4$; thus
- for any $(a, b) \in R$, we see that $4\,(a, b) = (4a, 4b) = ([0]_2 , [0]_4) = 0_R$.

Since the characteristic of $\mathbb{Z}_4$ is 4, we cannot go smaller than that.

Building finite fields

The standard method of building a finite field is different from what we will do here, but the method used here is an interesting application of quotient rings.

Notation 6·84. Our notation for a finite field with n elements is $\mathbb{F}_n$.

Example 6·85. You have already seen a finite field with nine elements (Example 3·60); here we build a finite field with sixteen elements.

To build $\mathbb{F}_{16}$, start with the polynomial ring $\mathbb{Z}_2[x]$. We claim that $f(x) = x^4 + x + 1$ does not factor in $\mathbb{Z}_2[x]$; if it did, it would have to factor as a product of either a linear and cubic polynomial, or as a product of two quadratic polynomials. The former is impossible, since neither 0 nor 1 is a root of f. As for the second, suppose that $f = (x^2 + ax + b)(x^2 + cx + d)$, where $a, b, c, d \in \mathbb{Z}_2$. Expanding the product, we have

$$\begin{aligned} x^4 + x + 1 = x^4 &+ (a + c)\,x^3 + (ac + b + d)\,x^2 \\ &+ (ad + bc)\,x + db. \end{aligned}$$

Equal polynomials have the same coefficients for like terms, giving us a system of linear equations,

$$\begin{aligned} a + c &= 0 \\ ac + b + d &= 0 \\ ad + bc &= 1 \\ bd &= 1. \end{aligned} \tag{6.1}$$

Recall that $b, d \in \mathbb{Z}_2$, so (6.1) means that $b = d = 1$; after all, the only other choice would be 0, which would contradict $bd = 1$. The system now simplifies to

$$a + c = 0 \tag{6.2}$$

$$ac + \cancel{1} + \cancel{1} = ac = 0$$

$$a + c = 1 \tag{6.3}$$

Equations 6.2 and 6.3 contradict! That shows f is irreducible, and Fact 3·69 tells us that we can build a field by taking $\mathbb{Z}_2[x]$ modulo f.

How many elements does this field have? Let $X \in R/I$; choose a representation $g + I$ of X where $g \in R$. Without loss of generality, we can assume that $\deg g < 4$, since if $\deg g \geqslant 4$ then we can divide and use the remainder, instead. There are thus four terms in g: c_3x^3, c_2x^2, c_1x^1, and c_0x^0. Each term's coefficient is either $[0]$ or $[1]$. This gives us $2^4 = 16$ distinct possibilities for X, and so 16 elements of R/I,

$$\begin{aligned} &I, & &1 + I, \\ &x + I, & &x + 1 + I, \\ &x^2 + I, & &x^2 + 1 + I, \\ &x^2 + x + I, & &x^2 + x + 1 + I, \\ &x^3 + I, & &x^3 + 1 + I, \\ &x^3 + x + I, & &x^3 + x + 1 + I, \\ &x^3 + x^2 + I, & &x^3 + x^2 + 1 + I, \\ &x^3 + x^2 + x + I, & &x^3 + x^2 + x + 1 + I. \end{aligned}$$

Question 6·86 .
Construct a field with 27 elements, and list them all.

Recalling the link between irreducible elements and ideals, we point out that

- $\mathbb{Z}_2$ is a field, so
- $\mathbb{Z}_2[x]$ is a principal ideal domain (Theorem 4·53(C)), so
- $\mathbb{Z}_2[x]$ is a unique factorization domain (Theorem 6·42), so
- $I = \langle f \rangle$ is a maximal ideal in $R = \mathbb{Z}_2[x]$ (Theorem 6·28(A)), and it just so happened that
- R/I turned out to be a field.

This illustrates Fact 6·70.

You may have noticed that we obtained $\mathbb{F}_9$ by starting in $\mathbb{Z}_3[x]$ and using an irreducible element of degree 2; we obtained $\mathbb{F}_{16}$ by starting in $\mathbb{Z}_2[x]$ and using an irreducible element of degree 4; you (hopefully) obtained $\mathbb{F}_{27}$ by starting in $\mathbb{Z}_3[x]$ and using a polynomial of degree 3. In turn, each gave us 3^2, 2^4, and 3^3 elements; that is, p^n elements where p is the characteristic and n is the degree.

You might wonder if this also generalizes to arbitrary finite fields: that is,

- start with $\mathbb{Z}_p[x]$,
- find a polynomial of degree n that does not factor in that ring, then
- build a quotient ring such that
- the field has p^n elements.

Yes and no. We do start with $\mathbb{Z}_p$, and all finite fields have p^n elements.

Theorem 6·87. *Suppose that $\mathbb{F}_n$ is a finite field with n elements. Then n is a power of an irreducible integer p, and the characteristic of $\mathbb{F}_n$ is p.*

Proof of Theorem 6·87: Let p be the characteristic of $\mathbb{F}_n$; by Theorem 6·75 on page 235, $\mathbb{F}_n$ is a vector space over $\mathbb{Z}_p$. The space has finitely many elements, so it has finite dimension over $\mathbb{Z}_p$. Let $m = \dim \mathbb{F}_n$. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be a basis of $\mathbb{F}_n$ over $\mathbb{Z}_p$; every linearly independent element of $\mathbb{F}_n$ has the form $a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m$, where $a_i \in \mathbb{Z}_p$ is arbitrary. As we have p choices for each a_i, there are p^m possible vectors, so $n = |\mathbb{F}_n| = p^m$. □

To construct $\mathbb{F}_{p^n}$ for *every* irreducible p and *every* $n \in \mathbb{N}^+$, however, we would need to find a polynomial of degree n that is irreducible over $\mathbb{F}_p$. It is not obvious that such polynomials exist for every possible p and n. That is the subject of Section 6·6.

Question 6·88 .
Does every infinite field have characteristic 0? To see why not, consider the set of all "rational functions" over $\mathbb{Z}_2$,

$$R_2(x) = \left\{ \frac{f(x)}{g(x)} \ : \ f, g \in \mathbb{Z}_2[x] \text{ and } g \neq 0 \right\}.$$

For instance,

$$0, \quad x^2, \quad , \frac{1}{x+1}, \quad \frac{x+1}{x^2+1} \qquad \in \qquad R_2(x).$$

As you might expect, we consider two rational functions f/g and p/q **equivalent** if $fq = pg$ as polynomials, so in fact

$$\frac{1}{x+1} = \frac{x+1}{x^2+1} \qquad \text{because} \qquad x^2+1 = (x+1)(x+1).$$

(Don't forget that in $\mathbb{Z}_2[x]$ we have $2x = 0$.)

(a) Show that the relation described above is in fact an equivalence relation.

(b) Show that the set of equivalence classes of this relation forms a field. We call this field $\mathbb{Z}_2(x)$.

(c) Explain why the characteristic of this field is 2.

(d) Explain why this means we can create an infinite field of characteristic p for *any* irreducible integer p.

6·6 Finite fields II

We saw in Section 6·5 that *if* a field is finite, then its size is p^n for some $n \in \mathbb{N}^+$ and some irreducible integer p. In this section, we show the converse: for every irreducible integer p and for every $n \in \mathbb{N}^+$, there exists a field with p^n elements. In this section, we show that for any polynomial $f \in \mathbb{F}[x]$, where $\mathbb{F}$ is a field of characteristic p,

- there exists a field $\mathbb{E}$ containing *one* root of f;
- there exists a field $\mathbb{E}$ where f factors into linear polynomials; and
- we can use this fact to build a finite field with p^n elements for any irreducible integer p, and for any $n \in \mathbb{N}^+$.

Before we proceed to the main topic of this section, we need a concept that we borrow from Calculus.

Definition 6·89. Let $f \in \mathbb{F}[x]$, and write $f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. The **formal derivative of** f is

$$f' = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

Proposition 6·90 (The product rule)**.** *Let $f \in \mathbb{F}[x]$, and suppose f factors as $f = pq$. Then $f' = p'q + pq'$.*

Proof. Write $p = \sum_{i=0}^{m} a_i x^i$ and $q = \sum_{j=0}^{n} b_j x^j$. First we write f in terms of the coefficients of p and q. By the distributive property,

$$f = pq = \sum_{i=0}^{m} \left[a_i x^i \sum_{j=0}^{n} b_j x^j \right] = \sum_{i=0}^{m} \left[\sum_{j=0}^{n} \left(a_i b_j \right) x^{i+j} \right].$$

If we collect like terms, we can rewrite this as

$$f = \sum_{k=0}^{m+n} \left[\left(\sum_{i+j=k} a_i b_j \right) x^k \right].$$

We can now examine the claim. By definition,

$$f' = \sum_{k=1}^{m+n} \left[k \left(\sum_{i+j=k} a_i b_j \right) x^{k-1} \right].$$

On the other hand,

$$\begin{aligned}
p'q + pq' &= \left(\sum_{i=1}^{m} i a_i x^{i-1} \right) \left(\sum_{j=0}^{n} b_j x^j \right) \\
&\quad + \left(\sum_{i=0}^{m} a_i x^i \right) \left(\sum_{j=1}^{n} j b_j x^{j-1} \right) \\
&= \sum_{k=1}^{m+n} \left[\left(\sum_{i+j=k} i a_i b_j \right) x^{k-1} \right] \\
&\quad + \sum_{k=1}^{m+n} \left[\left(\sum_{i+j=k} j a_i b_j \right) x^{k-1} \right] \\
&= \sum_{k=1}^{m+n} \left[\left(\sum_{i+j=k} (i+j)\, a_i b_j \right) x^{k-1} \right] \\
&= \sum_{k=1}^{m+n} \left[\left(\sum_{i+j=k} k a_i b_j \right) x^{k-1} \right] \\
&= f'.
\end{aligned}$$

□

We can now prove the main idea of this section.

The existence of finite fields

Theorem 6·91. *For any irreducible integer p, and for any $n \in \mathbb{N}^+$, there exists a field with p^n elements.*

Proof. First, suppose $p = 2$. If $n = 1$, the field $\mathbb{Z}_2$ proves the theorem. If $n = 2$, the field $\mathbb{Z}_2/\langle x^2 + x + 1\rangle$ proves the theorem. We may therefore assume that $p \neq 2$ or $n \neq 1, 2$.

Let $f = x^{p^n} - x \in \mathbb{Z}_p[x]$. By Kronecker's Theorem, there exists a field $\mathbb{D}$ such that $\mathbb{Z}_p \subseteq \mathbb{D}$, and f factors into linear polynomials over $\mathbb{D}$. Let $\mathbb{E} = \{\alpha \in \mathbb{D} : f(\alpha) = 0\}$. We claim that $\mathbb{E}$ has p^n elements, and that $\mathbb{E}$ is a field.

To see that $\mathbb{E}$ has p^n elements, it suffices to show that f has no repeated linear factors. Recall that $f = x^{p^n} - x$. The definition of a formal derivative tells us that

$$f' = p^n x^{p^n - 1} - 1.$$

In $\mathbb{Z}_p$, $p^n = 0$, so we can simplify f' as

$$f' = 0 - 1 = -1.$$

When we assumed that f had a repeated linear factor, we concluded that $x - a$ divides f'. However, we see now that $f' = -1$, and $x - a$ certainly does *not* divide -1, since $\deg(x - a) = 1 > 0 = \deg(-1)$. That assumption leads to a contradiction; so, f has no repeated linear factors.

We now show that $\mathbb{E}$ is a field. By its very definition, $\mathbb{E}$ consists of elements of $\mathbb{D}$; thus, $\mathbb{E} \subseteq \mathbb{D}$. We know that $\mathbb{D}$ is a field, and thus a ring; we can therefore use the Subring Theorem to show that $\mathbb{E}$ is a ring. Once we have that, we have to find an inverse for any nonzero element of $\mathbb{E}$.

For the Subring Theorem, let $a, b \in \mathbb{E}$. We must show that ab and $a - b$ are both roots of f; they would then be elements of $\mathbb{E}$ by definition of the latter. You will show in Question 6.93(a) that ab is a root of f. For subtraction, we claim that

$$(a - b)^{p^n} = a^{p^n} - b^{p^n}.$$

We proceed by induction.

Inductive base: Assume $n = 1$. Observe that

$$(a - b)^p = a^p + \sum_{i=1}^{p-1} (-1)^i \binom{p}{i} a^i b^{p-i} + (-1)^p b^p.$$

By assumption, p is an irreducible integer, so its only divisors in $\mathbb{N}$ are itself and 1. For any $i \in \mathbb{N}^+$, then, the integer

$$\binom{p}{i} = \frac{p!}{i!\,(p-i)!}$$

can be factored into the two integers

$$\binom{p}{i} = p \cdot \frac{(p-1)!}{i!\,(p-i)!};$$

the fraction $\frac{(p-1)!}{i!(p-i)!}$ is an integer precisely because no element of the denominator can divide p. Using Question 6.93(b), we can rewrite $(a-b)^p$ as

$$\begin{aligned}(a-b)^p &= a^p + \sum_{i=1}^{p-1} (-1)^i \frac{p!}{i!\,(p-i)!} a^i b^{p-i} + (-1)^p b^p \\ &= a^p + p \cdot \sum_{i=1}^{p-1} (-1)^i \frac{(p-1)!}{i!\,(p-i)!} a^i b^{p-i} + (-1)^p b^p \\ &= a^p + 0 + (-1)^p b^p \\ &= a^p + (-1)^p b^p.\end{aligned}$$

If $p = 2$, then $-1 = 1$, so either way we have $a^p - b^p$, as desired.

Inductive hypothesis: Assume that $(a-b)^{p^n} = a^{p^n} - b^{p^n}$.

Inductive step: Applying the properties of exponents,

$$\begin{aligned}(a-b)^{p^{n+1}} &= \left[(a-b)^{p^n}\right]^p \\ &= \left(a^{p^n} - b^{p^n}\right)^p = a^{p^{n+1}} - b^{p^{n+1}},\end{aligned}$$

where the final step uses the base case. Thus

$$(a-b)^{p^n} - (a-b) = \left(a^{p^n} - b^{p^n}\right) - (a-b).$$

Again, a and b are roots of f, so $a^{p^n} = a$ and $b^{p^n} = b$, so

$$(a-b)^{p^n} - (a-b) = (a-b) - (a-b) = 0.$$

We see that $a - b$ is a root of f, and therefore $a - b \in \mathbb{E}$.

Finally, we show that every nonzero element of $\mathbb{E}$ has an inverse in $\mathbb{E}$. Let $a \in \mathbb{E} \setminus \{0\}$; by definition, $a \in \mathbb{D}$. Since $\mathbb{D}$ is a field, there exists an inverse of a in $\mathbb{D}$; call it b. By definition of $\mathbb{E}$, a is a root of f; that is, $a^{p^n} - a = 0$. Multiply both sides of this equation by b^2, and rewrite to obtain $a^{p^n-2} = b$. Using the substitutions $b = a^{p^n-2}$ and $a^{p^n} = a$ in $f(b)$ shows that:

$$\begin{aligned}f(b) &= b^{p^n} - b \\ &= \left(a^{p^n-2}\right)^{p^n} - a^{p^n-2} \\ &= \left(a^{p^n} \cdot a^{-2}\right)^{p^n} - a^{p^n-2} \\ &= \left(a^{p^n}\right)^{p^n} \left(a^{p^n}\right)^{-2} - a^{p^n-2} \\ &= a^{p^n} \cdot a^{-2} - a^{p^n-2} \\ &= a^{p^n-2} - a^{p^n-2} \\ &= 0.\end{aligned}$$

We have shown that b is a root of f. By definition, $b \in \mathbb{E}$. Since $b = a^{-1}$ and a was an arbitrary element of $\mathbb{E} \setminus \{0\}$, *every* nonzero element of $\mathbb{E}$ has its inverse in $\mathbb{E}$.

We have shown that

- $\mathbb{E}$ has p^n elements;
- it is a ring, since it is closed under multiplication and subtraction; and
- it is a field, since every nonzero element has a multiplicative inverse in $\mathbb{E}$.

In other words, $\mathbb{E}$ is a field with p^n elements. □

Euler's theorems

The existence of finite fields means affords us some nice theorems that generalize Euler's Theorem.

Euler's Theorem for arbitrary finite fields. *If p is irreducible and $f(x) = x^{p^n} - x$, then $f(a) = 0$ for all $a \in \mathbb{Z}_p$.*

The proof is an exercise:

Question 6·92 .

Let $\mathbb{F}_n$ be a finite field of size n. (We now know such critters exist.)

(a) Use a corollary to Lagrange's Theorem to explain why $a^{n-1} = 1$ for every *nonzero* $a \in \mathbb{F}_n$.

(b) Explain how we know $n = p^k$ for some $k \in \mathbb{N}^+$.

(c) Combine (a) and (b) to show Euler's Theorem for arbitrary finite fields.

We can view this result a different way, too.

Euler's Theorem for polynomials. *Let p be an irreducible integer. For all $a \in \mathbb{F}_p$ and for all $n \in \mathbb{N}^+$, $a^{p^n} - a = 0$, and thus $a^{p^n} = a$ and in $\mathbb{Z}_p[x]$, we have*

$$x^p - x = \prod_{a \in \mathbb{Z}_p} (x - a).$$

Proof. Recall the $\mathbb{F}[x]$ is a unique factorization domain for any field $\mathbb{F}$; $\mathbb{Z}_p$ is a field, so $x^p - x$ has a unique factorization. By Euler's Theorem for arbitrary finite fields, a is a root of $x^p - x$ for every $a \in \mathbb{Z}_p \cong \mathbb{F}_p$. Apply the Factor Theorem to complete the proof. □

We can generalize Euler's Theorem a little further.

Fermat's Little Theorem on polynomials. *Let $k \in \mathbb{N}^+$. In $\mathbb{F}_{p^k}[x]$, we have*

$$x^{p^k} - x = \prod_{a \in \mathbb{F}_{p^k}} (x - a).$$

Proof. We first claim that every $a \in \mathbb{F}_{p^k}$ is a root of $x^{p^k} - x$. This is obvious when $a = 0$, so assume a lies in $\mathbb{F}_{p^k} \setminus \{0\}$. By definition, the nonzero elements of $\mathbb{F}_{p^k}$ form a group under multiplication. By Corollary 4·113 to Lagrange's Theorem, the order of a divides $\left|\mathbb{F}_{p^k} \setminus \{0\}\right| = p^k - 1$, so $a^{p^k - 1} = 1$. Multiplying both sides by a, we have $a^{p^k} = a$, which we can rewrite as $a^{p^k} - a = 0$. By definition, a is a root of $x^{p^d} - x$. The Factor Theorem implies that $x - a$ is a factor of $x^{p^k} - x$. If $a, b \in \mathbb{F}_{p^k}$ are distinct, so are $x - a, x - b \in \mathbb{F}_{p^k}$, so there are at least p^k such factors. The fact that $\mathbb{F}_{p^k}[x]$ is a unique factorization domain completes the proof. □

Question 6·93 .
Let p be an irreducible integer and $f(x) = x^{p^n} - x \in \mathbb{Z}_p[x]$. Define $\mathbb{E} = \mathbb{Z}_p[x] / \langle f \rangle$.

(a) Show that $pa = 0$ for all $a \in \mathbb{E}$.

(b) Show that if $f(a) = f(b) = 0$, then $f(ab) = 0$.

6·7 Polynomial factorization in finite fields

We now turn to the question of factoring polynomials in $R[x]$. This material comes primarily from [3]. Keep in mind that the goal of these notes is merely to show you how the ideas studied so far combine into this problem, so the algorithms we study won't be cutting-edge practice, though they're not bad, either.

This section factors polynomials whose coefficients come from finite fields, as that is somewhat easier than factoring polynomials whose coefficients come from the integers. We put that off to the next section.

Factorization of $f \in R[x]$ requires the following steps.

- **Squarefree factorization** is the process of removing multiples of factors of f; that is, if $g^a \mid f$, then we want to work with $\frac{f}{g^{a-1}}$, of which only g is a factor.
- **Distinct degree factorization** is the process of factoring a squarefree polynomial f into polynomials $g_1, \ldots, g_m$ such that if g_i factors as $g_i = h_1 \cdots h_n$, then $\deg h_1 = \cdots \deg h_n$.
- **Equal degree factorization** is the process of factoring each distinct degree factor g_i into its equal degree factors $h_1, \ldots, h_n$.

Example 6·94. Suppose $R = \mathbb{Z}_2$. Let

$$f(x) = x^{16} + x^{13} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^5 + x^2.$$

You can see that $g(x) = x^2$ is a factor of f, so f is not squarefree. (It is not typically this easy.) Squarefree factorization identifies this factor and removes it, reducing the problem to factoring

$$g(x) = x^2 \quad \text{and} \quad h(x) = x^{14} + x^{11} + x^9 + x^8 + x^7 + x^6 + x^5 + x^3 + 1.$$

Distinct degree factorization factors h as

$$\left(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1\right)\left(x^8 + x^7 + x^5 + x^4 + x^3 + x + 1\right). \tag{6.4}$$

Equal degree factorization focuses on the second two factors, giving us

$$\left[\left(x^3 + x + 1\right)\left(x^3 + x^2 + 1\right)\right]\left[\left(x^4 + x + 1\right)\left(x^4 + x^2 + 1\right)\right]. \tag{6.5}$$

Notice how the second and third factors in (6.5), which come from the second factor of (6.4), have the same degree. Likewise, the second and third factors of (6.5), which have the same degree, come from the third factor of (6.4).

For the rest of this section, we assume that $p \in \mathbb{N}$ is irreducible and $f \in \mathbb{Z}_p[x]$.

It would be nice to proceed in order, but the approach we take requires us to perform distinct- and equal-degree factorization first.

Distinct degree factorization.

We accomplish distinct-degree factorization via Fermat's Little Theorem on polyomials.

Example 6·95. Suppose $p = 5$. You already know from basic algebra that

$$\begin{aligned} x^5 - x &= x\left(x^4 - 1\right) \\ &= x\left(x^2 - 1\right)\left(x^2 + 1\right) \\ &= x(x-1)(x+1)\left(x^2+1\right). \end{aligned}$$

We are working in $\mathbb{Z}_5$, so $1 = -4$. Thus $x + 1 = x - 4$, and $(x-2)(x-3) = (x^2 - 5x + 6) = (x^2 + 1)$. This means that we can write

$$x^5 - x = x(x-1)(x-2)(x-3)(x-4) = \prod_{a \in \mathbb{Z}_5} (x - a),$$

as claimed.

Generalization of Fermat's Little Theorem for polynomials. *Let $k, q \in \mathbb{N}^+$, and $a = qk$. Then $x^{p^a} - x$ is the product of all monic irreducible polynomials in $\mathbb{F}_{p^k}[x]$ whose degree divides q.*

Proof. Let $f \in \mathbb{F}_{p^k}[x]$ be monic and irreducible of degree n. We will show that

$$f \mid \left(x^{p^a} - x\right) \quad \iff \quad n \mid a.$$

Assume first that f divides $x^{p^a} - x$. By unique factorization and Fermat's Little Theorem on polynomials, f factors into linear polynomials $x - c$, where $c \in \mathbb{F}_{p^a}$. Let α be any one of the corresponding roots, and let $\mathbb{E} = \mathbb{F}_{p^k}(\alpha)$. Using the basis B of Theorem 6·75, we see that $|\mathbb{E}| = \left(p^k\right)^n = p^{kn}$, since it has $|B| = n$ basis elements, and p^k choices for each coefficient of a basis element.

Now, $\mathbb{F}_{p^a}$ is the extension of $\mathbb{E}$ by the remaining roots of $x^{p^a} - x$, one after the other. By reasoning similar to that for $\mathbb{E}$, we see that $p^a = \left|\mathbb{F}_{p^a}\right| = |\mathbb{E}|^b = p^{bkn}$ for some $b \in \mathbb{N}^+$. Rewriting the extreme sides of that equation, we have $p^{bkn} = p^a = p^{qk}$, whence $bkn = qk$ and $n \mid q$.

Conversely, assume $n \mid q$. A straightforward computation verifies that

$$p^q - 1 = \left(p^n - 1\right)\left(p^{q-n} + p^{q-2n} + \cdots + p^n + 1\right).$$

Let $r = p^{q-n} + p^{q-2n} + \cdots + p^n + 1$; a straightforward computation verifies that

$$x^{p^q - 1} - 1 = \left(x^{p^n - 1} - 1\right)\left(x^{(p^n-1)(r-1)} + x^{(p^n-1)(r-2)} + \cdots + x^{p^n - 1} + 1\right).$$

Algorithm 6.1 Distinct degree factorization

inputs
$f \in \mathbb{Z}_p[x]$, squarefree and monic, of degree $n > 0$
outputs
$p_1, \ldots, p_m \in \mathbb{Z}_p[x]$, a distinct-degree factorization of f
do
Let $h_0 = x$
Let $f_0 = f$
Let $i = 0$
while $f_i \neq 1$ **do**
Increment i
Let h_i be the remainder of division of h_{i-1}^p by f
Let $p_i = \gcd(h_i - x, f_{i-1})$
Let $f_i = \frac{f_{i-1}}{p_i}$
Let $m = i$
return $p_1, \ldots, p_m$

Rewrite this as

$$x^{p^q} - x = \left(x^{p^n} - x\right)\left(x^{(p^n-1)(r-1)} + x^{(p^n-1)(r-2)} + \cdots + x^{p^n-1} + 1\right). \tag{6.6}$$

Construct $\mathbb{F}_{p^{kn}} = \mathbb{F}_{p^k}[x] / \langle f \rangle$, and let α be the corresponding root $x + \langle f \rangle$ of f. Fermat's Little Theorem tells us $\alpha^{p^n} = \alpha$. Equation (6.6) tells us $x^{p^n} - x$ divides $x^{p^q} - x$, so $x - \alpha$ is also a root of $x^{p^q} - x$. Similar reasoning implies $x^{p^q} - x$ divides $x^{p^a} - x$, so α is also a root of $x^{p^a} - x$. Thus, $(x - \alpha) \mid \gcd\left(f, x^{p^a} - x\right)$ in $\mathbb{F}_{p^n}$. By hypothesis, f is irreducible; the only divisors it has are 1 and f itself. But $x - \alpha$ divides the gcd, implying that $\gcd\left(f, x^{p^a} - x\right) = f$; in other words, $f \mid \left(x^{p^a} - x\right)$, as claimed. □

The Generalization of Fermat's Little Theorem for polynomials suggests an "easy" algorithm to compute the distinct degree factorization of $f \in \mathbb{Z}_p[x]$. See algorithm 6.1.

Theorem 6·96. *Algorithm 6.1 terminates with each p_i the product of the factors of f that are all of degree i.*

Proof. Note that the second and third steps of the loop are an optimization of the computation of $\gcd\left(x^{p^i} - x, f\right)$; you can see this by thinking about how the Euclidean algorithm would compute the gcd. So termination is guaranteed by the fact that eventually $\deg h_i^p > \deg f_i$: the generalization of Fermat's Little Theorem for polynomials implies that at this point, all distinct degree factors of f have been removed. Correctness is guaranteed by the fact that in each step we are computing $\gcd\left(x^{p^i} - x, f\right)$. □

Example 6·97. Returning to $\mathbb{Z}_5[x]$, let's look at

$$f = x(x+3)\left(x^3 + 4\right).$$

Do *not assume* whether this factorization is into irreducible elements. Expanded, $f = x^5 + 3x^4 + 4x^2 + 2x$. When we plug it into algorithm 6.1, the following occurs:

- For $i = 1$,
  - the remainder of division of $h_0^5 = x^5$ by f is $h_1 = 2x^4 + x^2 + 3x$;
  - $p_1 = x^3 + 2x^2 + 2x$;
  - $f_1 = x^2 + x + 1$.
- For $i = 2$,
  - the remainder of division of $h_1^5 = 2x^{20} + x^{10} + 3x^5$ by f is $h_2 = x$;
  - $p_2 = \gcd(0, f_1) = f_1$;
  - $f_2 = 1$.

Thus the distinct degree factorization of f is

$$f = \left(x^3 + 2x^2 + 2x\right)\left(x^2 + x + 1\right).$$

This demonstrates that the original factorization was not into irreducible elements, since $x(x+3)$ is not equal to either of the two new factors, so that $x^3 + 4$ must have a linear factor as well.

Question 6·98 .
Compute the distinct degree factorization of $f = x^5 + x^4 + 2x^3 + 2x^2 + 2x + 1$ in $\mathbb{Z}_5[x]$. This factorization is in irreducible elements; explain how we know this.

Question 6·99 .
Suppose that we don't want the factors of f, but only its roots. Explain how we can use $\gcd(x^p - x, f)$ to give us the maximum number of roots of f in $\mathbb{Z}_p$. Use the polynomial from Example 6.98 to illustrate your argument.

Equal degree factorization

Once we have a distinct degree factorization of $f \in \mathbb{Z}_p[x]$ as $f = p_1 \cdots p_m$, where each p_i is the product of the factors of degree i of a squarefree polynomial f, we need to factor each p_i into its irreducible factors. Here we consider the case that p is an odd prime; the case where $p = 2$ requires different methods.

Take any p_i, and let its factorization into irreducible polynomials of degree i be $p_i = q_1 \cdots q_n$. Suppose we select at random some $h \in \mathbb{Z}_p[x]$ with $\deg h < n$. If p_i and h share a common factor, then $\gcd(p_i, h) \neq 1$, and we have found a factor of p_i. Otherwise, we will try the following. Since each q_j is irreducible and of degree i, $\langle q_j \rangle$ is a maximal ideal in $\mathbb{Z}_p[x]$, so $\mathbb{Z}_p[x] / \langle q_j \rangle$ is a field with p^i elements. Denote it by $\mathbb{F}$.

Lemma 6·100. *Let G be the multiplicative group of nonzero elements of $\mathbb{F}$; that is, $G = \mathbb{F} \setminus \{0\}$. Let $a = \frac{p^i - 1}{2}$, and let $\varphi : G \to G$ by $\varphi(g) = g^a$.*

(A) φ is a group homomorphism of G.

(B) Its image, $\varphi(G)$, consists of the square roots of unity.

(C) $|\ker\varphi| = a$.

Proof. From the definition of a field, G is an abelian group under multiplication.

(A) Let $g, h \in G$. Since G is abelian,

$$\begin{aligned}\varphi(gh) = (gh)^a &= \underbrace{(gh)(gh)\cdots(gh)}_{a \text{ copies}} \\ &= \underbrace{(g\cdot g\cdots g)}_{a \text{ copies}} \cdot \underbrace{(h\cdot h\cdots h)}_{a \text{ copies}} \\ &= g^a h^a = \varphi(g)\,\varphi(h).\end{aligned}$$

(B) Let $y \in \varphi(G)$; by definition, there exists $g \in G$ such that

$$y = \varphi(g) = g^a.$$

Corollary 4·113 to Lagrange's Theorem, with the fact that $|G| = p^i - 1$, implies that

$$y^2 = (g^a)^2 = \left(g^{\frac{p^i-1}{2}}\right)^2 = g^{p^i-1} = 1.$$

We see that y is a square root of unity. We chose $y \in \varphi(G)$ arbitrarily, so *every* element of $\varphi(G)$ is a square root of unity.

(C) Observe that $g \in \ker\varphi$ implies $g^a = 1$, or $g^a - 1 = 0$. That makes g an ath root of unity. Since $g \in \ker\varphi$ was chosen arbitrarily, $\ker\varphi$ consists of ath roots of unity. By the Factor Theorem, each $g \in \ker\varphi$ corresponds to a linear factor $x - g$ of $x^a - 1$. There can be at most a such factors, so there can be at most a distinct elements of $\ker\varphi$; that is, $|\ker\varphi| \leqslant a$. Since $\varphi(G)$ consists of the square roots of unity, similar reasoning implies that there are at most two elements in $\varphi(G)$. Since G has $p^i - 1$ elements, the Isomorphism Theorem tells us that $G/\ker\varphi \cong \varphi(G)$, so $|G/\ker\varphi| = |\varphi(G)|$. That gives us

$$p^i - 1 = |G| = |\ker\varphi|\,|\varphi(G)| \leqslant a \cdot 2 = \frac{p^i-1}{\cancel{2}} \cdot \cancel{2} = p^i - 1.$$

The inequality is actually an equality, forcing $|\ker\varphi| = a$. □

To see how Lemma 6·100 is useful, consider a nonzero coset in $\mathbb{F}$,

$$[h] = h + \langle q_j \rangle \in \mathbb{F}.$$

As a field, $\mathbb{F}$ can have no zero divisors, so h can have no common factor with q_j. As q_j is irreducible, this gives us $h \notin \langle q_j \rangle$, so $[h] \neq 0_{\mathbb{F}}$, so $[h] \in G$. Raising $[h]$ to the ath power gives us an element of $\varphi(G)$. Part (B) of the lemma tells us that $\varphi(G)$ consists of the square roots of unity in G, so $[h]^a$ is a square root of $1_{\mathbb{F}}$, either $1_{\mathbb{F}}$ or $-1_{\mathbb{F}}$. If $[h]^a = 1_{\mathbb{F}}$, then $[h]^a - 1_{\mathbb{F}} = 0_{\mathbb{F}}$. Recall that $\mathbb{F}$ is a quotient ring, and $[h] = h + \langle q_j \rangle$. Thus

$$(h^a - 1) + \langle q_j \rangle = [h]^a - 1_{\mathbb{F}} = 0_{\mathbb{F}} \in \langle q_j \rangle.$$

Algorithm 6.2 Equal-degree factorization

inputs
$f \in \mathbb{Z}_p[x]$, where p is irreducible and odd, f is squarefree, $n = \deg f$, and all factors of f are of degree d
outputs
a factor q_i of f
do
Let $q = 1$
while $q = 1$ **do**
Let $h \in \mathbb{Z}_p[x] \setminus \mathbb{Z}_p$, with $\deg h < n$
Let $q = \gcd(h, f)$
if $q = 1$ **then**
Let h be the remainder from division of $h^{\frac{p^d-1}{2}}$ by f
Let $q = \gcd(h - 1, f)$
return q

This is a phenomenal consequence! Equality of cosets implies that $h^a - 1 \in \langle q_j \rangle$, so q_j divides $h^a - 1$. This means that $h^a - 1$ has at least q_j in common with p_i! Taking the greatest common divisor of $h^a - 1$ and p_i extracts the greatest common factor, which may be a multiple of q_j. This leads us to Algorithm 6.2. Note that there we have written f instead of p_i and d instead of i.

Algorithm 6.2 is a little different from previous algorithms, in that it requires us to select a random element. Not all choices of h have either a common factor with p_i, or an image $\varphi([h]) = 1_{\mathbb{F}}$. To get $q \neq 1$, we have to be "lucky". If we're extraordinarily unlucky, Algorithm 6.2 might never terminate. But this is highly unlikely, for two reasons. First, Lemma 6·100(C) implies that the number of elements $g \in G$ such that $\varphi(g) = 1$ is a. We have to have $\gcd(h, p_i) = 1$ to be unlucky, so $[h] \in G$. Observe that

$$a = \frac{p^i - 1}{2} = \frac{|G|}{2},$$

so we have less than 50% probability of being unlucky, and the cumulative probability decreases with each iteration. In addition, we can (in theory) keep track of which polynomials we have computed, ensuring that we never use an "unlucky" polynomial more than once.

Keep in mind that Algorithm 6.2 only returns *one* factor, and that factor might not be irreducible! This is not a problem, since

- we can repeat the algorithm on f/g to extract another factor of f;
- if $\deg q = d$, then q is irreducible; otherwise;
- $d < \deg q < n$, so we can repeat the algorithm in q to extract a smaller factor.

Since the degree of f or q decreases each time we feed it as input to the algorithm, the Well-Ordering Principle implies that we will eventually conclude with an irreducible factor.

Example 6·101. Recall from Example 6·97 that

$$f = x(x+3)\left(x^3+4\right) \in \mathbb{Z}_5[x]$$

gave us the distinct degree factorization

$$f = \left(x^3+2x^2+2x\right)\left(x^2+x+1\right).$$

The second polynomial is in fact the one irreducible quadratic factor of f; the first polynomial, $p_1 = x^3+2x^2+2x$, is the product of the irreducible linear factors of f. We use algorithm 6.2 to factor the linear factors.

- We have to pick $h \in \mathbb{Z}_5[x]$ with $\deg h < \deg p_1 = 3$. Let $h = x^2+3$.
  - Using the Euclidean algorithm, we find that h and f are relatively prime. (In particular, $r_1 = f-(x+2)h = 4x+4$, $r_2 = h-(4x+1)r_1 = 4$.)
  - The remainder of division of $h^{\frac{5^1-1}{2}}$ by f is $3x^2+4x+4$.
  - Now $q = \gcd\left((3x^2+4x+4)-1, p_1\right) = x+4$.
  - Return $x+4$ as a factor of p_1.

We did not know this factor from the outset! In fact, $f = x(x+3)(x+4)(x^2+x+1)$.

As with Algorithm 6.1, we need efficient algorithms to compute gcd's and exponents in order to perform Algorithm 6.2. Doing these as efficiently as possible is beyond the scope of these notes, but we do in fact have relatively efficient algorithms to do both: the Euclidean algorithm (Algorithm 5.1 on page 181) and fast exponentiation (Section 5·5).

Question 6·102 .
Use the distinct degree factorization of Example 6·97 and the fact that $f = x(x+3)(x^3+4)$ to find a complete factorization of f, using only the fact that you now know three irreducible factors f (two linear, one quadratic).

Squarefree factorization over a field of nonzero characteristic

Another approach to squarefree factorization is to combine the previous two algorithms in such a way as to guarantee that, once we identify an irreducible factor, we remove all powers of that factor from f before proceeding to the next factor. See Algorithm 6.3.

Example 6·103. In Question 6.104 you will try (and fail) to perform a distinct degree factorization on $f = x^5+x^3$ using only Algorithm 6.1. Suppose that we use algorithm 6.3 to factor f instead.

- Since f is monic, $b = 1$.
- With $i = 1$, distinct-degree factorization gives us $h_1 = 4x^3$, $q_1 = x^3+x$, $f_1 = x^2$.

Algorithm 6.3 Squarefree factorization in $\mathbb{Z}_p[x]$

inputs
$f \in \mathbb{Z}_p[x]$
outputs
An irreducible factorization $f = bp_1^{\alpha_1} \cdots p_m^{\alpha_m}$
do
Let $b = \mathrm{lc}(f)$
Let $h_0 = x$
Let $f_0 = b^{-1} \cdot f$ {After this step, f is monic}
Let $i = j = 0$
while $f_i \neq 1$ **do**
{One step of distinct degree factorization}
Increment i
Let h_i be the remainder of division of h_{i-1}^p by f
Let $q_i = \gcd(h_i - x, f_{i-1})$
Let $f_i = \frac{f_{i-1}}{q_i}$
{Find the equal degree factors of q_i}
while $q_i \neq 1$ **do**
Increment j
Find a degree-i factor p_j of q_i using algorithm 6.2
Let $q_i = \frac{q_i}{p_j}$
{Divide out all copies of p_j from f_i}
Let $\alpha_j = 1$
while p_j divides f_i **do**
Increment α_j
Let $f_i = \frac{f_i}{p_j}$
Let $m = j$
return $b, p_1, \ldots, p_m, \alpha_1, \ldots, \alpha_m$

- Suppose that the first factor that Algorithm 6.2 gives us is x. We can then divide f_1 twice by x, so $\alpha_j = 3$ and we conclude the innermost loop with $f_1 = 1$.
- Algorithm 6.2 subsequently gives us the remaining factors $x + 2$ and $x + 3$, none of which divides f_1 more than once..

The algorithm thus terminates with $b = 1, p_1 = x, p_2 = x+2, p_3 = x_3, \alpha_1 = 3$, and $\alpha_2 = \alpha_3 = 1$.

Question 6·104 .
Explain why Algorithm 6.1 might not work for $f = x^5 + x^3$. Then try the algorithm on f in $\mathbb{Z}_5[x]$, and explain why the result is incorrect.

6·8 Factoring integer polynomials

We conclude, at the end of this chapter, with factorization in $\mathbb{Z}[x]$. The previous section showed how to factor a polynomial in an arbitrary finite field whose characteristic is an odd irreducible integer. We can use this technique to factor a polynomial $f \in \mathbb{Z}[x]$. As in the previous section, this method is not necessarily the most efficient, but it does illustrate techniques that are used in practice.

We show this using the example

$$f = x^4 + 8x^3 - 33x^2 + 120x - 720.$$

Suppose f factors as

$$f = p_1^{\alpha_1} \cdots p_m^{\alpha_m}.$$

Now let $p \in \mathbb{N}^+$ be odd and irreducible, and consider $\widehat{f} \in \mathbb{Z}_p[x]$ such that the coefficients of $\widehat{f}$ are the coefficients of f mapped to their cosets in $\mathbb{Z}_p$. That is,

$$\widehat{f} = [1]_p x^4 + [8]_p x^3 + [-33]_p x^2 + [120]_p x + [-720]_p.$$

By the properties of arithmetic in $\mathbb{Z}_p$, we know that $\widehat{f}$ will factor as

$$\widehat{f} = \widehat{p}_1^{\alpha_1} \cdots \widehat{p}_m^{\alpha_m},$$

where the coefficients of each $\widehat{p}_i$ are the coefficients of p_i mapped to their cosets in $\mathbb{Z}_p$. As we will see, these $\widehat{p}_i$ might not be irreducible for each choice of p; we might have instead

$$\widehat{f} = \widehat{q}_1^{\beta_1} \cdots \widehat{q}_n^{\beta_n}$$

where each $\widehat{q}_i$ divides some $\widehat{p}_j$. Nevertheless, we will be able to recover the irreducible factors of f even from these factors; it will simply be more complicated. There are two possible solutions to this issue: using one big irreducible p, or several small irreducibles along with the Chinese Remainder Theorem.

Squarefree factorization over a field of characteristic zero

We first pause to discuss squarefree factorization in this context. When our ground field has characteristic 0, we can compute the formal derivative f' of f, then $g = \gcd(f, f')$. The quotient $\frac{f}{g}$ is then squarefree.

Example 6·105. Recall the polynomial of Example 6·94,

$$f(x) = x^{16} + x^{13} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^5 + x^2.$$

Its formal derivative is

$$f'(x) = 16x^{15} + 13x^{12} + 11x^{10} + 10x^9 + 9x^8 + 8x^7 + 7x^6 + 5x^4 + 2x.$$

The Euclidean algorithm tells us $g = \gcd(f, f') = x$, so f is not squarefree; as indicated earlier, we can continue by factoring both g (which in this case is trivial) and $h = f/g$.

This example also explains why we didn't use the formal derivative in the previous section: over $\mathbb{Z}_2$ a *lot* of terms in the derivative become zero! Which ones? the terms derived from those with even powers:

$$f'(x) \equiv x^{12} + x^{10} + x^8 + x^6 + x^4.$$

In this case, the gcd is x^2, and while we can factor that out of f, we cannot reduce x^2 itself to squarefree form, because its derivative is $2x \equiv 0$.

Question 6·106 .
Show that $\frac{f}{g}$ is squarefree if $f \in \mathbb{C}[x]$, f' is the usual derivative from Calculus, and $g = \gcd(f, f')$.

One big irreducible.

One approach is to choose an odd, irreducible $p \in \mathbb{N}^+$ sufficiently large that, once we factor $\widehat{f}$, the coefficient a_i of any p_i is either the corresponding coefficient in $\widehat{p}_i$ or (on account of the modulus) the largest negative integer corresponding to it. Sophisticated methods to obtain a good p exist, but for our purposes it suffices to choose p approximately twice the size of the maximum coefficient of f.

Example 6·107. The maximum coefficient in the example f given above is 720. There are several irreducible integers larger than 1440 and "close" to it. We'll try the closest one, 1447. Using the techniques of the previous section, we obtain the factorization in $\mathbb{Z}_{1447}[x]$

$$\widehat{f} = (x + 12)(x + 1443)\left(x^2 + 15\right) \in \mathbb{Z}_{1447}[x].$$

It is "obvious" that this cannot be the correct factorization in $\mathbb{Z}[x]$, because 1443 is too large. On the other hand, properties of modular arithmetic tell us that

$$\widehat{f} = (x + 12)(x - 4)\left(x^2 + 15\right) \in \mathbb{Z}_{1447}[x].$$

In fact,

$$f = (x + 12)(x - 4)\left(x^2 + 15\right) \in \mathbb{Z}[x].$$

This is why we chose an irreducible number that is approximately twice the largest coefficient of f: it will recover negative factors as integers that are "too large".

We mentioned above that we can get "false positives" in the finite field.

Example 6·108. Let $f = x^2 + 1$. In $\mathbb{Z}_5[x]$, this factors as $x^2 + [1]_5 = (x + [2]_5)(x + [3]_5)$, but certainly $f \neq (x+2)(x+3)$ in $\mathbb{Z}[x]$.

Avoiding this problem requires techniques that are beyond the scope of these notes. However, it is certainly easy enough to verify whether a factor of $\widehat{f}$ is a factor of f using division; once we find all the factors $\widehat{q}_j$ of $\widehat{f}$ that do not give us factors p_i of f, we can try combinations of them until they give us the correct factor. Unfortunately, this can be very time-consuming, which is why in general one would want to avoid this problem entirely.

Several small primes.

For various reasons, we may not want to try factorization modulo one large prime; in this case, it would be possible to factor using several small primes, then recover f using the Chinese Remainder Theorem. Recall that the Chinese Remainder Theorem tells us that if $\gcd(m_i, m_j) = 1$ for each $1 \leqslant i < j \leqslant n$, then we can find x satisfying

$$\begin{cases} [x] &= [\alpha_1] \text{ in } \mathbb{Z}_{m_1}; \\ [x] &= [\alpha_2] \text{ in } \mathbb{Z}_{m_2}; \\ &\vdots \\ [x] &= [\alpha_n] \text{ in } \mathbb{Z}_{m_n}; \end{cases}$$

and $[x]$ is unique in $\mathbb{Z}_N$ where $N = m_1 \cdots m_n$. If we choose $m_1, \ldots, m_n$ to be all irreducible, they will certainly satisfy $\gcd(m_i, m_j) = 1$; if we factor f in each $\mathbb{Z}_{m_i}$, we can use the Chinese Remainder Theorem to recover the coefficients of each p_i from the corresponding $\widehat{q}_j$.

Example 6·109. Returning to the polynomial given previously; we would like a unique solution in $\mathbb{Z}_{720}$ (or so). Unfortunately, the factorization $720 = 2^4 \cdot 3^2 \cdot 5$ is not very convenient for factorization. We can, however, use $3 \cdot 5 \cdot 7 \cdot 11 = 1155$:

- in $\mathbb{Z}_3[x]$, $\widehat{f} = x^3(x+2)$;
- in $\mathbb{Z}_5[x]$, $\widehat{f} = (x+1)(x+2)x^2$;
- in $\mathbb{Z}_7[x]$, $\widehat{f} = (x+3)(x+5)(x^2+1)$; and
- in $\mathbb{Z}_{11}[x]$, $\widehat{f} = (x+1)(x+7)(x^2+4)$.

If we examine all these factorizations, we can see that there appears to be a "false positive" in $\mathbb{Z}_3[x]$; we should have

$$f = (x+a)(x+b)\left(x^2+c\right).$$

The easiest of the coefficients to recover will be c, since it is unambiguous that

$$\begin{cases} c = [0]_3 \\ c = [0]_5 \\ c = [1]_7 \\ c = [4]_{11} \end{cases}$$

In fact, the Chinese Remainder Theorem tells us that $c = [15] \in \mathbb{Z}_{1155}$.

Recovering a and b is more difficult, as we have to guess "correctly" which arrangement of the coefficients in the finite fields gives us the arrangement corresponding to $\mathbb{Z}$. For example, the system

$$\begin{cases} b = [0]_3 \\ b = [1]_5 \\ b = [3]_7 \\ b = [1]_{11} \end{cases}$$

gives us $b = [276]_{1155}$, which turns out to be wrong, but the system

$$\begin{cases} b = [0]_3 \\ b = [2]_5 \\ b = [5]_7 \\ b = [1]_{11} \end{cases}$$

gives us $b = [12]_{1155}$, the correct coefficient in $\mathbb{Z}$.

The drawback to this approach is that, in the worst case, we would try $2^4 = 16$ combinations before we can know whether we have found the correct one. In practice, therefore, sophisticated criteria and techniques are used to reassemble f.

Question 6·110 .

Factor $x^7 + 8x^6 + 5x^5 + 53x^4 - 26x^3 + 93x^2 - 96x + 18$ using each of the two approaches described here.

Question 6·111 .

Let $f(x) = x^8 + 5x^7 + 9x^5 + 53x^4 + 40x^3 + 72x + 360$. We want to factor f over $\mathbb{Z}$ by first factoring over $\mathbb{Z}_p$ for some "good" values of p.

(a) Suppose we try to factor f over $\mathbb{Z}_3$. Someone might argue that this is actually a bad idea, because it gives *false positives;* that is, it allows *too much* factorization. *Why?*
Hint: Think about an "obvious" factorization of f when you write its coefficients modulo 3, and whether this "obvious" factorization also occurs over $\mathbb{Z}$.

(b) Based on the answer to part (a), what would be *bad* values of p?

(c) If we wanted to factor f over $\mathbb{Z}_p$ for several irreducibles p, then reconstruct the factorization over $\mathbb{Z}$ using the Chinese Remainder Theorem, without using for p any of the moduli you identified in (b), *and* you wanted to use

(i) the *smallest* p possible, how many such p would you want to use, and what are they?

(ii) the *fewest* p possible, how many such p would you want to use, and what are they?

Chapter 7

Some important, noncommutative groups and rings

We've identified a number of common structures shared by certain systems. Most of the systems we've studied have enjoyed commutative operations, but we generally worked without that assumption, when possible. We now consider systems that are built with a noncommutative operation. These appear in many interesting situations.

7·1 Functions

The systems we will consider in this chapter are organized primarily around *functions.* We've already defined a function, but we haven't made much use of the definition, so it can't hurt to remind ourselves what they are.

> A **function** is any relation $F \subsetneq S \times T$ such that every $s \in S$ corresponds to exactly one $(s,t) \in F$. If F is a function, we write $F : S \to T$ instead of $F \subseteq S \times T$, and $F(s) = t$ instead of $(s,t) \in F$. When $F(s) = t$, we call t the **image** of s, and s the **preimage** of t. Notice that we can *never* have $F(s) = t$ and $F(s) = u$ whenever $t \neq u$.

The **domain of** F in the definition above is S; the **range of** F is T in the definition above; and the **image of** F, or $\mathrm{Img}(F)$, is the subset of T whose elements actually have a preimage. More precisely,

$$\mathrm{Img}(F) = \{t \in T : f(s) = t \text{ for some } s \in S\}.$$

In this chapter, we generally assume the domain and range of a function are the same; in such cases, we say that F is a function **on** S.

Two functions are equal when they map every element of the same domain to the same image. That is, if $f : S \to T$ and $g : S \to T$, then $f = g$ if and only if $f(s) = g(s)$ for every $s \in S$.

Remark 7·1. A subtle difference distinguishes our definition of a function from the one you've encountered before: $F(s)$ must be defined for *every* $s \in S$. Thus, the following relations *are not* functions according to our definition:

- $F : \mathbb{N} \to \mathbb{N}$ by $F(n) = \sqrt{n}$, because $\sqrt{2} \notin \mathbb{N}$;
- $F : \mathbb{R} \to \mathbb{R}$ by $F(r) = 1/r$, because $1/0$ is undefined.

We *never* use the word "function" *unless* we know it is defined for every element of its domain. The usual convention is to use lower-case letters for function names, but try not to forget that functions are really a special subset of a Cartesian product.

Addition and multiplication of functions

Let S and T be sets where addition or multiplication is defined, though S and T need not be rings or groups. Let $f, g : S \to T$. We "add" two functions f and g and obtain a new function $f + g$ by identifying $(f + g)(s)$ with the sum of the images of s. That is,

$$(f + g)(s) = f(s) + g(s).$$

We "multiply" two functions f and g in a similar way. That is,

$$(fg)(s) = f(s) \cdot g(s).$$

Since f and g are both defined for all $s \in S$, the sum and product of f and g is also a function.

Example 7·2. Let $S = \mathbb{Z}$, and let f and g be functions on S such that $f(s) = s^2$ and $f(s) = s - 2$. Then $f + g : S \to S$ according to the rule

$$(f + g)(s) = f(s) + g(s) = s^2 + (s - 2),$$

while $fg : S \to S$ according to the rule

$$(fg)(s) = f(s) \cdot g(s) = s^2(s - 2).$$

Fact 7·3. *If T is a ring under addition and multiplication of its elements, then the set $\mathcal{F}$ of all functions mapping S to T is a ring under addition and multiplication of functions.*

Notice that S need not be a ring itself.

Why is this true? We already showed closure above. We now show some of the other properties, leaving the rest as an exercise. Assume that S is a ring under addition and multiplication of its elements.

associativity of addition? Let $f, g, h \in \mathcal{F}$. We need to show that $f + (g + h) = (f + g) + h$. This is true if the two functions $f + (g + h)$ and $(f + g) + h$ map every element of S to the same image. So, let $s \in S$. We use the fact that S is a ring to show that the associative property carries over to $\mathcal{F}$. For the left-hand sum,

$$(f + (g + h))(s) = f(s) + (g + h)(s) = f(s) + [g(s) + h(s)].$$

For the right-hand sum,

$$((f + g) + h)(s) = (f + g)(s) + h(s) = [f(s) + g(s)] + h(s).$$

We have translated from a problem about *functions* f, g, and h to one about $f(s)$, $g(s)$, and $h(s)$, which are *elements of the ring* T. Ring addition is associative, so

$$f(s) + [g(s) + h(s)] = [f(s) + g(s)] + h(s).$$

Substitution gives us

$$(f + (g + h))(s) = ((f + g) + h)(s).$$

Recall that s is an arbitrary element of S, so the above equation is true for *every* element of S, showing that the functions $f + (g + h)$ and $(f + g) + h$ are equal. Addition of functions is associative.

additive identity? Since T is a ring, it has an additive identity, 0. We claim that the additive identity of $\mathcal{F}$ is the **zero function**, $z : S \to S$ by $z(s) = 0$ for all $s \in S$. (We usually write this function simply as 0, but for the sake of pedantry, we use more function-like notation. Pedantry has its uses sometimes.) To show that this is indeed the additive identity, we must show that $z + f = f$ and $f + z = f$ for all $f \in \mathcal{F}$. So, let $f \in \mathcal{F}$. We must show that show that $z + f$, $f + z$, and f map every element of S to the same image. So, let $s \in S$. By substitution,

$$(z + f)(s) = z(s) + f(s) = 0 + f(s) = f(s)$$

and

$$(f + z)(s) = f(s) + z(s) = f(s) + 0 = f(s).$$

Substitution gives us

$$(f + z)(s) = f(s) = (z + f)(s).$$

Recall that s is an arbitrary element of S, so the above equation is true for *every* element of S, showing that the functions $f + z$, $z + f$, and f are equal. This shows that z really is the additive identity of addition of functions.

□

Question 7·4 .

Show that the remaining properties of a commutative ring are true; that is, addition of functions is commutative; every element has an additive inverse; multiplication of functions is associative; there is a multiplicative identity function; and multiplication of functions distributes over addition of functions. Show also that multiplication of functions is commutative if and only if T is a commutative ring.

Question 7·5 .

Show that $\mathcal{F}$ is *almost never* a field, even if T is a field! (To answer this question fully, you should give a very specific criterion that determines when $\mathcal{F}$ is a field, and that also shows there are "very few" $\mathcal{F}$ that can satisfy this requirement.)

Functions under composition

It is not uncommon to compose functions, yet even if students do it routinely, they often forget the term. Let S and T be *sets* (neither necessarily a ring) and suppose $f : S \to T$ and $g : T \to U$. The **composition** $g \circ f$ of functions f and g is a function that maps any $s \in S$ to $g(f(s))$.

Example 7·6. Suppose $S = \mathbb{N}$ and $f, g : S \to S$ by $f(s) = s^2$, $g(t) = t - 2$. Then $g \circ f : S \to S$ according to the rule

$$(g \circ f)(s) = g(f(s)) = g\left(s^2\right) = s^2 - 2,$$

while $f \circ g : S \to S$ according to the rule

$$(f \circ g)(s) = f(g(s)) = f(s-2) = (s-2)^2.$$

Under certain conditions, composition of functions can also be an operation on functions. The main point is that g needs to be defined on the image of f. This is easy to satisfy if the domain and range of both functions are the same, as in the example above.

So, let S be an arbitrary set, and let $\mathcal{F}$ be the set of all functions on S. Which properties of an operation does composition satisfy?

closure? Let $f, g \in \mathcal{F}$. Is $g \circ f \in \mathcal{F}$ also? Well, yes: g is defined for all $t \in S$, so for any $s \in S$, we know that $(g \circ f)(s) \in S$ regardless of the value of $f(s) = t$. Hence, $g \circ f$ will be defined for all $s \in S$, and is a function on S, so an element of $\mathcal{F}$.

associative? Let $f, g, h \in \mathcal{F}$. Is $h \circ (g \circ f) = (h \circ g) \circ f$? To find out, we have to make sure that *every* $s \in S$ has the same destination, regardless of which path it takes. So let $s \in S$. The left path, $h \circ (g \circ f)$, asks us to evaluate $(g \circ f)(s)$ first, and *then* evaluate h on the result. Let $u = (g \circ f)(s)$ and $v = h(u)$. The right path, $(h \circ g) \circ f$, asks us to evaluate $f(s)$ first, and *then* evaluate $h \circ g$ on the result. Let $t = f(s)$.

The question comes down to checking whether $v = (h \circ g)(t)$. By definition, $(h \circ g)(s) = h(g(t))$. What is $g(t)$? Recall that $u = (g \circ f)(s) = g(f(s)) = g(t)$. So, $g(t) = u$, which means $h(g(t)) = h(u) = v$. We have now verified that $v = (h \circ g)(t)$, so indeed

$$(h \circ (g \circ f))(s) = ((h \circ g) \circ f)(s).$$

Recall that s was chosen arbitrarily, so $h \circ (g \circ f) = (h \circ g) \circ f$ regardless of the preimage $s \in S$. The two functions are equal, and composition is associative.

What about the identity property? Define $\iota : S \to S$ by $\iota(s) = s$. As this is defined for all $s \in S$, $\iota \in \mathcal{F}$.

Question 7·7 .

Show that the function ι is also an identity for composition; that is, $\iota \circ f = f$ and $f \circ \iota = f$ for all $f \in \mathcal{F}$.

Question 7·8 .
On the other hand, show that not all functions have inverse elements. That is, find a set S and a function f on S such that the inverse of f, *if it exists,* is not a function *on S. Hint:* One such function appears earlier in this section.

Question 7·9 .
Also, show that composition of functions is not always commutative. That is, find a set S and two functions f and g on S such that $f \circ g \neq g \circ f$.

We have just shown the following important fact.

Fact 7·10. *The set of all functions on a set S is a monoid under the operation of composition. In general, it is not a group.*

Having said that, *some* sets of functions on a set S *are* a group; one has to choose the right subset of $\mathcal{F}$. Most of this chapter considers such special sets, but we first turn to an example you may not have expected.

Differentiation and integration

In Calculus, you learned about *differentiation* and *integration*. In some cases, we can consider them to be functions *of functions!* For now, we'll make life easy and work (mostly) with polynomials, which are functions. Let R be any ring.

Fact 7·11. *Differentiation with respect to x is a function on $R[x]$; that is, $\frac{d}{dx} : R[x] \to R[x]$. If R is a field of characteristic zero, then integration with respect to x is a function on $R[x]$.*

Why? Keeping in mind that elements of R are constant, this becomes straightforward with the familiar rules of differentiation and integration:

$$\frac{d}{dx}\left(r_n x^n + \cdots + r_1 x + r_0\right) = n r_n x^{n-1} + \cdots + 2 r_2 x + r_1 \in R[x]$$

and, if R is a field,

$$\int r_n x^n + \cdots + r_1 x + r_0 \, dx = (n+1)^{-1} r_n \cdot x^{n+1} + \cdots + 2^{-1} r_1 \cdot x^2 + r_0 x \in R[x].$$

Either way, the result is an element of $R[x]$, and the formula is deterministic, so differentiation and integration are defined for every element of the polynomial ring, showing that they are functions.

(Notice that we *do not* tack on an arbitrary constant of integration, but use zero, instead. This is okay for our purposes.) □

The upshot of this is that if you take the set of all functions of $R[x]$, then $\frac{d}{dx}$ and $\int dx$ are included. Building on what we wrote above, then, expressions such as $0 + \int dx$ and $\iota \circ \frac{d}{dx}$ actually make sense.

Question 7·12 .
Why must the field have characteristic zero for $\int dx$ to be a function?

7·2 Permutations

Certain applications of mathematics involve the rearrangement of a list of n elements. It is common to refer to such rearrangements as *permutations*.

Definition 7·13. A **list** is a sequence. Let V be any finite list. A **permutation** is a one-to-one, onto function whose domain and range are both V.

Throughout this section, V is a list of n elements, unless we say otherwise.

Example 7·14. If V is a list of all the elements in a finite group or ring, an isomorphism on V is a permutation.

The order of the elements matters in a permutation: the lists $(a, d, k, r) \neq (a, k, d, r)$ even though $\{a, d, k, r\} = \{a, k, d, r\}$. For the sake of convenience, we usually write V as a list of natural numbers between 1 and $|V|$, but it can be any finite list.

Example 7·15. The permutation of the list (red, green, blue) to (green, red, blue) is equivalent to the permutation of the list $(1, 2, 3)$ to $(2, 1, 3)$. We care only about the change in the entries' positions, not in the values of those entries.

The importance of permutations is twofold. First, group theory is a pretty neat and useful thing in itself, and we will see eventually that all finite groups can be modeled by groups of permutations. Anything that can model every possible group is by that very fact important.

Besides that, permutations relate to the factorization of polynomials. The polynomial $x^4 - 1$ can be factored over $\mathbb{C}$ as

$$(x + 1)(x - 1)(x + i)(x - i),$$

and

$$(x - 1)(x + 1)(x - i)(x + i).$$

On account of the commutative property, it doesn't matter what order we list the factors; this corresponds to a permutation, and is related to another idea that we will study, called field extensions. Field extensions can be used to solve polynomials equations, and since the order of the extensions doesn't really matter, permutations are important to determining the structure of the extension that solves a polynomial.

Permutations as functions

Example 7·16. Let $S = (a, d, k, r)$. Define a permutation on the elements of S by

$$f(x) = \begin{cases} r, & x = a; \\ a, & x = d; \\ k, & x = k; \\ d, & x = r. \end{cases}$$

Notice that f is one-to-one, and $f(S) = (r, a, k, d)$.

We can represent the same permutation on $V = (1, 2, 3, 4)$, a generic list of four elements. Define a permutation on the elements of V by

$$\pi(i) = \begin{cases} 2, & i = 1; \\ 4, & i = 2; \\ 3, & i = 3; \\ 1, & i = 4. \end{cases}$$

Here π is one-to-one, and $\pi(i) = j$ is interpreted as "the jth element of the permuted list is the ith element of the original list." You could visualize this as

position i in original list		position j in permuted list
1	$\rightarrow$	2
2	$\rightarrow$	4
3	$\rightarrow$	3
4	$\rightarrow$	1

Thus $\pi(V) = (4, 1, 3, 2)$. If you look back at $f(S)$, you will see that in fact the first element of the permuted list, $f(S)$, is the fourth element of the original list, S.

It should not surprise you that the identity function is a "do-nothing" permutation, just as it was a "do-nothing" symmetry of the triangle in Section 3·6.

Proposition 7·17. *Let V be a set of n elements. The function $I : V \to V$ by $I(x) = x$ is a permutation on V. In addition, for any permutation α on V, $I \circ \alpha = \alpha$ and $\alpha \circ I = \alpha$.*

Question 7·18 .
Why is Proposition 7·17 true?

As functions, the composition of two permutations is also a function. It gets a little better...

Lemma 7·19. *The composition of two permutations on a list of n elements is a permutation on the same list.*

Proof. Let V be a set of n elements, and α, β permutations of V. Let $\gamma = \alpha \circ \beta$. We claim that γ is a permutation. To show this, we must show that γ is a one-to-one function whose domain and range are both V. By definition, the domain and range of γ are both V; it remains to show that γ is one-to-one. Let $x, y \in V$ and assume that $\gamma(x) = \gamma(y)$; substituting the definition of γ,

$$\alpha(\beta(x)) = \alpha(\beta(y)).$$

Because they are permutations, α and β are one-to-one functions. Since α is one-to-one, we can simplify the above equation to

$$\beta(x) = \beta(y);$$

and since β is one-to-one, we can simplify the above equation to

$$x = y.$$

We assumed that $\gamma(x) = \gamma(y)$, and found that this forced $x = y$. By definition, γ is a one-to-one function. We already explained why its domain and range are both V, so γ is a permutation. □

In Example 7·16, we wrote a permutation as a piecewise function. This is burdensome; we would like a more efficient way to denote permutations.

Notation 7·20. The **tabular notation** for a permutation on a list of n elements is a $2 \times n$ matrix

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}$$

indicating that $\alpha(1) = \alpha_1$, $\alpha(2) = \alpha_2$, …, $\alpha(n) = \alpha_n$. Again, $\alpha(i) = j$ indicates that the jth element of the permuted list is the ith element of the original list.

Example 7·21. Recall V and π from Example 7·16. In tabular notation,

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

because π moves

- the element in the first position to the second;
- the element in the second position to the fourth;
- the element in the third position nowhere; and
- the element in the fourth position to the first.

Then

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

Notice that the tabular notation for π looks similar to the table in Example 7·16.

We can also use π to permute different lists, so long as the new lists have four elements:

$$\pi(3, 2, 1, 4) = (4, 3, 1, 2);$$
$$\pi(2, 4, 3, 1) = (1, 2, 3, 4);$$
$$\pi(a, b, c, d) = (d, a, c, b).$$

Question 7·22 .

For the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 2 & 4 & 6 & 3 \end{pmatrix},$$

(a) Evaluate $\alpha(1, 2, 3, 4, 5, 6)$.

(b) Evaluate $\alpha(1, 5, 2, 4, 6, 3)$.

(c) Evaluate $\alpha(6, 3, 5, 2, 1, 4)$.

Groups of permutations

Permutations form groups in a very natural way.

Definition 7·23. For $n \geqslant 2$, denote by S_n the set of all permutations of a list of n elements.

Example 7·24. For $n = 2, 3$ we have

$$S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$$

$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.$$

Is there some structure to S_n? By definition, a permutation is a one-to-one function. Fact 7·10 on page 262 tells us that the set of functions on a set is a monoid under composition of functions. The identity function is one-to-one, and the composition of one-to-one functions is also one-to-one, so S_n has an identity and is closed under composition. In addition, S_n inherits the associative property from the larger set of functions. Already, then, we can conclude that S_n is a monoid. However, one-to-one functions have inverses, which leads us to ask whether S_n is also a group.

Theorem 7·25. *For all $n \geqslant 2$ $(S_n, \circ)$ is a group.*

Notation 7·26. Normally we just write S_n, assuming that the operation is composition of functions. It is common to refer to S_n as the **symmetric group** of n elements.

Proof. Let $n \geqslant 2$. We have to show that S_n satisfies the properties of a group under the operation of composition of functions. Proposition 7·17 tells us that the identity function acts as an identity in S_n, and Lemma 7·19 tells us that S_n is closed under composition. We showed in Section 7·1 that functions are associative; as a subset of functions, S_n satisfies the associative property.

We still have to show that S_n satisfies the inverse property. Let V be a list of n elements. Let $\alpha \in S_n$. By definition of a permutation, α is one-to-one; since V is finite, α is onto. Since α is one-to-one, it has an inverse function α^{-1}, which satisfies the relationship that, for every $v \in V$,

$$\alpha^{-1}(\alpha(v)) = v \quad \text{and} \quad \alpha\left(\alpha^{-1}(v)\right) = v.$$

Since $I(v) = v$ for every $v \in V$, we have shown that $\alpha^{-1} \circ \alpha = \alpha \circ \alpha^{-1} = I$. The function α^{-1} is also a one-to-one, onto function on V, so $\alpha^{-1} \in S_n$! We chose α as an arbitrary permutation of n elements, so S_n satisfies the inverse property.

As claimed, S_n satisfies all four properties of a group. □

Question 7·27 .

Compute the order of

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

A final question: *how large is each S_n?* In other words, "how many permutations are there of n elements?" A counting argument called *the multiplication principle* shows that there are

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$$

such permutations. Why? Given any list of n elements,

- we have n positions to move the first element, including its current position;
- we have $n-1$ positions to move the second element, since the first element has already taken one spot;
- we have $n-2$ positions to move the third element, since the first and second elements have already take two spots;
- etc.

We have shown the following.

Lemma 7·28. *For each $n \in \mathbb{N}^+$, $|S_n| = n!$*

Question 7·29 .
How many elements are in S_4 and S_5? Try listing all the elements of S_4.

A hint of things to come.

You won't work extensively with groups of permutations just yet, but we mentioned earlier that *all* finite groups are really groups of permutations, and it's convenient to close here with an example of this.

Fact 7·30. *$S_3 \cong D_3$ as groups.*

Why? We map D_3 to S_3 by considering how each symmetry of the triangle permutes the three vertices. As long as this new point of view preserves the operation, the fact that D_3 and S_3 have six elements shows that such a map is a bijection, and we are done.

We take as our guide the labeling of the triangle's vertices from Section 3·6. For any sym-

metry $\sigma \in S_3$ map

$$f(\sigma) = \begin{cases} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \text{if } \sigma = \iota; \\ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, & \text{if } \sigma = \varphi; \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, & \text{if } \sigma = \rho; \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, & \text{if } \sigma = \rho^2; \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, & \text{if } \sigma = \rho\varphi; \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, & \text{if } \sigma = \rho^2\varphi. \end{cases}$$

This map is clearly a bijection. □

Question 7·31 .

Verify that $f(\varphi^2) = [f(\varphi)]^2, f(\rho^2) = [f(\rho)]^2, f(\rho^3) = [f(\rho)]^3$. (This will be a big deal in a moment.)

Explanation of Fact 7·30, continued: We must verify that f preserves the operation. We have thirty-six products to check, but with a little cleverness we can reduce this significantly. For instance, regardless of the value of σ,0

$$f(\sigma\iota) \underset{\text{subst}}{=} f(\sigma) = f(\sigma) \circ I,$$

and the same can be shown easily for $f(\iota\sigma)$, giving us 11 products more or less for free. That leaves twenty-five products to check. This quickly becomes a bit tedious, but remember that D_3 enjoys the property $\varphi\rho = \rho^2\varphi$. If this map is an isomorphism, we should expect a corresponding relationship in their images. To help us out, let's name $\alpha = f(\varphi)$ and $\beta = f(\rho)$. Indeed,

$$f(\varphi) f(\rho) = \alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

and by Question 7.31,

$$f(\rho^2) f(\varphi) = \beta^2\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

This allows us to rewrite all products of S_3 where $f(\varphi)$ appears before $f(\rho)$ exactly the same way we rewrite all products of D_3 where φ appears before ρ. We can verify the remaining twenty operations by mere substitution and application of these properties. For instance,

$$f(\varphi\rho^2) = f(\rho^4\varphi) = f((\rho^3\rho)\varphi) = f(\rho\varphi) = \beta\alpha$$

while

$$f(\varphi) f\left(\rho^2\right) = \alpha\beta^2 = (\alpha\beta)\beta = \left(\beta^2\alpha\right)\beta = \beta^2(\alpha\beta) = \beta^2\left(\beta^2\alpha\right) = \beta^4\alpha = \left(\beta^3\beta\right)\alpha = \beta\alpha,$$

so the two are equal, as desired. □

Question 7·32 .
Find an explicit isomorphism from S_2 to $\mathbb{Z}_2$. (It's not nearly as involved as Fact 7·30.)

7·3 Morphisms

We saw earlier that the set of all functions on a set S is a monoid under composition, even when S is not. What of the set $\mathcal{F}$ of all homomorphisms on a monoid S, or the set of all isomorphisms on S? Remember that "on S" means the domain *and range* are both S.

Homomorphisms

Theorem 7·33. *$\mathcal{F}$ is a monoid under composition.*

Proof. The proof of every property is identical to that of Fact 7·10, *except* for the closure and identity properties. The challenge with the identity property is easy to dispose of, as Question 4.74 shows the identity homomorphism I acts as an identity under composition. The challenge with the closure property is that the operation is *composition* of functions; so, we have to show that the set is closed under composition.

We begin with closure; we need to show that for any homomorphisms f and g on S, $f \circ g$ is also a homomorphism on S. To that end, let $f, g \in \mathcal{F}$, and define $h = f \circ g$. We need to show that $h \in \mathcal{F}$; in other words, $h(st) = h(s)\,h(t)$ for any $s, t \in S$. So, let $s, t \in S$. By substitution,

$$h(st) = (f \circ g)(st) = f(g(st)).$$

By definition, $g(st) = g(s)\,g(t)$, so we can rewrite the equation above as

$$h(st) = f(g(s)\,g(t)).$$

Let $x = g(s)$ and $y = g(t)$. By definition, $f(xy) = f(x) f(y)$, so we can rewrite the equation above as

$$h(st) = f(xy) = f(x) f(y) = [f(g(s))]\,[f(g(t))] = [(f \circ g)(s)]\,[(f \circ g)(t)] = h(s)\,h(t).$$

Recall that s and t were arbitrary in S; the ends of this equation show that h satisfies the homomorphism property. Since $h = f \circ g$, we have shown that the composition of f and g, two homomorphisms on S, is itself a homomorphism S. Since f and g were arbitrary homomorphisms on S, we have shown that composition of homomorphisms satisfies *closure* in the set of all homomorphisms on S. That is, $\mathcal{F}$ is closed.

We have shown that the set of all homomorphisms on a set S satisfies the associative, closure, and identity properties. This set forms a monoid. □

Is it also a group?

Fact 7·34. *If S is a monoid with more than one element, the set of homomorphisms on S is not a group.*

Why not? Suppose S has more than one element. Define a function f on S as $f(s) = я$; that is, f maps every element of S to the identity. We claim that f is a homomorphism: It clearly preserves the identity, since $f(я) = я$. As for preserving the operation: for any $s, t \in S$ we have

$$f(st) = я \qquad \text{and} \qquad f(s)f(t) = яя = я,$$

so $f(st) = f(s)f(t)$. On the other hand, f is not one-to-one; to see why, let $s, t \in S$ be *distinct* elements of S; that is, $s \neq t$. By definition, $f(s) = я$ and $f(t) = я$. As f is not one-to-one, its inverse cannot be a *function,* let alone a homomorphism. Thus, f^{-1} is not in the set $\mathcal{F}$ of homomorphisms on S. We have an element f of $\mathcal{F}$ with no inverse in $\mathcal{F}$; by definition, $\mathcal{F}$ is not a group. □

Question 7·35 .
The only monoid with one element is the set $\{я\}$, whose only element is the identity under whatever the operation may be. It is not a very interesting monoid, except that it sometimes serves as an exception to what would otherwise be a rule. Is this one of those cases? Is the set $\mathcal{F}$ of homomorphisms on $\{я\}$ a group?

Isomorphisms

We saw that the set $\mathcal{F}$ of homomorphisms on a monoid of at least two elements did not form a group, effectively because at least one homomorphism was not one-to-one. What if we looked only at one-to-one homomorphisms? For convenience, write $\mathcal{G}$ for this set.

Fact 7·36. *The set of one-to-one homomorphisms is also not a group under composition of functions.*

Why not? Consider the monoid of integers under addition. You found in Question 2.55(a) that the doubling function is an isomorphism between $\mathbb{Z}$ and $2\mathbb{Z}$. Since $2\mathbb{Z} \subsetneq \mathbb{Z}$, the doubling function is a homomorphism on $\mathbb{Z}$.

This has big implications. The doubling function is a one-to-one homomorphism on $\mathbb{Z}$, so it is in $\mathcal{G}$. If $\mathcal{G}$ is to be a group, the doubling function must have an inverse function in $\mathcal{G}$. What would that inverse function be? Let's call it g; to be an element of $\mathcal{G}$, g must be defined for every element of $\mathbb{Z}$. Sometimes, this is obvious: $\ldots, g(-2) = -1, g(0) = 0, g(2) = 1, g(4) = 2, \ldots$. But what is $g(1)$? It has to be an integer, but we just saw g covers every integer by some *even* integer. That means $g(1)$ *cannot exist,* because no possible images are left! But if $g(1)$ does not exist, g is not defined on all elements of S, so $g \notin \mathcal{G}$!

Now, g is the only possible inverse for the doubling function; since $g \notin \mathcal{G}$, the doubling function has no inverse in $\mathcal{G}$, despite being an element of $\mathcal{G}$ itself. The only possible conclusion is that $\mathcal{G}$ cannot be a group. □

The problem, of course, is that the doubling function is not onto. In Fact 7·34, the problem was that not all homomorphisms are one-to-one; in Fact 7·36, the problem was that not all one-to-one homomorphisms are onto. In a last-ditch attempt to obtain a *group,* we'll look at

the set of homomorphisms on S that are both one-to-one and onto. That is, we'll look at the set of isomorphisms.

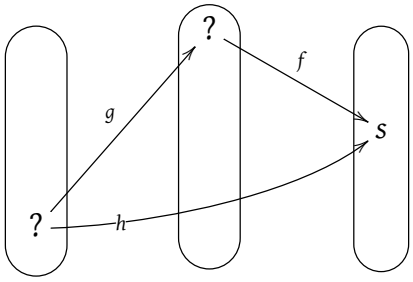
(This may look like an act of desperation, and from a certain point of view, it is. However, it is also how a lot of research works. You ask a question about a set, and when you find that it doesn't work for all elements of that set, you identify a subset for which you hope it does work. Unless you're completely off-base, this eventually leads to an interesting result. Along the way, you get additional interesting results, namely, why it does *not* work in other subsets.)

Theorem 7·37. *The set $\mathcal{H}$ of isomorphisms on a monoid is a group under composition of functions.*

Notice that we get a *group,* even though the underlying set is a monoid.

Proof. As with Theorem 7·33, we need only worry about that which is not covered by Fact 7·10: closure and inverse proprerties.

For **closure**, we need to show that composition of isomorphisms is still an *iso*morphism. Let $f, g \in \mathcal{H}$; that is, let f and g be any two isomorphisms on S. Let $h = f \circ g$; we need to show that $h \in \mathcal{H}$; in other words, that h is an isomorphism. We have already showed that it is a homomorphism in the proof of Theorem 7·33, so we need only show that it is one-to-one and onto. We show first that it is onto; let $s \in S$. We need to find an element of S whose image under h is s; we'll try to use the structure of h to do this. Composition means that we can calculate h by going through g and f:



Because f is onto, we can find $t \in S$ such that $f(t) = s$; because g is onto, we can find $u \in S$ such that $g(u) = t$. By definition,

$$h(u) = (f \circ g)(u) = f(g(u)) = f(t) = s,$$

so we have found an element of S that h maps to s. Since s was arbitrary in S, h is onto.

To show that h is one-to-one, let $s, t \in S$, and suppose

$$h(s) = h(t).$$

We need to show that $s = t$. By definition, $h(s) = (f \circ g)(s) = f(g(s))$, and likewise $h(t) = f(g(t))$, so by substitution into our supposition we know

$$f(g(s)) = f(g(t)).$$

To make the next steps a little easier, write $y = g(s)$ and $z = g(t)$. The equation immediately above becomes $f(y) = f(z)$. By hypothesis, f is one-to-one; by definition, $y = z$. We now have

$$g(s) = g(t).$$

Again, g is one-to-one; by definition,

$$s = t.$$

Since s and t are arbitrary in S, this holds for *all* elements of S.

We have shown that $h = f \circ g$ is one-to-one and onto; since f and g were arbitrary in $\mathcal{H}$, we see that $\mathcal{H}$ is closed under composition of functions.

It remains to show the **inverse** property. Let $f \in \mathcal{H}$. By definition, f is a one-to-one function, so it has an inverse function on S, f^{-1}.

Question 7·38 .
Show that f^{-1} is a function on S, and that it is also one-to-one and onto.

Proof of Theorem 7·37, continued: the inverse property. We need to show that f^{-1} is a homomorphism. To that end, let $s, t \in S$; we need to show that $f^{-1}(st) = f^{-1}(s) f^{-1}(t)$. To accomplish this, we change our perspective, and exploit the fact that f is a homomorphism. Let $x, y, z \in S$ such that $f(x) = st$, $f(y) = s$, and $f(z) = t$. In the Cayley table, we know the row and column headers match up, but not the body of the table.

$\ast$	$\cdots$	z	$\cdots$
$\vdots$			
y		yz	
$\vdots$			

$\overset{f}{\longrightarrow}$
$\overset{f^{-1}}{\longleftarrow}$

$\ast$	$\cdots$	t	$\cdots$
$\vdots$			
s		$f(x)$	
$\vdots$			

If we can show that $yz = x$, the body of the table also matches up, and f^{-1} is a homomorphism.

Recall that f is an isomorphism; by the homomorphism property,

$$f(yz) = f(y) f(z) = st.$$

However, we also know that

$$f(x) = st,$$

so by substitution,

$$f(x) = f(yz).$$

Isomorphisms are one-to-one, which makes f one-to-one. By definition, $x = yz$. This was precisely our goal, as substitution gives us

$$f^{-1}(st) = f^{-1}(s) f^{-1}(t).$$

We have show that f^{-1}, a one-to-one and onto function on S, is also a homomorphism. That makes it an isomorphism, so $f^{-1} \in \mathcal{H}$. We chose f arbitrarily in $\mathcal{H}$, so $\mathcal{H}$ satisfies the inverse property. We have shown that the set $\mathcal{H}$ of isomorphisms on S satisfies the four properties of a group. □

The set of isomorphisms on a monoid or group — that is, its domain and range are the same — is important enough to merit a special name. We call it the set of **automorphisms** on the monoid or group, and denote it $\mathrm{Aut}(S)$.

Question 7·39 .

Recall from Question that $f : \mathbb{Z} \to \mathbb{Z}_d$ is a homomorphism of rings, but not an isomorphism.

(a) Determine $\ker f$. (It might help to use a specific value of n first.)

(b) Indicate how we know that $\mathbb{Z}/\ker f \cong \mathbb{Z}_n$. (Eventually, we will show that $G/\ker f \cong H$ for *any* homomorphism $f : G \longrightarrow H$ that is onto.)

Chapter 8

Groups of permutations

You met permutations in Section 7·2. The main goal of this chapter is to show that groups of permutations are, in some sense, "all there is" to group theory, so in some sense, Section 8·2 is the point. Section 8·1 develops a convenient way to denote permutations, and 8·3 introduce you to two special classes of groups of permutation. We conclude with a great example of an application of symmetry groups in Section 8·4.

8·1 Cycle notation

Tabular notation of permutations is rather burdensome; a simpler and more compact notation is possible.

Cycles

Definition 8·1. A **cycle** is a vector

$$\alpha = (\alpha_1\ \alpha_2\ \cdots\ \alpha_n)$$

that corresponds to the permutation where the entry in position α_1 is moved to position α_2; the entry in position α_2 is moved to position α_3, ... and the element in position α_n is moved to position α_1. If a position is not listed in α, then the entry in that position is not moved. We call such positions **stationary**. For the identity permutation where no entry is moved, we write

$$\mathfrak{a} = I = (1)\,.$$

The fact that the permutation α moves the entry in position α_n to position α_1 is the reason we call it a *cycle;* applying it repeatedly cycles the list of elements around, and on the nth application the list returns to its original order.

Example 8·2. Example 7·21 considered the following permutation in tabular notation,

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

To write it as a cycle, we can start with any position of the list we like. However, the convention is to start with the smallest position affected by the permutation. Since π moves elements out of position 1, we start with

$$\pi = (1\ ?)\,.$$

The second entry in the cycle tells us where π moves the element in the position indicated by the first entry, 1. From the tabular notation, we see that π moves the element in position 1 to position 2, so

$$\pi = (1\ 2\ ?)\,.$$

The third entry of cycle notation tells us where π moves the element in the position indicated by the second entry. The second entry indicates position 2. From the tabular notation, we see that π moves the element in position 2 to position 4, so

$$\pi = (1\ 2\ 4\ ?)\,.$$

The fourth entry of cycle notation tells us where π moves the element in the position indicated by the third entry. The third element indicates position 4. From the tabular notation, we see that π moves the element in position 4 to position 1, so you might feel the temptation to write

$$\pi = (1\ 2\ 4\ 1\ ?)\,,$$

but that misses the entire point of cycle notation: we have now returned to the position indicated by the first entry, so we close the cycle:

$$\pi = (1\ 2\ 4)\,.$$

The cycle $(1\ 2\ 4)$, indicates that

- the element in position 1 of a list moves to position 2;
- the element in position 2 of a list moves to position 4;
- the element in position 4 of a list moves to position 1.

What about the element in position 3? Since it doesn't appear in the cycle notation, it must be stationary. This agrees with what we wrote in the piecewise and tabular notations for π.

Question 8·3 .

Write every element of S_3 in cycle notation.

Question 8·4 .

Let $\alpha = (1\ 2\ 3)$ and $\beta = (2\ 3)$. In this problem, you will verify two things about α and β.

(a) Verify that they are the same as the permutations designated α and β in the proof of Fact 7·30 on page 267.

(b) All elements of S_3 can be written as compositions of α and β; that is, every element of S_3 (in cycle form) has the form $\alpha^i\beta^j$ where $0 \leqslant i < 3$ and $0 \leqslant j < 2$.

Question 8·5 .
For α and β as defined in Question 8.4, show that $\beta \circ \alpha = \alpha^2 \circ \beta$. (Notice that $\alpha, \beta \in S_n$ for all $n > 2$, so as a consequence of this exercise S_n is not abelian for $n > 2$.)

Question 8·6 .
Write the Cayley table for S_3, with every element in the form $\alpha^i\beta^j$ where $0 \leqslant i < 3$ and $0 \leqslant j < 2$. Question 8.5 will make your life easier.

Question 8·7 .
In Fact 7·30, we showed $D_3 \cong S_3$ by mapping ρ and φ of D_3 to elements α and β of S_3. We used tabular notation at that time. Work through the proof again, this time using cycle notation instead of tabular notation.

Not all permutations can be written as one cycle.

Example 8·8. Consider the permutation in tabular notation

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

We can easily start the cycle with $\alpha = (1\ 2)$, and this captures the behavior on the elements in the first and second positions of a list, but what about the third and fourth positions? We cannot write $(1\ 2\ 3\ 4)$; that would imply that the element in the second position is moved to the third, and the element in the fourth position is moved to the first.

To solve this difficulty, we develop a simple arithmetic of cycles.

Cycle arithmetic

What operation should we apply to cycles? Cycles represent permutations; permutations are functions; functions can be *composed*. Hence, the appropriate operation is *composition*.

Example 8·9. Consider the cycles

$$\beta = (2\ 3\ 4) \quad \text{and} \quad \gamma = (1\ 2\ 4).$$

What is the cycle notation for

$$\beta \circ \gamma = (2\ 3\ 4) \circ (1\ 2\ 4)?$$

Cycles represent permutations, and permutations are closed under composition, tellingus $\beta \circ \gamma$ is also a permutation. With any luck, it will be a permutation that we can write as a cycle. What we need to do, then, is determine how the permutation $\beta \circ \gamma$ moves a list of four elements around. If that permutation can be represented as a cycle, then we've answered the question.

Since an element in the first position is moved, we should be able to write

$$\beta \circ \gamma = (1\ ?).$$

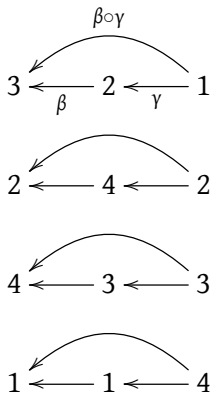


Figure 8·1: Diagram of how $\beta\circ\gamma$ modifies a list of four elements, for $\beta = (2\ 3\ 4)$ and $\gamma = (1\ 2\ 4)$.

Where is this first element moved? Let's apply the definition of composition: $\beta\circ\gamma$ means, "first apply γ; then apply β." Figure 8·1 show the basic idea; we refer to it throughout the example. The first cycle entry considers the first element of the list, or the top row of Figure 8·1; γ moves the first element to the second position, and β then moves it to the third. It must be that $\beta \circ \gamma$ moves an element from the first position to the third. We now know that

$$\beta \circ \gamma = (1\ 3\ ?)\,.$$

The second cycle entry should tell us where $\beta \circ \gamma$ moves an element that starts in the *third* position (*not* the second), illustrated by the *third* row of Figure 8·1. Applying the definition of composition again, we know that γ moves an element from the third position to… well, nowhere, actually. So an element in the third position *doesn't* move under γ; if we then apply β, however, it moves to the fourth position. It must be that $\beta \circ \gamma$ moves an element from the third position to the fourth. We now know that

$$\beta \circ \gamma = (1\ 3\ 4\ ?)\,.$$

The third cycle entry should tell tell us where $\beta\circ\gamma$ moves and element that starts in the fourth position, illustrated by the fourth row of Figure 8·1. We know that γ moves an element in the fourth position to the first position (4 is at the end of the cycle, so it moves to the beginning), and β moves elements in the first position… well, nowhere, actually. So $\beta \circ \gamma$ moves elements from the fourth position to the first position. This completes the cycle, so we now know that

$$\beta \circ \gamma = (1\ 3\ 4)\,.$$

Haven't we missed something? What about an element that starts in the second position? Since γ moves elements in the second position to the fourth, and β moves elements from the fourth position to the second, they undo each other, and the second position is stationary. It is, therefore, *absolutely correct* that 2 does not appear in the cycle notation of $\beta \circ \gamma$, and we see this in the *second* row of Figure 8·1.

Another phenomenon occurs when each permutation moves elements that the other does not.

Example 8·10. Consider the two cycles

$$\beta = (1\ 3) \quad \text{and} \quad \gamma = (2\ 4)\,.$$

There is no way to simplify $\beta \circ \gamma$ into a *single* cycle, because β operates only on the first and third elements of a list, and γ operates only on the second and fourth elements of a list. The only way to write them is as the composition of two cycles,

$$\beta \circ \gamma = (1\ 3) \circ (2\ 4)\,.$$

This motivates the following.

Definition 8·11. Two cycles are **disjoint** if none of their entries is common.

Disjoint cycles enjoy an important property: their permutations commute under composition.

Lemma 8·12. *Let α, β be two disjoint cycles. Then $\alpha \circ \beta = \beta \circ \alpha$.*

Proof. Let $n \in \mathbb{N}^+$ be the largest entry in α or β. Let $V = (1, 2, \ldots, n)$. Let $i \in V$. We consider the following cases:

Case 1. $\alpha(i) \neq i$.

Let $j = \alpha(i)$. The definition of cycle notation implies that j appears immediately after i in the cycle α. The definition of "disjoint" means that, since i and j are entries of α, they cannot be entries of β. By definition of cycle notation, $\beta(i) = i$ and $\beta(j) = j$. Hence

$$(\alpha \circ \beta)(i) = \alpha(\beta(i)) = \alpha(i) = j = \beta(j) = \beta(\alpha(i)) = (\beta \circ \alpha)(i)\,.$$

Case 2. $\alpha(i) = i$.

Subcase (a): $\beta(i) = i$.

We have $(\alpha \circ \beta)(i) = i = (\beta \circ \alpha)(i)$.

Subcase (b): $\beta(i) \neq i$.

Let $j = \beta(i)$. The definition of cycle notation implies that j appears immediately after i in the cycle β. The definition of "disjoint" means that, since i and j are entries of β, they cannot be entries of α. By definition of cycle notation, $\alpha(j) = j$. Hence

$$(\alpha \circ \beta)(i) = \alpha(\beta(i)) = \alpha(j) = j = \beta(i) = \beta(\alpha(i)) = (\beta \circ \alpha)(i)\,.$$

In both cases, we had $(\alpha \circ \beta)(i) = (\beta \circ \alpha)(i)$. Since i was arbitrary, $\alpha \circ \beta = \beta \circ \alpha$. □

Notation 8·13. Since the composition of two disjoint cycles $\alpha \circ \beta$ cannot be simplified, we normally write it without the circle; for example,

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

By Lemma 8·12, we can also write this as

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

That said, the usual convention for cycles is to write the smallest entry of a cycle first, and to order cycles by their first entries. We prefer

$$(1\ 4)\,(2\ 3)$$

to either of

$$(1\ 4)\,(3\ 2) \quad \text{or} \quad (2\ 3)\,(1\ 4)\,.$$

The convention for writing a permutation in cycle form is the following:

1. The first entry in each cycle is the cycle's smallest.
2. Simplify the composition of non-disjoint cycles, discarding those of length 1.
3. The remaining cycles are disjoint. They commute by Lemma 8·12; write them in order of the cycles' first entries.

Example 8·14. We return to Example 8·8, with

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

To write this permutation in cycle notation, we begin again with

$$\alpha = (1\ 2)\ldots ?$$

Since α also moves entries in positions 3 and 4, we need to add a second cycle. We start with the smallest position whose entry changes position, 3:

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

Since α moves the element in position 3 to position 4, we write

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

Now α moves the element in position 4 to position 3, so we can close the second cycle:

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

Now α moves no more entries, so the cycle notation is complete.

Permutations as cycles

We have come to the main result of this section.

Theorem 8·15. *Every permutation can be written as a composition of disjoint cycles.*

The proof is constructive; we build the cycle notation for the permutation.

Proof. Let $\pi \in S_n$. If $\pi(i) = i$ for all $i = 1, \ldots, n$, then we can write $\pi = (1)$. Otherwise, we can find $i_1 \in \{1, \ldots, n\}$ such that $\pi(i_1) \neq i_1$. Let i_1 be the smallest such number. As S_n is finite, we know $\pi^k = (1) = I$ for some $k \in \{2, \ldots, n\}$, and then $\pi^k(i_1) = I(i_1) = i_1$. Let

$$\alpha^{(1)} = \left(i_1 \ \pi(i_1) \ \ \pi(\pi(i_1)) \ \cdots \ \pi^{k-1}(i_1)\right).$$

By construction, $\alpha^{(1)}$ correctly describes how π moves elements in positions i_1, $\pi(i_1)$, …, $\pi^{k-1}(i_1)$.

If $\pi = \alpha^{(1)}$, then we are done. Otherwise, we can find $i_2 \in \{1, \ldots, n\} \setminus \left\{i_1, \pi(i_1), \ldots, \pi^{k-1}(i_1)\right\}$ such that $\pi(i_2) \neq i_2$. Choose the smallest such i_2, and let $\alpha^{(2)} = \left(i_2 \ \pi(i_2) \ \ \pi(\pi(i_2)) \ \cdots \ \pi^{\ell-1}(i_2)\right)$, where, as before $\pi^{\ell}(i_2) = i_2$.

Repeat this process until every non-stationary element of V appears in a cycle, generating $\alpha^{(3)}$, …, $\alpha^{(m)}$ for non-stationary $i_3 \notin \alpha^{(1)}, \alpha^{(2)}$, $i_4 \notin \alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}$, and so on until $i_m \notin \alpha^{(1)}, \ldots, \alpha^{(m-1)}$. There are only finitely many numbers π can move, so this process will not continue indefinitely, and concludes with a finite list of cycles.

The remainder of the proof consists of two claims.

Claim 1: Each cycles is disjoint from any other.

By way of contradiction, assume that two cycles $\alpha^{(i)}$ and $\alpha^{(j)}$ are not disjoint. Without loss of generality, assume $i < j$. Let c be the first entry in $\alpha^{(j)}$ that also appears in $\alpha^{(i)}$; by construction, it is not the first element of $\alpha^{(j)}$, so let a be the entry that precedes c in $\alpha^{(i)}$, and b the entry that precedes c in $\alpha^{(j)}$. By construction, $\pi(a) = c = \pi(b)$. Permutations are one-to-one, so $a = b$, but then b appears in $\alpha^{(i)}$, contradicting our choice of c as the *first* entry of $\alpha^{(j)}$ that appears in $\alpha^{(i)}$. Hence, $\alpha^{(i)}$ and $\alpha^{(j)}$ are disjoint.

Claim 2: $\pi = \alpha^{(1)}\alpha^{(2)} \cdots \alpha^{(m)}$.

Let $i \in V$. We consider two cases.

If $\pi(i) = i$, then i could not have been used to begin construction of an α. Since π is a one-to-one function, we cannot have $\pi(k) = i$ for any $k \neq i$, either. By construction, i appears in none of the $\alpha^{(j)}$. By substitution, the expression $\left(\alpha^{(1)}\alpha^{(2)} \cdots \alpha^{(m)}\right)(i)$ simplifies to

$$\left(\alpha^{(1)}\alpha^{(2)} \cdots \alpha^{(m)}\right)(i) = \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots \alpha^{(m)}(i)\right)\right) = i = \pi(i).$$

So far, so good.

Assume, then, that $\pi(i) \neq i$. By construction, i appears in $\alpha^{(j)}$ for some $j = 1, 2, \ldots, m$. By definition, $\alpha^{(j)}(i) = \pi(i)$, so both i and $\pi(i)$ appear in $\alpha^{(j)}$. By Claim 1, i and $\pi(i)$ do not appear in $\alpha^{(k)}$ whenever $k \neq j$. That guarantees $\alpha^{(k)}(i) = i$ and $\alpha^{(k)}(\pi(i)) = \pi(i)$. By substitution,

the expression $\left(\alpha^{(1)}\alpha^{(2)}\cdots\alpha^{(m)}\right)(i)$ simplifies to

$$\begin{aligned}\left(\alpha^{(1)}\alpha^{(2)}\cdots\alpha^{(m)}\right)(i) &= \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots\alpha^{(m-1)}\left(\alpha^{(m)}(i)\right)\right)\right)\\ &\vdots\\ &= \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots\alpha^{(j-1)}\left(\alpha^{(j)}(i)\right)\right)\right)\\ &= \alpha^{(1)}\left(\alpha^{(2)}\left(\cdots\alpha^{(j-1)}(\pi(i))\right)\right)\\ &\vdots\\ &= \pi(i).\end{aligned}$$

We have shown that

$$\left(\alpha^{(1)}\alpha^{(2)}\cdots\alpha^{(m)}\right)(i) = \pi(i).$$

Since i is arbitrary, $\pi = \alpha^{(1)}\circ\alpha^{(2)}\circ\cdots\circ\alpha^{(m)}$. That is, π is a composition of cycles. Since π was arbitrary, every permutation is a composition of cycles. □

Example 8·16. Consider the following permutation written in tabular notation,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\\ 7 & 5 & 3 & 2 & 4 & 8 & 1 & 6\end{pmatrix}.$$

The proof of Theorem 8·15 constructs the cycles

$$\begin{aligned}\alpha^{(1)} &= (1\ 7)\\ \alpha^{(2)} &= (2\ 5\ 4)\\ \alpha^{(3)} &= (6\ 8).\end{aligned}$$

Notice that $\alpha^{(1)}$, $\alpha^{(2)}$, and $\alpha^{(3)}$ are disjoint. In addition, the only element of $V = (1, 2, \ldots, 8)$ that does not appear in an α is 3, because $\pi(3) = 3$. Inspection verifies that

$$\pi = \alpha^{(1)}\alpha^{(2)}\alpha^{(3)}.$$

We conclude with some examples of simplifying the composition of permutations.

Example 8·17. Let $\alpha = (1\ 3)(2\ 4)$ and $\beta = (1\ 3\ 2\ 4)$. Notice that $\alpha \neq \beta$; check this on $V = (1, 2, 3, 4)$ if this isn't clear. In addition, α and β are not disjoint.

1. We compute the cycle notation for $\gamma = \alpha\circ\beta$. We start with the smallest entry moved by either α or β:

   $$\gamma = \begin{pmatrix} 1 & ? \end{pmatrix}.$$

   The notation $\alpha\circ\beta$ means to apply β first, *then* α. What does β do with the entry in position 1? It moves it to position 3. Subsequently, α moves the entry in position 3 back to the entry in position 1. The next entry in the first cycle of γ should thus be 1, but that's also the first entry in the cycle, so we close the cycle. So far, we have

   $$\gamma = (1)\ldots?$$

We aren't finished, since α and β also move other entries around. The next smallest entry moved by either α or β is 2, so

$$\gamma = (1)\,(2\ ?)\,.$$

Now β moves the entry in position 2 to the entry in position 4, and α moves the entry in position 4 to the entry in position 2. The next entry in the second cycle of γ should thus be 2, but that's also the first entry in the second cycle, so we close the cycle. So far, we have

$$\gamma = (1)\,(2)\ldots ?$$

Next, β moves the entry in position 3, so

$$\gamma = (1)\,(2)\,(3\ ?)\,.$$

Where does β move the entry in position 3? To the entry in position 2. Subsequently, α moves the entry in position 2 to the entry in position 4. We now have

$$\gamma = (1)\,(2)\,(3\ 4\ ?)\,.$$

You can probably guess that 4, as the largest possible entry, will close the cycle, but to be safe we'll check: β moves the entry in position 4 to the entry in position 1, and α moves the entry in position 1 to the entry in position 3. The next entry of the third cycle will be 3, but this is also the first entry of the third cycle, so we close the third cycle and

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

Finally, we simplify γ by not writing cycles of length 1, so

$$\gamma = (3\ 4)\,.$$

Hence

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

2. Now we compute the cycle notation for $\beta \circ \alpha$, but with less detail. Again we start with 1, which α moves to 3, and β then moves to 2. So we start with

$$\beta \circ \alpha = (1\ 2\ ?)\,.$$

Next, α moves 2 to 4, and β moves 4 to 1. This closes the first cycle:

$$\beta \circ \alpha = (1\ 2)\ldots ?$$

We start the next cycle with position 3: α moves it to position 1, which β moves back to position 3. This generates a length-one cycle, so there is no need to add anything. Likewise, the element in position 4 is also stable under $\beta \circ \alpha$. Hence we need write no more cycles;

$$\beta \circ \alpha = (1\ 2)\,.$$

3. Let's look also at $\beta \circ \gamma$ where $\gamma = (1\ 4)$. We start with 1, which γ moves to 4, and then β moves to 1. Since $\beta \circ \gamma$ moves 1 to itself, we don't have to write 1 in the cycle. The next smallest number that appears is 2: γ doesn't move it, and β moves 2 to 4. We start with

$$\beta \circ \gamma = (2\ 4\ ?)\,.$$

Next, γ moves 4 to 1, and β moves 1 to 3. This adds another element to the cycle:

$$\beta \circ \gamma = (2\ 4\ 3\ ?)\,.$$

We already know that 1 won't appear in the cycle, so you might guess that we should not close the cycle. To be certain, we consider what $\beta \circ \gamma$ does to 3: γ doesn't move it, and β moves 3 to 2. The cycle is now complete:

$$\beta \circ \gamma = (2\ 4\ 3)\,.$$

Question 8·18 .
List the elements of S_4 using cycle notation.

Question 8·19 .
Identify at least one normal subgroup of S_3, and at least one subgroup that is not normal.

Question 8·20 .
Compute the cyclic subgroup of S_4 generated by $\alpha = (1\ 3\ 4\ 2)$. Compare your answer to that of Question 7.27.

Question 8·21 .
Let $\alpha = (\alpha_1\ \alpha_2\ \cdots\ \alpha_m) \in S_n$. (Note $m \leqslant n$.) Show that we can write α^{-1} as

$$\beta = (\alpha_1\ \alpha_m\ \alpha_{m-1}\ \cdots\ \alpha_2)\,.$$

For example, if $\alpha = (2\ 3\ 5\ 6)$, $\alpha^{-1} = (2\ 6\ 5\ 3)$.

8·2 Cayley's remarkable result

The mathematician Arthur Cayley discovered a lovely fact about the permutation groups. Its effective consequence is that the theory of finite groups is equivalent to the study of groups of permutations.

Cayley's Theorem. *Every group of order n is isomorphic to a subgroup of S_n.*

Before we prove the theorem, we use an example to illustrate the idea behind the proof.

Example 8·22. Consider the Klein 4-group; this group has four elements, so Cayley's Theorem tells us that it must be isomorphic to a subgroup of S_4. We build an isomorphism by looking at the Cayley table for the Klein 4-group:

$\times$	$я$	a	b	ab
$я$	$я$	a	b	ab
a	a	$я$	ab	b
b	b	ab	$я$	a
ab	ab	b	a	$я$

To find a permutation appropriate to each element, we first label the elements:

$$\begin{aligned} я &\leftrightsquigarrow 1, \\ a &\leftrightsquigarrow 2, \\ b &\leftrightsquigarrow 3, \\ ab &\leftrightsquigarrow 4. \end{aligned}$$

Using tabular notation for permutations, we define a map f from the Klein 4-group to S_4 by

$$f(x) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \ell(x \cdot я) & \ell(x \cdot a) & \ell(x \cdot b) & \ell(x \cdot ab) \end{pmatrix}, \tag{8.1}$$

where $\ell(y)$ is the label that corresponds to y.

This notation affords us a powerful means of expression, but can be hard to read. Suppose f maps an element x of the Klein 4-group to the permutation $\sigma = (1\ 2)(3\ 4)$ of S_4 Any permutation of S_4 is a one-to-one function on a list of 4 elements, say $(1, 2, 3, 4)$. By definition, $\sigma(2) = 1$. Since $\sigma = f(x)$, we can likewise write, $(f(x))(2) = 1$. This double-evaluation can be hard to look at; it does not say "$f(x)$ times 2," but rather, "$f(x)$ of 2"? To avoid confusion, we adopt the following notation to emphasize that $f(x)$ is a permutation, and thus a function:

$$f(x) = f_x.$$

It's much easier to look at $f_x(2)$ and understand we want $f_x(2) = \sigma(1)$.

Let's compute f_a:

$$f_a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \ell(a \cdot я) & \ell(a \cdot a) & \ell(a \cdot b) & \ell(a \cdot ab) \end{pmatrix}.$$

The first entry has the value $\ell(a \cdot e) = \ell(a) = 2$, telling us that

$$f_a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & \ell(a \cdot a) & \ell(a \cdot b) & \ell(a \cdot ab) \end{pmatrix}.$$

The next entry has the value $\ell(a \cdot a) = \ell(a^2) = \ell(я) = 1$, telling us that

$$f_a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & \ell(a \cdot b) & \ell(a \cdot ab) \end{pmatrix}.$$

The third entry has the value $\ell(a \cdot b) = \ell(ab) = 4$, telling us that

$$f_a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & \ell(a \cdot ab) \end{pmatrix}.$$

The final entry has the value $\ell(a \cdot ab) = \ell(a^2 b) = \ell(b) = 3$, telling us that

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

So applying the formula in equation (8.1) definitely gives us a permutation.

Look closely. We could have filled out the bottom row of the permutation by looking above at the Klein 4-group's Cayley table, locating the row for the multiples of a (the second row of the multiplication table), and filling in the labels for the entries in that row! After all,

the row corresponding to a is *precisely*
the row of products $a \cdot y$ for all elements y of the group!

Doing this or applying equation (8.1) to the other elements of the Klein 4-group tells us

$$\begin{aligned} f_я &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = (1) \\ f_b &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \\ f_{ab} &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix}. \end{aligned}$$

The result is a subset of S_4; or, in cycle notation,

$$\begin{aligned} W &= \{f_я, f_a, f_b, f_{ab}\} \\ &= \{(1), (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}. \end{aligned}$$

Verifying that W is a group, and therefore a subgroup of S_4, is straightforward; you will do so in the homework. In fact, it is a consequence of the fact that f is a homomorphism. Strictly speaking, f is really an isomorphism. Inspection shows that f is one-to-one and onto; the hard part is the homomorphism property. We will use a little cleverness for this. Let x, y in the Klein 4-group.

- Recall that f_x, f_y, and f_{xy} are permutations, and by definition one-to-one, onto functions on a list of four elements.
- Notice that ℓ is also a one-to-one function, and it has an inverse. Just as $\ell(z)$ is the label of z, $\ell^{-1}(m)$ is the group element labeled by the number m. For instance, $\ell^{-1}(3) = b$.
- Since f_x is a permutation of a list of four elements, we can look at $f_x(m)$ as the position where f_x moves the list element in the mth position.
- By definition, f_x moves m to $\ell(z)$ where z is the product of x and the element in the mth position. Written differently, $z = x \cdot \ell^{-1}(m)$, so

$$f_x(m) = \ell\left(x\ell^{-1}(m)\right). \tag{8.2}$$

  Similar statements hold for f_y and f_{xy}.

- Applying these facts, we observe that

$$\begin{aligned}
\left(f_x \circ f_y\right)(m) &= f_x\left(f_y(m)\right) && \text{(def. of comp.)}\\
&= f_x\left(\ell\left(y \cdot \ell^{-1}(m)\right)\right) && \text{(def. of } f_y\text{)}\\
&= \ell\left(x \cdot \ell^{-1}\left(\ell\left(y \cdot \ell^{-1}(m)\right)\right)\right) && \text{(def. of } f_x\text{)}\\
&= \ell\left(x \cdot \left(y \cdot \ell^{-1}(m)\right)\right) && (\ell^{-1}, \ell \text{ inverses})\\
&= \ell\left(xy \cdot \ell^{-1}(m)\right) && \text{(assoc. prop.)}\\
&= f_{xy}(m). && \text{(def. of } f_{xy}\text{)}
\end{aligned}$$

- Since m was arbitrary in $\{1, 2, 3, 4\}$, f_{xy} and $f_x \circ f_y$ are identical functions.
- Since $f_x f_y = f_x \circ f_y$, we have $f_{xy} = f_x f_y$.
- Since x, y were arbitrary in the Klein 4-group, this holds for the entire group.

We conclude that f is a homomorphism; since it is one-to-one and onto, f is an isomorphism.

You should read through Example 8·22 carefully two or three times, and make sure you understand it, since in the homework you will construct a similar isomorphism for a different group, and also because we do the same thing now in the proof of Cayley's Theorem.

Proof of Cayley's Theorem. Let G be a finite group of n elements. Label the elements in any order $G = \{g_1, g_2, \ldots, g_n\}$ and denote $\ell(g_i) = i$. Define a relation

$$f : G \to S_n \qquad \text{by} \qquad f(g) = \begin{pmatrix} 1 & 2 & \cdots & n \\ \ell(g \cdot g_1) & \ell(g \cdot g_2) & \cdots & \ell(g \cdot g_n) \end{pmatrix}.$$

By definition, this assigns to each $g \in G$ the permutation whose second row of the tabular notation contains, in order, the labels for each entry in the row of the Cayley table corresponding to g. By this fact, we know that f is one-to-one and onto (see also Question 2.40 on page 58). The proof that f is a homomorphism is identical to the proof for Example 8·22: nothing in that argument required x, y, or z to be elements of the Klein 4-group; the proof was for a general group! Hence f is an isomorphism, and $G \cong f(G) < S_n$. □

What's so remarkable about this result? One way of looking at it is the following: since every finite group is isomorphic to a subgroup of a group of permutations, *you can learn everything you need to know about finite groups by studying the groups of permutations!*

In theory, I could go back and rewrite these notes, introducing the reader first to lists, then to permutations, then to S_2, to S_3, to the subgroups of S_4 that correspond to the cyclic group of order 4 and the Klein 4-group, and so forth, making no reference to these other groups, nor to the dihedral group, nor to any other finite group that we have studied. But it is more natural to think in terms other than permutations (geometry for D_n is helpful); and it can be tedious to work only with permutations. While Cayley's Theorem has its uses, it does not suggest that we should always consider groups of permutations in place of the more natural representations.

Question 8·23 .
For this problem, you may need to review the group D_4 of Question 3.118. In Example 8·22 we found W, a subgroup of S_4 that is isomorphic to the Klein 4-group. It turns out that W maps to a subgroup V of D_4, as well. Draw the geometric representations for each element of V, using a square and writing labels in the appropriate places, as we did in Figure 3·6 on page 107.

Question 8·24 .
Apply Cayley's Theorem to find a subgroup of S_4 that is isomorphic to $\mathbb{Z}_4$. Write the permutations in both tabular and cycle notations.

Question 8·25 .
The subgroup of S_4 that you identified in Question 8.24 maps to a subgroup of D_4, as well. Draw the geometric representations for each element of this subgroup, using a square with labeled vertices, and arcs to show where the vertices move.

Question 8·26 .
Since S_3 has six elements, we know it is isomorphic to a subgroup of S_6. In fact, it can be isomorphic to more than one subgroup; Cayley's Theorem tells us only that it is isomorphic to *at least* one. Identify a subgroup A of S_6 such that $S_3 \cong A$, yet A is *not* the image of the isomorphism used in the proof of Cayley's Theorem.

8·3 Alternating groups

The *alternating groups* are a special subgroup of the permutations that play an important role in upcoming topics. We define them using a property of a permutation that does not change regardless of how we can write it. A property like this is called **invariant**.

Transpositions

The particular invariant we consider depends on the shortest non-trivial cycles.

Definition 8·27. Let $n \in \mathbb{N}^+$. An **n-cycle** is a permutation that can be written as one cycle with n entries. A **transposition** is a 2-cycle.

Example 8·28. The permutation $(1\ 2\ 3) \in S_3$ is a 3-cycle. The permutation $(2\ 3) \in S_3$ is a transposition. The permutation $(1\ 3)\,(2\ 4) \in S_4$ cannot be written as only one n-cycle for any $n \in \mathbb{N}^+$: it is the composition of two disjoint transpositions.

Fact 8·29. *Any transposition is its own inverse.*

Why? You do it! See Question 8.30 □

Question 8·30 .
Show that:

(a) the inverse of any transposition is a transposition; and

(b) if we can write the permutation π as $\pi = \tau_1\tau_2\cdots\tau_k$, where each τ_i is a transposition, then $\pi^{-1} = \tau_k\tau_{k-1}\cdots\tau_1$.

Any permutation can be written with many different numbers of cycles; after all, any one-cycle is the identity:

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

A neat trick allows us to write every permutation as a composition of transpositions.

Example 8·31. Verify that

- $(1\ 2\ 3) = (1\ 3)\,(1\ 2)$;
- $(1\ 4\ 8\ 2\ 3) = (1\ 3)\,(1\ 2)\,(1\ 8)\,(1\ 4)$; and
- $(1) = (1\ 2)\,(1\ 2)$.

Do you see the relationship between the n-cycle and the corresponding transpositions?

Lemma 8·32. *Any permutation can be written as a composition of transpositions.*

Proof. You do it! See Question 8.33. □

Question 8·33 .
Show that any permutation can be written as a product of transpositions.

Remark 8·34. Given an expression of σ as a product of transpositions, say $\sigma = \tau_1\cdots\tau_n$, it is clear from Fact 8·29 that we can write $\sigma^{-1} = \tau_n\cdots\tau_1$, as an application of the associative property yields

$$\begin{aligned}(\tau_1\cdots\tau_n)\,(\tau_n\cdots\tau_1) &= (\tau_1\cdots\tau_{n-1})\,(\tau_n\tau_n)\,(\tau_{n-1}\cdots\tau_1)\\ &= (\tau_1\cdots\tau_{n-1})\,\begin{pmatrix}1\end{pmatrix}\,(\tau_{n-1}\cdots\tau_1)\\ &\vdots\\ &= (1)\,.\end{aligned}$$

At this point it is worth revisiting Example 8·31. Can we write $\begin{pmatrix}1 & 2 & 3\end{pmatrix}$ with many different numbers of *transpositions?* Yes:

$$\begin{aligned}(1\ 2\ 3) &= (1\ 3)\,(1\ 2)\\ &= (1\ 3)\,(1\ 2)\,(2\ 3)\,(2\ 3)\\ &= (1\ 3)\,(1\ 2)\,(1\ 3)\,(1\ 3)\\ &= \cdots .\end{aligned}$$

Nevertheless, there is a difference to writing $(1\ 2\ 3)$ as any number of cycles: no matter how we try, we seem able to write it only as an *even* number of transpositions. Similarly,

$$\begin{aligned}(2\ 3) &= (2\ 3)\,(2\ 3)\,(2\ 3)\\ &= (2\ 3)\,(1\ 2)\,(1\ 3)\,(1\ 3)\,(1\ 2) = \cdots.\end{aligned}$$

No matter how we try, we seem able to write it only as an *odd* number of transpositions.

Is this always the case?

Even and odd permutations

Theorem 8·35. *Let $\alpha \in S_n$.*

- *If α can be written as the composition of an even number of transpositions, then it cannot be written as the composition of an odd number of transpositions.*
- *If α can be written as the composition of an odd number of transpositions, then it cannot be written as the composition of an even number of transpositions.*

Proof. Define the polynomials

$$g = \prod_{1\leqslant i<j\leqslant n} \left(x_i - x_j\right) \qquad \text{and} \qquad g_\alpha = \prod_{1\leqslant i<j\leqslant n} \left(x_{\alpha(i)} - x_{\alpha(j)}\right).$$

The value of g_α depends on the permutation α; in particular, it depends only on what α does to each pair i and j. That makes g_α invariant under the representation of α. Its value does not depend on how we write α in terms of transpositions!

How is g_α related to g? Sometimes they agree; for example, if $\alpha = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}$ then

$$g = (x_1 - x_2)\,(x_1 - x_3)\,(x_2 - x_3)$$

and

$$\begin{aligned}g_\alpha &= (x_3 - x_1)\,(x_3 - x_2)\,(x_1 - x_2)\\ &= [(-1)\,(x_1 - x_3)]\,[(-1)\,(x_2 - x_3)]\,(x_1 - x_2)\\ &= g.\end{aligned}$$

Is it always the case that $g_\alpha = g$? Not necessarily: if $\alpha = \begin{pmatrix} 1 & 2 \end{pmatrix}$ then $g = x_1 - x_2$ and $g_\alpha = x_2 - x_1 \neq g$. In this case , $g_\alpha = -g$.

Question 8·36 .

We pause the proof a moment to ask you a question. Compute g_α for the permutations $(1\ 3)\,(2\ 4)$ and $(1\ 3\ 2\ 4)$. Use the value of g_α to determine which of the two permutations is odd, and which is even?

Proof of Theorem 8·35 (continued). We cannot guarantee $g_\alpha = g$, but can we write g_α in terms of g? Try the following. Lemma 8·32 tells us α is a composition of transpositions, so let's think about what happens when we compute g_τ for any transposition $\tau = \begin{pmatrix} i & j \end{pmatrix}$. Without loss of generality, we may assume that $i < j$. Let k be another positive integer.

- We know that $x_i - x_j$ is a factor of g. After applying τ, $x_i - x_j$ is no longer a factor of g_τ; rather, $x_j - x_i$. Observe that $x_j - x_i = -\left(x_i - x_j\right)$.
- If $i < j < k$, then $x_i - x_k$ and $x_j - x_k$ are factors of g. After applying τ, $x_j - x_k$ and $x_i - x_k$ are factors of g_τ. While the order has changed, the factors have not.
- If $k < i < j$, then $x_k - x_i$ and $x_k - x_j$ are factors of g. After applying τ, $x_k - x_j$ and $x_k - x_i$ are factors of g_τ. Again, the order has changed, but the factors have not.
- If $i < k < j$, then $x_i - x_k$ and $x_k - x_j$ are factors of g. After applying τ, $x_j - x_k$ and $x_k - x_i$ are factors of g_τ. The factors have changed, but the changes cancel:

$$\left(x_j - x_k\right)\left(x_k - x_i\right) = \left[-\left(x_k - x_j\right)\right]\left[-\left(x_i - x_k\right)\right] = \left(x_i - x_k\right)\left(x_k - x_j\right).$$

To summarize: $x_i - x_j$ is the only factor whose change of sign makes a difference in g and g_τ. We see that $g_\tau = -g$.

Excellent! We have characterized the relationship between g_α and g whenever α is a transposition! Return to the general case, where α is an arbitrary permutation. From Lemma 8·32, α is a composition of transpositions. Choose transpositions $\tau_1, \tau_2, \ldots, \tau_m$ such that $\alpha = \tau_1\tau_2\cdots\tau_m$. Using substitution and the observation we just made,

$$g_\alpha = g_{\tau_1\cdots\tau_m} = -g_{\tau_1\cdots\tau_{m-1}} = (-1)^2 g_{\tau_1\cdots\tau_{m-2}} = \cdots = (-1)^m g.$$

In short,

$$g_\alpha = (-1)^m g. \tag{8.3}$$

Recall that g_α depends only on α, *and not on its representation.* Assume α can be written as an even number of transpositions; say, $\alpha = \tau_1 \cdots \tau_{2m}$. Formula (8.3) tells us that $g_\alpha = (-1)^{2m} g = g$. If we could *also* write α as an odd number of transpositions, say, $\alpha = \mu_1 \cdots \mu_{2m+1}$, then $g_\alpha = (-1)^{2k+1} g$. Substitution gives us $(-1)^{2m} g = (-1)^{2k+1} g = \left[(-1)^2\right]^k (-g) = -g$, a contradiction. Hence, α cannot be written as an odd number of transpositions.

A similar argument shows that if α can be written as an odd number of transpositions, then it cannot be written as an even number of transpositions. Since $\alpha \in S_n$ was arbitrary, the claim holds. □

Lemma 8·32 tells us any permutation can be written as a composition of transpositions, and Theorem 8·35 tells us that for any given permutation, this number is always either an even or odd number of transpositions. The number itself is invariant, but whether it is even or odd (its **parity**) is not.

Definition 8·37. If a permutation can be written with an even number of permutations, then we say the permutation is **even**. Otherwise, we say the permutation is **odd**.

Example 8·38. The permutation $\rho = (1\ 2\ 3) \in S_3$ is even, since as we saw earlier $\rho = (1\ 3)\,(1\ 2)$. So is the permutation $\iota = (1) = (1\ 2)\,(1\ 2)$. The permutation $\varphi = (2\ 3)$ is odd.

Question 8·39 .
Recall the polynomials g and g_α defined in the proof of Theorem 8·35. The **sign function** $\operatorname{sgn}(\alpha)$ indicates the relationship,

$$g = \operatorname{sgn}(\alpha) \cdot g_\alpha.$$

Another way of saying this is that

$$\operatorname{sgn}(\alpha) = \begin{cases} 1, & \alpha \in A_n; \\ -1, & \alpha \notin A_n. \end{cases}$$

Show that for any two cycles α, β,

$$(-1)^{\operatorname{sgn}(\alpha\beta)} = (-1)^{\operatorname{sgn}(\alpha)}\,(-1)^{\operatorname{sgn}(\beta)}.$$

Explain why sgn is a homomorphism from S_n to the group under multiplication $\{\pm 1\}$. What is its kernel?

The alternating groups

The invariance of a permutation's parity allows us to identify a new kind of group.

Definition 8·40. Let $n \in \mathbb{N}^+$ and $n \geqslant 2$. Let $A_n = \{\alpha \in S_n : \alpha \text{ is even}\}$. We call A_n **the set of alternating permutations**.

Remark 8·41. While this A_3 is not the "A_3" of Example 4·134 on page 163, the isomorphism between D_3 and S_3 maps one to the other, so they are isomorphic.

Question 8·42 .
List the elements of A_2, A_3, and A_4 in cycle notation.

This set is, in fact, a subgroup of the symmetric group.

Theorem 8·43. *For all $n \geqslant 2$, $A_n < S_n$.*

We prove this two different ways.

First proof: directly. Let $n \geqslant 2$. By Lemma 4·4, the elements of A_n are associative under composition. Observe that $(1) = (1\ 2)\,(1\ 2) \in A_n$, so A_n has an identity. For any $\pi \in A_n$, write it as an even number of transpositions, say $\pi = \tau_1\tau_2\cdots\tau_{2k}$; by Question 8.30, $\tau^{-1} = \tau_{2k}\tau_{2k-1}\cdots\tau_1 \in A_n$, so A_n satisfies the inverse property. Finally, for any $\pi, \sigma \in A_n$, write them as an even number of transpositions, say $\pi = \tau_1\tau_2\cdots\tau_{2k}$ and $\sigma = \upsilon_1\upsilon_2\cdots\upsilon_{2\ell}$. Their product is $\pi\sigma = (\tau_1\cdots\tau_{2k})\,(\upsilon_1\cdots\upsilon_{2\ell})$, which is clearly an even number of transpositions, so A_n is closed under composition of function. Since it satisfies the four properties of a group, $A_n < S_n$. □

Second proof: using the Subgroup Theorem. Let $n \geqslant 2$. Observe that $(1) = (1\ 2)(1\ 2)$, so $A_n \neq \varnothing$. Let $x, y \in A_n$. By definition, we can write $x = \sigma_1 \cdots \sigma_{2m}$ and $y = \tau_1 \cdots \tau_{2n}$, where $m, n \in \mathbb{Z}$ and each σ_i or τ_j is a transposition. From Remark 8·34,

$$y^{-1} = \tau_{2n} \cdots \tau_1,$$

so

$$xy^{-1} = (\sigma_1 \cdots \sigma_{2m})(\tau_{2n} \cdots \tau_1).$$

This is a composition of $2m + 2n = 2(m + n)$ transpositions, which shows $xy^{-1} \in A_n$. By the Subgroup Theorem, $A_n < S_n$. □

How large is A_n, relative to S_n?

Theorem 8·44. *For any $n \geqslant 2$, there are half as many even permutations as there are permutations. That is, $|A_n| = |S_n|/2$.*

Proof. We show that there are two cosets of $A_n < S_n$, then apply Lagrange's Theorem.

Let $X \in S_n/A_n$. Let $\alpha \in S_n$ such that $X = \alpha A_n$. If α is an even permutation, then Lemma 4·103 on page 152 implies that $X = A_n$. Otherwise, α is odd. Let β be any other odd permutation. Write out the odd number of transpositions of α^{-1}, followed by the odd number of transpositions of β, to see that $\alpha^{-1}\beta$ is an even permutation. Hence, $\alpha^{-1}\beta \in A_n$, and by Lemma 4·103, $\alpha A_n = \beta A_n$.

We have shown that even permutations lie in one coset (A_n itself) while odd permutations lie in one different coset (we can write it as $(1\ 2)\,A_n$). From Lemma 8·32, any permutation is either even or odd. These two cosets of A_n thus partition S_n. By Lagrange's Theorem,

$$\frac{|S_n|}{|A_n|} = |S_n/A_n| = 2,$$

and a little algebra rewrites this equation as $|A_n| = |S_n|/2$. □

Corollary 8·45. *For any $n \geqslant 2$, $A_n \lhd S_n$.*

Proof. You do it! See Question 8.46. □

Question 8·46 .
Show that for any $n \geqslant 2$, $A_n \lhd S_n$.

8·4 The 15-puzzle

The 15-puzzle resembles a 4×4 square, with all the squares numbered, except one. The numbering starts in the upper left and proceeds consecutively until the lower right; the only squares that aren't in order are the last two, which are swapped:

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

The only permissible moves are those where one "slides" a square left, right, above, or below the empty square. The following moves are permissible from the starting position above:

1	2	3	4
5	6	7	8
9	10	11	12
13	15		***14***

or

1	2	3	4
5	6	7	8
9	10	11	
13	15	14	***12***

.

The following moves are *not*:

1	2	3	4
5	6	7	8
9	10		12
13	15	14	***11***

or

1	2	3	4
5	6	7	8
9	10	11	12
13	***14***	***15***	

.

(You may have played with a similar toy as a child.) The challenge is to find a way to rearrange the squares so that they are in order, like so:

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

This section shows the challenge to be impossible.

How? Since the problem is one of rearranging a list of elements, it is a problem of permutations. Every permissible move consists of transpositions $\tau = (x\ y)$ in S_{16} where:

- $x < y$;
- one of x or y is the position of the empty square in the current list; and
- legal moves imply that either
  - $y = x + 1$ and $4 \nmid x$; or
  - $y = x + 4$.

Example 8·47. The legal moves illustrated above correspond to the transpositions

- $(15\ 16)$, because square 14 was in position 15, and the empty space was in position 16: notice that $16 = 15 + 1$; and
- $(12\ 16)$, because square 12 was in position 12, and the empty space was in position 16: notice that $16 = 12 + 4$.

The illegal moves illustrated above correspond to the transpositions

- $(11\ 16)$, because square 11 was in position 11, and the empty space was in position 16: notice that $16 = 11 + 5$; and

- $(13\ 14)$, because neither 13 nor 14 contains the empty square.

Likewise $(12\ 13)$ would be an illegal move in any configuration, because it crosses rows: even though $y = 13 = 12 + 1 = x + 1, x = 12 = 3 \times 4$.

How can we use this to show that it is impossible to solve 15-puzzle? We take two steps. The first shows that *if* there is a solution, it must belong to a particular group.

Lemma 8·48. *Any solution to the 15-puzzle is a permutation $\sigma \in A_{16}$.*

Proof. Any permissible move corresponds to a transposition τ as described above. Any solution contains the empty square in the lower right hand corner. As a consequence,

- if $(x\ y)$ is a move left, then the empty square must eventually return to the rightmost row, so there must eventually be a corresponding move right, $(x'\ y')$; and
- if $(x\ y)$ is a move up, the empty square must eventually return to the bottom row, so there must eventually be a corresponding move down, $(x'\ y')$.

Thus, moves come in pairs. The upshot is that any solution to the 15-puzzle must be a permutation σ defined by an even number of transpositions. By Theorem 8·35 and Definitions 8·37 and 8·40, $\sigma \in A_{16}$. □

The second step is to explain why the initial configuration makes this impossible.

Theorem 8·49. *The 15-puzzle has no solution.*

Proof. By way of contradiction, assume that it has a solution σ. By Lemma 8·48, $\sigma \in A_{16}$. Because A_{16} is a subgroup of S_{16}, and hence a group in its own right, $\sigma^{-1} \in A_{16}$. Notice $\sigma^{-1}\sigma = \iota$, the permutation which corresponds to the configuration of the solution.

Now σ^{-1} is a permutation corresponding to the moves that change the arrangement

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

into the arrangement

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

which corresponds to $(14\ 15)$. Remember that a permutation's parity is invariant; you will show below that a permutation and its inverse have the same parity. We conclude that $\operatorname{sgn}\sigma = \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(14\ 15)$, contradicting Lemma 8·48. □

Question 8·50 .

Let $\sigma \in S_n$. How do we know that $\text{sgn}\sigma = \text{sgn}\,(\sigma^{-1})$?

As a historical note, the 15-puzzle was developed in 1878 by an American puzzle maker, who promised a $1,000 reward to the first person to solve it. Most probably, the puzzle maker knew that no one would ever solve it: if we account for inflation, the reward would correspond to $22,265 in 2008 dollars.[1]

The textbook [2] contains a more general discussions of solving puzzles of this sort using algebra.

Question 8·51 .

Determine which of these configurations, if any, is solvable by the same rules as the 15-puzzle:

1	2	3	4
5	6	7	8
9	10	12	11
13	14	15	

,

1	2	3	4
5	10	6	8
13	9	7	11
14	15	12	

,

3	6	4	7
1	2	12	8
5	15	10	14
9	13	11	

.

[1] According to the website `www.measuringworth.com/ppowerus/result.php`.

Chapter 9

Solving polynomials by radicals

In this chapter, we take a few steps into *Galois theory,* a major impetus for the development of algebra. The subject's development began with the effort to generalize the quadratic formula,

$$ax^2 + bx + c = 0 \quad \Longrightarrow \quad x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

to higher-degree polynomials. This formula requires arithmetic operations (addition, subtraction, multiplication, division) *and* one "algebraic" operation, the radical. When an equation can be solved by these five operations, we say that it can be *solved by radicals.*

Renaissance mathematicians discovered formulas that extend this elegant approach to cubic and quartic polynomials, so that in principle one can describe a "cubic formula" and a "quartic formula," though they are much more complicated than the quadratic. Quintic polynomials turned out to be more difficult — because, as Ruffini argued, Abel proved, and finally Galois elaborated, it is *impossible* to solve every quintic polynomial by radicals; not every polynomial root can be described in this way!

We explore only the theory which explains this failure. Polynomials will lie over a ground field $\mathbb{F}$ of characteristic zero: so, no zero divisors; if $na = 0$, then $n = 0$ or $a = 0$. This assumption rules out all the clockwork fields, since $a \in \mathbb{F}_{p^k}$ implies that $pa = 0$.

9·1 Radical extensions of a field

Section 3·1 showed that we could use the polynomial $x^2 + 1$ over $\mathbb{R}$ to build a new field, "$\mathbb{C}$", over which the polynomial $x^2 + 1$ has a root. This new field acts as an extension of $\mathbb{R}$ in the sense that we can find a subfield of "$\mathbb{C}$" that is isomorphic to $\mathbb{R}$. We developed this further in Sections 6·2, 6·5, and 6·6: any irreducible polynomial generates a maximal ideal, which we can use to build a new field that contains a root of the polynomial. (See in particular Theorem 6·30, Fact 6·70, and Theorem 6·72.

This section "extends" our results a bit.

Extending a field by a root

Since $\mathbb{F}$ is a subfield of the ring $\mathbb{F}[x]$, we can view it as a subfield of the field $\mathbb{E} = \mathbb{F}[x] / \langle f \rangle$. At any rate, it is certainly isomorphic to a subfield of the latter field, which has a root of

f, which means we are not unreasonable in stating that there exists a superfield of $\mathbb{F}$ that contains a root α of f.

Definition 9·1. Let f be an irreducible polynomial over a field $\mathbb{F}$, and let α be a root of f that is not in $\mathbb{F}$. We call the field $\mathbb{E} = \mathbb{F}(\alpha)$ an **algebraic extension of** $\mathbb{F}$, and say that we obtain $\mathbb{E}$ from $\mathbb{F}$ by **adjoining** α. If f is irreducible and $d = \deg f$, we say that $\mathbb{E}$ is an **extension of degree** d **(over** $\mathbb{F}$**)**. If there exists $m \in \mathbb{N}^+$ such that $\alpha^m \in \mathbb{F}$, then we say that $\mathbb{E}$ is a **radical extension of** $\mathbb{F}$. We will also relax our precision sometimes to say that a *sequence* of field extensions is an algebraic (or radical) extension if every extension in the sequence is algebraic (or radical).

This terminology allows us to give our main question more precision:

Is an algebraic extension always radical?

Were the answer "Yes," then we could always solve polynomials by radicals; we would simply construct a finite sequence of radical extensions

$$\mathbb{E}_1 = \mathbb{Q}(\alpha_1), \quad \mathbb{E}_2 = \mathbb{Q}(\alpha_1)(\alpha_2), \quad \cdots \quad \mathbb{E}_k = \mathbb{Q}(\alpha_1)\cdots(\alpha_k)$$

where each α_i has the form

$$\alpha_i = \sqrt[n]{\beta} \quad \text{for some } \beta \in \mathbb{E}_{i-1},$$

and every root of the polynomial would appear in $\mathbb{E}_k$.

The purpose of this chapter is to show that the answer is, in fact, "No."

Question 9·2 .

Let α be a root of an irreducible polynomial $f \in \mathbb{F}[x]$, with $\deg f \geqslant 2$. Explain how we know that $\mathbb{F}(\alpha)$ satisfies both the properties of a field and $\mathbb{F} \subsetneq \mathbb{F}(\alpha) \subseteq \mathbb{E}$, where $\mathbb{E}$ is any field that contains both $\mathbb{F}$ and α.

You may wonder whether the degree of an algebraic extension is well-defined; after all, α could be the root of two different irreducible polynomials of different degree. In fact, this cannot happen.

Fact 9·3. *Let $f, g \in \mathbb{F}[x]$ be two irreducible polynomials with a common root $\alpha \notin \mathbb{F}$. Then $\deg f = \deg g$.*

Why? Recall that $\mathbb{F}[x]$ is a Euclidean domain, and compute a gcd p of f and g. Since f is irreducible and $p \mid f$, p is either a unit or an associate of f. By Question 6.51 and Fact 6·55 we know that Bézout's Lemma applies, so we can find $h_1, h_2 \in \mathbb{F}[x]$ such that $p = h_1 f + h_2 g$. By substitution,

$$p(\alpha) = h_1(\alpha) f(\alpha) + h_2(\alpha) g(\alpha) = h_1(\alpha) \cdot 0 + h_2(\alpha) \cdot 0 = 0.$$

You will show in Question 9.4 that a polynomial cannot both have a root and be a unit, so p is not a unit; it must be an associate of f; say $p = af$ where $a \in \mathbb{F}[x]$ is a unit. Since g is irreducible

and $p \mid g$, by substitution $af \mid g$, so $f \mid a^{-1}g$. Recall that $\mathbb{F}[x]$ is a unique factorization domain, so the irreducible f is prime, so $f \mid a^{-1}$ or $f \mid g$. You will show in Question that only a unit can divide a unit, and an irreducible is by definition not a unit, so $f \mid g$. Since g is irreducible, f must be an associate of g. You will also show in Question 9.4 that the degree of a unit is zero, so $\deg f = \deg g$. □

Question 9·4 .

The explanation for all three properties hinted at in the proof above is nearly identical.

(a) Let $f \in \mathbb{F}[x]$. Show that if f has a root $\alpha \in \mathbb{F}$, then f cannot be a unit. *Hint:* Proceed by contradiction. Think about the value of f and its inverse at α.

(b) Let R be a ring and $a, r \in R$, where a is a unit. Show that if $r \mid a$ then r is also a unit.

(c) Let $\mathbb{F}[x]$ be the ring of polynomials over the field $\mathbb{F}$, and $a \in \mathbb{F}[x]$ a unit. Show that $\deg a = 0$.

Question 9·5 .

An alternate approach to defining the degree of an extension is as follows. Let α be the root of an irreducible polynomial f over $\mathbb{F}$.

(a) Show that $\mathbb{F}(\alpha)$ is a vector space over $\mathbb{F}$.

(b) Show that the vector space is finite dimensional. *Hint:* Use f to show that only finitely many powers of α are linearly independent.

Define $\deg \mathbb{F}(\alpha)$ to be the dimension of $\mathbb{F}(\alpha)$ over $\mathbb{F}$.

(c) Explain why the value of $\deg \mathbb{F}(\alpha)$ computed this way gives the identical result as the value indicated above.

Example 9·6. Let $f = x^5 - 2x^3 - 3x^2 + 6$. This factors over $\mathbb{Q}$ as $(x^2 - 2)(x^3 - 3)$. Both factors are irreducible over $\mathbb{Q}$. From what we wrote above, there exists a radical extension of degree 2 of $\mathbb{Q}$ that contains a root of $x^2 - 2$; call the corresponding root $\alpha = \sqrt{2}$, so that instead of writing $\mathbb{Q}(\alpha)$, we can write $\mathbb{Q}\left(\sqrt{2}\right)$, instead.

What do elements of $\mathbb{Q}(\alpha)$ "look" like? By the definition of a ring extension, we know that elements of this field have the form $a + b\sqrt{2} + c\sqrt{2}^2 + \cdots$. Now, $\sqrt{2}$ is a root of $x^2 - 2$, which means that $\sqrt{2}^2 - 2 = 0$, which we can rewrite as $\sqrt{2}^2 = 2$. Hence, we can assume that elements of $\mathbb{Q}\left[\sqrt{2}\right]$ really have the form $a + b\sqrt{2}$, since we just saw how higher powers of $\sqrt{2}$ reduce either to an element of $\mathbb{Q}$, or to a rational multiple of $\sqrt{2}$ itself.

It might not be obvious that such elements have multiplicative inverses, but they do. You can see this either by working with the isomorphic quotient field $\mathbb{Q}[x] / \langle x^2 - 2 \rangle$, or in this case solving a straightforward linear equation. For the nonzero element $a + b\sqrt{2}$ to have an inverse $c + d\sqrt{2}$, we need

$$1 = \left(a + b\sqrt{2}\right)\left(c + d\sqrt{2}\right) = (ac + 2bd) + (ad + bc)\sqrt{2}.$$

Since $1 = 1 + 0\sqrt{2}$, we know we can find an inverse if

$$ac + 2bd = 1 \quad \text{and} \quad ad + bc = 0.$$

Since $a + b\sqrt{2}$ is nonzero, we can assume that $a \neq 0$ or $b \neq 0$. If $a \neq 0$, then we can solve the two equations to see that

$$c = \frac{1 - 2bd}{a} \quad \text{and} \quad d = -\frac{bc}{a}.$$

Notice that this solution satisfies $c, d \in \mathbb{Q}$, since the rationals are a field. If $a = 0$, on the other hand, those equations simplify to

$$2bd = 1 \quad \text{and} \quad bc = 0,$$

so that $d = 1/(2b)$ and $c = 0$. To make sure you understand that, use this principle to find the inverses of $1 - 2\sqrt{2}$ and $3\sqrt{2}$.

Does $x^3 - 3$ factor over this extension field? If so, then it has at least one linear factor, $x - \beta$. This makes β a root of $x^3 - 3$, so we can resolve the question by asking, does $x^3 - 3$ have a root in $\mathbb{Q}\left(\sqrt{2}\right)$? If so, it has the form $x = a + b\sqrt{2}$, and we can rewrite the polynomial as

$$\begin{aligned} 0 = x^3 - 3 &= \left(a + b\sqrt{2}\right)^3 - 3 \\ &= a^3 + 3a^2b\sqrt{2} + 6ab^2 + 2b^3\sqrt{2} - 3 \\ &= \left(a^3 + 6ab^2 - 3\right) + \left(3a^2b + 2b^3\right)\sqrt{2}. \end{aligned}$$

In other words,

$$\sqrt{2} = \frac{-a^3 - 6ab^2 + 3}{3a^2b + 2b^3}.$$

Remember that $a, b \in \mathbb{Q}$, so addition, subtraction, and multiplication, are closed, and division is closed so long as the divisor is nonzero. If the divisor in this expression is in fact nonzero — that is, $3a^2b + 2b^3 \neq 0$ — then the equation above tells us that $\sqrt{2} \in \mathbb{Q}$. **We know that this is false!** The divisor must, therefore, be zero, which means that

$$b\left(3a^2 + 2b^2\right) = 3a^2b + 2b^3 = 0 \quad \Longrightarrow \quad b = 0 \text{ or } 3a^2 + 2b^2 = 0.$$

If $b = 0$, then $x \in \mathbb{Q}$. That is, $x^3 - 3$ has a rational root. **We know that this is false!** If $b \neq 0$, on the other hand, then $3a^2 + 2b^2 = 0$, which we can rewrite as $a/b = \sqrt{-2/3}$. Since $a, b \in \mathbb{Q}$, we conclude that $\sqrt{-2/3} \in \mathbb{Q}$. Again, **we know that this is false!** All the possibilities lead us to a contradiction, so we conclude that $x^3 - 3$ does not factor over the extension field $\mathbb{Q}\left(\sqrt{2}\right)$.

As before, we can extend $\mathbb{Q}\left(\sqrt{2}\right)$ by a root of $x^3 - 2$; call it $\sqrt[3]{3}$. We now have the extension field $\mathbb{E} = \mathbb{Q}\left(\sqrt{2}\right)\left(\sqrt[3]{3}\right)$. Have we found all the roots f now? For the factor $x^2 - 2$, we certainly have, since $x^2 - 2 = \left(x - \sqrt{2}\right)\left(x + \sqrt{2}\right)$. For the other factor, we are not quite done; we have,

$$x^3 - 3 = \left(x - \sqrt[3]{3}\right)\left(x^2 + x\sqrt[3]{3} + \sqrt[3]{9}\right),$$

and this latter polynomial does not factor. To see why not, let's use the quadratic equation to find what the roots *should* be:

$$\begin{aligned} x^2 + x\sqrt[3]{3} + \sqrt[3]{9} = 0 \qquad & \\ x &= \frac{-\sqrt[3]{3} \pm \sqrt{\left(\sqrt[3]{3}\right)^2 - 4\sqrt[3]{9}}}{2} \\ &= \frac{-\sqrt[3]{3} \pm \sqrt{-3\sqrt[3]{9}}}{2} \\ &= \frac{-\sqrt[3]{3} \pm i\sqrt{3}\sqrt[3]{3}}{2} \\ &= -\sqrt[3]{3}\left(\frac{1}{2} \pm i\frac{\sqrt{3}}{2}\right). \end{aligned}$$

So we are still missing the cube roots of unity.

Question 9·7 .
Find the smallest extension field of $\mathbb{Q}$ where $f(x) = x^7 - 2x^4 - x^3 + 2$ factors completely. *Hint:* f is not irreducible over $\mathbb{Q}$, so try to factor it completely over $\mathbb{Q}$ before working on extensions.

In the example, we construct $\mathbb{Q}\left(\sqrt{2}\right)$, whose degree over $\mathbb{Q}$ is 2, and $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$, whose degree over $\mathbb{Q}\left(\sqrt{2}\right)$ is 3. How should we determine the degree of $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$ over $\mathbb{Q}$? You might think to add the degrees, but then you would lose an important relationship between the degree of an extension and the dimension of the extension as a vector space over the base field. Elements of $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$ can be written as

$$a + b\sqrt{2} + c\sqrt[3]{3} + d\sqrt[3]{9} + e\sqrt{2}\sqrt[3]{3} + f\sqrt{2}\sqrt[3]{9};$$

each term is linearly independent of the others, so that $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$ is a vector space of dimension 6 over $\mathbb{Q}$. In the same way, $\mathbb{Q}\left(\sqrt{2}\right)$ was a vector space of dimension 2 over $\mathbb{Q}$, and $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$ was a vector space of dimension 3 over $\mathbb{Q}\left(\sqrt{2}\right)$. Given that link, it makes better sense to define the degree of $\mathbb{Q}\left(\sqrt{2}, \sqrt[3]{3}\right)$ over $\mathbb{Q}$ as 6.

Definition 9·8. Let $\mathbb{F}$ be a field, and

$$\mathbb{F} = \mathbb{E}_0 \subsetneq \mathbb{E}_1 \subsetneq \mathbb{E}_2 \subsetneq \cdots \subsetneq \mathbb{E}_m$$

a chain of algebraic extensions. Denote the **degree** of $\mathbb{E}_i$ over $\mathbb{E}_{i-1}$ as $[\mathbb{E}_i : \mathbb{E}_{i-1}]$; we define the **degree** of $\mathbb{E}_m$ over $\mathbb{F}$ as

$$[\mathbb{E}_m : \mathbb{E}_{m-1}]\,[\mathbb{E}_{m-1} : \mathbb{E}_{m-2}] \cdots [\mathbb{E}_2 : \mathbb{E}_1]\,[\mathbb{E}_1 : \mathbb{E}_0].$$

Question 9·9 .
Show that Definition 9·8 is well-defined; that is, if there is more than one way to fill in the dots of the chain of algebraic extensions $\mathbb{E}_0 \subsetneq \cdots \subsetneq \mathbb{E}_m$ (for instance,

$$\mathbb{Q} \subsetneq \mathbb{Q}\left(\sqrt{2}\right) \subsetneq \mathbb{Q}\left(\sqrt{2}, \sqrt{5}\right) \subsetneq \mathbb{Q}\left(\sqrt{2}, \sqrt{5}, \sqrt{7}\right)$$

or

$$\mathbb{Q} \subsetneq \mathbb{Q}\left(\sqrt{5}\right) \subsetneq \mathbb{Q}\left(\sqrt{5}, \sqrt{7}\right) \subsetneq \mathbb{Q}\left(\sqrt{2}, \sqrt{5}, \sqrt{7}\right),$$

then we do not obtain different degrees by considering different chains. *Hint:* Use the alternate definition of the degree of an extension, given in Question 9.5.

Remark. The usual study of Galois theory considers field extensions in a more general case; that is, not just as the roots of irreducible polynomials. Our restriction tries to keep the affair in a very concrete setting; this allows us some latitude with the degree of an extension that the more general case does not enjoy.

Complex roots

As a cube root of unity popped up in the example above, you would be wise to conjecture that the roots of unity play a fundamental role here. In fact, we can obtain radical roots by adjoining both a "principal" root, and a sensible "root of unity."

Recall from Theorems 3·26 and 3·28 that $1, \omega, \omega^2, \ldots, \omega^{n-1}$ are all n-th roots of unity, where ω has the form

$$\omega = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right),$$

and from Lemma 3·27 we see further that

$$\omega^m = \cos\left(\frac{2\pi m}{n}\right) + i \sin\left(\frac{2\pi m}{n}\right).$$

This pattern extends beyond the roots of unity.

Theorem 9·10. *If α is a root of an irreducible polynomial $x^n - a \in \mathbb{Q}[x]$, then all other roots of $x^n - a$ have the form $\alpha \cdot \omega^m$, where ω is a primitive n-th root of unity and $m \in \{1, \ldots, n-1\}$.*

Proof. Assume that α is a root of an irreducible polynomial $x^n - a \in \mathbb{Q}[x]$. By substitution and definition of the primitive n-th root,

$$(\alpha\omega^m)^n - a = \alpha^n (\omega^n)^m - a = \alpha^n \cdot 1^m - a.$$

By hypothesis, $\alpha^n - a = 0$, so

$$(\alpha\omega^m)^n - a = 0.$$

By definition, $\alpha\omega^m$ is a root of $x^n - a$.

Very well, but why must this form characterize *all* the roots of $x^n - a$? Using the Factor Theorem, we see that $x^n - a$ can have no more than n roots, and we just found n such distinct roots,

$$\alpha,\ \alpha\omega,\ \alpha\omega^2,\ \ldots,\ \alpha\omega^{n-1}.$$

□

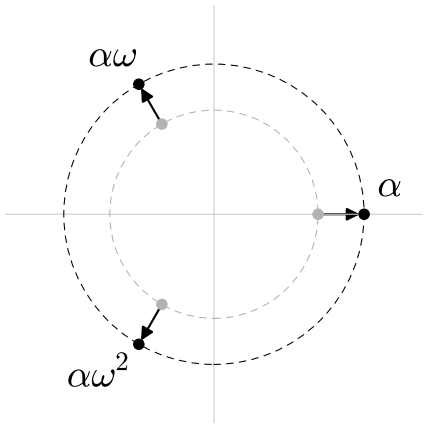


Figure 9·1: The roots of $x^3 - 3$, obtained using one root and the cube roots of unity.

Example 9·11. Returning to the question of the roots of $x^3 - 3$, we defined one root to be $\sqrt[3]{3}$. The other roots are, therefore,

$$\sqrt[3]{3}\left[\cos\left(\frac{2\pi}{3}\right) + i\sin\left(\frac{2\pi}{3}\right)\right] \quad \text{and} \quad \sqrt[3]{3}\left[\cos\left(\frac{4\pi}{3}\right) + i\sin\left(\frac{4\pi}{3}\right)\right]$$

or, after evaluating these trigonometric functions,

$$\sqrt[3]{3}\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) \quad \text{and} \quad \sqrt[3]{3}\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right).$$

If you look back at the result of the quadratic equation, you will find that this does indeed describe the missing roots. Figure 9·1 shows how the primitive cube roots of unity "scale out" to give us the roots of $x^3 - 3$.

Thus, the extension of $\mathbb{Q}$ to a field containing all the roots of $x^5 - 2x^3 - 3x^2 + 6$ is the field $\mathbb{Q}\left(\sqrt{2}\right)\left(\sqrt[3]{3}\right)(\omega)$, where ω is any primitive cube root of unity.

(You may wonder: have we actually captured all the roots? After all, we didn't extend by a primitive *square* root of unity. This is because there is only one primitive square root of unity, -1, and it appears in $\mathbb{Q}$ already.)

At this point, we encounter a problem: what if we had proceeded in a different order? In the example given, we adjoined $\sqrt{2}$ first, then $\sqrt[3]{3}$, and finally ω. Suppose we were to adjoin them in a different order — say, $\sqrt[3]{3}$ first, then ω, and finally $\sqrt{2}$? How would that work out?

As long as we adjoin all the roots, we arrive at the same field. For this reason, we write $\mathbb{Q}(\alpha_1, \ldots, \alpha_n) = \mathbb{Q}(\alpha_1)(\alpha_2)\cdots(\alpha_n)$ as a shorthand. However, Theorem 9·10 implies that $\mathbb{Q}\left(\sqrt[3]{3}\right)$ by itself does not contain all the roots of $x^3 - 3$; it contains only $\sqrt[3]{3}$. We could adjoin the other roots, $\mathbb{Q}\left(\sqrt[3]{3}, \omega\sqrt[3]{3}, \omega^2\sqrt[3]{3}\right)$, but there is another, simpler way. To obtain all the roots of $x^3 - 3$, we can first adjoin a primitive cube root of unity, then $\sqrt[3]{3}$. Typically, we adjoin a primitive cube root of unity *first,* obtaining $\mathbb{Q}(\omega)\left(\sqrt[3]{3}\right)$, or $\mathbb{Q}\left(\omega, \sqrt[3]{3}\right)$. This certainly gives us $\sqrt[3]{3}$, $\omega\sqrt[3]{3}$, and $\omega^2\sqrt[3]{3}$.

You might wonder if this doesn't give us *too much.* After all, $\omega \in \mathbb{Q}\left(\omega, \sqrt[3]{3}\right)$, but it isn't obviously an element of $\mathbb{Q}\left(\sqrt[3]{3}, \omega\sqrt[3]{3}, \omega^2\sqrt[3]{3}\right)$. You will show in the exercises that, in fact,

$\mathbb{Q}\left(\omega, \sqrt[3]{3}\right) = \mathbb{Q}\left(\sqrt[3]{3}, \omega\sqrt[3]{3}, \omega^2\sqrt[3]{3}\right)$, and the more general notion also holds: if we adjoin a primitive n-th root of unity ω and $\sqrt[n]{a}$, we end up with exactly the field $\mathbb{Q}\left(\sqrt[n]{a}, \omega\sqrt[n]{a}, \ldots, \omega^{n-1}\sqrt[n]{a}\right)$ — nothing more, nothing less.

Question 9·12 .

Suppose that $\alpha^n \in \mathbb{Q}$, $\alpha^i \notin \mathbb{Q}$ for $1 \leqslant i < n$, and ω is a primitive n-th root of unity. Show that $\mathbb{Q}\left(\alpha, \omega\alpha, \ldots, \omega^{n-1}\alpha\right) = \mathbb{Q}\left(\omega, \alpha\right)$.

9·2 The symmetries of the roots of a polynomial

Let $\mathbb{F}$ be a field, and $f \in \mathbb{F}[x]$ of degree 2. We can show by the Factor Theorem that f has at most 2 roots in $\mathbb{F}$. (See Exercise 6.56.) Suppose that f does have 2 roots in $\mathbb{F}$; we can then write $f(x) = (x-\alpha_1)(x-\alpha_2)$. If we expand this product, we obtain $f(x) = x^2 - (\alpha_1+\alpha_2)x + \alpha_1\alpha_2$. Likewise, if f is of degree 3, it can have at most 3 roots in $\mathbb{F}$; we can write $f(x) = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$, which expands to

$$f(x) = x^3 - (\alpha_1+\alpha_2+\alpha_3)x + (\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)x - \alpha_1\alpha_2\alpha_3.$$

In general, if f is of degree n and has n roots in $\mathbb{F}$, we can write

$$f(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n),$$

which expands to

$$f(x) \quad = \quad x^n \; + \; \left(\sum_{i=1}^{n}\alpha_i\right)x^{n-1} \; + \; \left(\sum_{i<j}\alpha_i\alpha_j\right)x^{n-2} \; + \; \cdots \; + \; \alpha_1\alpha_2\cdots\alpha_n.$$

Every coefficient is a sum of terms that are products of roots. *Permuting the roots does not change the sum.*

Example 9·13. Look at the coefficient of x in the cubic polynomial above. One of the terms is $\alpha_1\alpha_3$. If we permute by (123), α_1 changes to α_2 and α_3 changes α_1. The result is $\alpha_2\alpha_1 = \alpha_1\alpha_2$, which also appears in that coefficient. Another term is $\alpha_2\alpha_3$; applying the same permutation gives us $\alpha_3\alpha_1 = \alpha_1\alpha_3$.

This gives rise to a special class of polynomial.

Definition 9·14. Let R be a ring and $f \in R[x_1, \ldots, x_n]$. For any $\sigma \in S_n$, write σf for the polynomial $g \in R[x_1, \ldots, x_n]$ obtained by replacing x_i by $x_{\sigma(i)}$. We say that f is a **symmetric polynomial** if $f = \sigma f$ for all $\sigma \in S_n$.

Example 9·15. Let $f(x) = x_1x_2 - x_1x_3$. This is not a symmetric polynomial, since for $\sigma = (1\ 3)$ we obtain

$$\sigma f = x_2x_3 - x_1x_3 \neq f.$$

Example 9·16. On the other hand, if $f(x) = x_1x_2x_3 + x_2x_3x_4 + x_1x_3x_4 + x_1x_2x_4$, every $\sigma \in S_4$ satisfies $\sigma f = f$. For example, if $\sigma = (1\ 4)$,

$$\sigma f = x_2x_3x_4 + x_1x_2x_3 + x_1x_3x_4 + x_1x_2x_4 = f.$$

Here, f is symmetric.

Question 9·17 .

The polynomial $f(x) = x^4 - 7x^2 + 10$ factors over $\mathbb{Q}$ as $(x^2 - 2)(x^2 - 5)$, and over $\mathbb{Q}\left(\sqrt{2}, \sqrt{5}\right)$ as $\left(x \pm \sqrt{2}\right)\left(x \pm \sqrt{5}\right)$.

(a) Compute the symmetric polynomials of the coefficients of a generic fourth-degree polynomial.

(b) Substitute the roots of f into the symmetric polynomials. Show that they simplify to the coefficients of f.

Theorem 9·18. *Let $f \in \mathbb{F}[x]$. The coefficient of any term of f is a symmetric polynomial of the roots of f. In particular, if $\deg f = n$, then the coefficient of x^i is the sum of all squarefree products of exactly $n - i$ roots.*

Proof. We proceed by induction on $n = \deg f$. Write $\alpha_1, \ldots, \alpha_n$ for the roots of f.

Inductive base: If $n = 2$, then $f(x) = (x - \alpha_1)(x - \alpha_2) = x^2 - (\alpha_1 + \alpha_2)x + \alpha_1\alpha_2$. The coefficient of x^2 is the sum of all products of $2 - 2 = 0$ roots; the coefficient of x is the sum of all squarefree products of $2 - 1 = 1$ roots, and the coefficient of x^0 is the sum of all squarefree products of $2 - 0 = 2$ roots.

Inductive hypothesis: Assume that the coefficients of the terms of any $(n-1)$-th degree polynomial have the form specified.

Inductive step: Let $g \in \mathbb{F}(\alpha_1) \cdots (\alpha_{n-1})$ such that $f(x) = g(x)(x - \alpha_n)$. Since $\deg g = n - 1$, the inductive hypothesis tells us that its terms are symmetric polynomials of its roots, in precisely the form specified. With that in mind, write

$$g(x) = x^{n-1} + \beta_{n-2}x^{n-2} + \cdots + \beta_0$$

where β_i is the sum of all squarefree products of $(n-1) - i$ roots $\alpha_1, \ldots, \alpha_{n-1}$. Expand the product $f(x) = g(x)(x - \alpha_n)$ to see that

$$\begin{aligned} f(x) &= \left(x^n + \beta_{n-2}x^{n-1} + \cdots + \beta_0 x\right) + \left(\alpha_n x^{n-1} + \alpha_n\beta_{n-2}x^{n-2} + \cdots + \alpha_n\beta_0\right) \\ &= x^n + \left(\beta_{n-2} + \alpha_n\right)x^{n-1} + \left(\beta_{n-3} + \alpha_n\beta_{n-2}x^{n-2}\right) + \cdots + \alpha_n\beta_0. \end{aligned}$$

- Since β_{n-2} is the sum of all squarefree products of $(n-1) - (n-2) = 1$ roots $\alpha_1, \ldots, \alpha_{n-1}$, we indeed have $\beta_{n-2} + \alpha_n$ as the sum of all products of 1 root in $\alpha_1, \ldots, \alpha_n$.
- Let $i \in \{2, 3, \ldots, n-1\}$. Since β_{n-i} is the sum of all squarefree products of $(n-1) - (n-i) = i - 1$ roots $\alpha_1, \ldots, \alpha_{n-1}$, we see that $\alpha_n\beta_{n-i}$ is the sum of all squarefree products of i roots $\alpha_1, \ldots, \alpha_n$ that contain precisely one α_n. Since β_{n-i-1} is the sum of all squarefree products of $(n-1) - (n-i-1) = i$ roots $\alpha_1, \ldots, \alpha_{n-1}$, and $\alpha_n\beta_{n-i}$ is the sum of all squarefree products of i roots $\alpha_1, \ldots, \alpha_n$ that contain precisely one α_n, we indeed have $\beta_{n-i-1} + \alpha_n\beta_{n-i}$ as the sum of all squarefree products of i roots in $\alpha_1, \ldots, \alpha_n$.

- Since β_0 is the sum of all squarefree products of $(n-1)-0 = n-1$ roots $\alpha_1, \ldots, \alpha_n$, we have $\beta_0 = \alpha_1 \cdots \alpha_{n-1}$. By substitution, $\alpha_n \beta_0 = \alpha_1 \cdots \alpha_n$. This is precisely the sum of all squarefree products of $n - 0 = n$ roots $\alpha_1, \ldots, \alpha_n$.

□

Another way to read Theorem 9·18 is that we can study the roots of polynomials by looking at permutations of them. In particular, the functions defined on $\mathbb{E} = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ that permute the roots but leave elements of $\mathbb{Q}$ fixed must be of paramount importance. We are especially interested in those functions that are isomorphisms on $\mathbb{E}$ itself; in other words, automorphisms on $\mathbb{E}$.

Example 9·19. Let $f = x^2 + 1$; we have $f \in \mathbb{Q}[x]$, but its roots are not in $\mathbb{Q}$. Let i be a root of f, and let $\mathbb{E} = \mathbb{Q}(i)$. By Exercise 9.5, $\mathbb{E}$ is a vector space over $\mathbb{Q}$, with basis $\{1, i\}$, so every element of $\mathbb{E}$ can be written as $a + bi$ where $a, b \in \mathbb{Q}$.

We are interested in the automorphisms of $\mathbb{E}$ that fix $\mathbb{Q}$. Let φ be any such automorphism; by definition, $\varphi(q) = q$ for any $q \in \mathbb{Q}$, while for any $w, z \in \mathbb{E} \backslash \mathbb{Q}$, $\varphi(w)\varphi(z) = \varphi(wz)$.

Let $z \in \mathbb{E}$, and choose $a, b \in \mathbb{Q}$ such that $z = a + bi$. The properties of a ring homomorphism imply that

$$\varphi(z) = \varphi(a + bi) = \varphi(a) + \varphi(bi) = \varphi(a) + \varphi(b)\varphi(i).$$

As stated, φ fixes $\mathbb{Q}$, so $\varphi(a) = a$ and $\varphi(b) = b$. By substitution,

$$\varphi(z) = a + b\varphi(i).$$

In other words, φ is determined completely by what it does to i.

What are the possible destinations of $\varphi(i)$? First notice that φ cannot map i to a rational number q, because φ is an automorphism, hence one-to-one, and φ fixes $\mathbb{Q}$, so $\varphi(q) = q$: we would have $\varphi(i) = \varphi(q)$, but $i \neq q$. The only thing we can choose for $\varphi(i)$ to satisfy this requirement is some $w = c + di \in \mathbb{E}$ where $c, d \in \mathbb{Q}$ and $d \neq 0$. On the other hand, the homomorphism property means that we *must* have

$$w^2 = \varphi(i)^2 = \varphi\left(i^2\right) = \varphi(-1) = -1.$$

(Again, φ fixes $\mathbb{Q}$, and $-1 \in \mathbb{Q}$.) That forces $w = \pm i$.

Can we use both? If $w = i$, then φ is the identity map, since $\varphi(z) = a + bi = z$. That certainly works. If $w = -i$, then $\varphi(z) = a - bi$, the **conjugation map**. You will show in the exercises that this is indeed a ring automorphism.

Question 9·20 .

Show that the conjugation map $\varphi(a + bi) = a - bi$ is a ring isomorphism in $\mathbb{C}$.

Question 9·21 .
Find all the automorphisms on $\mathbb{E}$ that fix $\mathbb{F}$.

(a) $\mathbb{F} = \mathbb{Q}; \mathbb{E} = \mathbb{Q}\left(\sqrt{2}\right)$

(b) $\mathbb{F} = \mathbb{Q}; \mathbb{E} = \mathbb{Q}\left(2\right)$

(c) $\mathbb{F} = \mathbb{Q}; \mathbb{E} = \mathbb{Q}\left(\sqrt{2}, \sqrt{5}\right)$

(d) $\mathbb{F} = \mathbb{Q}; \mathbb{E} = \mathbb{Q}\left(i, \sqrt{2}\right)$

9·3 Galois groups

In the previous section, we observed that permuting a polynomials' roots does not change its coefficients, and that suggests a connection with permutations.

Isomorphisms of field extensions that permute the roots

Let's look, therefore, at formulating functions that combine these two. Let $f \in \mathbb{Q}[x]$ have degree n, and let $\mathbb{E}$ be a field that extends $\mathbb{Q}$ by all the roots of f. For any permutation $\sigma \in S_n$, define a function $\varphi : \mathbb{E} \longrightarrow \mathbb{E}$ such that φ acts as the identity on elements of $\mathbb{Q}$ (we say that φ **fixes** $\mathbb{Q}$), but permutes the roots of f. We place one condition on φ: it must be an isomorphism; after all, *we want the same field structure*. This imposes a condition on σ, as well.

Example 9·22. In the previous section, we used $f(x) = x^5 - 2x^3 - 3x^2 + 6$. That gave us $\mathbb{E} = \mathbb{Q}\left(\sqrt{2}, \omega, \sqrt[3]{3}\right)$, where ω is a primitive cube root of unity. The roots of f are $\alpha_1 = \sqrt{2}$, $\alpha_2 = -\sqrt{2}$, $\alpha_3 = \sqrt[3]{3}$, $\alpha_4 = \omega\sqrt[3]{3}$, and $\alpha_5 = \omega^2\sqrt[3]{3}$. Which permutations of the roots will we allow?

One example to try is $(1\ 2)$; this would switch $\sqrt{2}$ and $-\sqrt{2}$ in any element of $\mathbb{E}$. Does it extend to an isomorphism? Any expression that does not contain $\pm\sqrt{2}$ is left untouched, so let's look at expressions that contain $\pm\sqrt{2}$. As a simple case, consider two elements of the elements of $\mathbb{Q}\left(\sqrt{2}\right) \subsetneq \mathbb{E}$. By Exercise 9.5, we can write any $x, y \in \mathbb{Q}\left(\sqrt{2}\right)$ as $x = a + b\sqrt{2}$ and $y = c + d\sqrt{2}$ for some $a, b, c, d \in \mathbb{Q}$. For addition, we have

$$\begin{aligned}\varphi(x+y) &= \varphi\left((a+c) + (b+d)\sqrt{2}\right)\\ &= (a+c) - (b+e)\sqrt{2}\\ &= \left(a - b\sqrt{2}\right) + \left(c - d\sqrt{2}\right)\\ &= \varphi(x) + \varphi(y).\end{aligned}$$

For multiplication, we have

$$\begin{aligned}\varphi(xy) &= \varphi\left((ac + 2bd) + (ad + bc)\sqrt{2}\right)\\ &= (ac + 2bd) - (ad + bc)\sqrt{2}\end{aligned}$$

and

$$\begin{aligned}\varphi(x)\,\varphi(y) &= \left(a - b\sqrt{2}\right)\left(c - d\sqrt{2}\right)\\ &= (ac + 2bd) - (ad + bc)\sqrt{2}\\ &= \varphi(xy).\end{aligned}$$

We have show that φ is a homomorphism; it should be clear that it is one-to-one and onto from the fact that all we did was switch $\pm\sqrt{2}$. Thus, φ is a field isomorphism on $\mathbb{E}$ that fixes $\mathbb{Q}$.

On the other hand, consider the permutation $(1\ 3)$, which would exchange $\sqrt{2}$ and $\sqrt[3]{3}$. This *cannot* be turned into an isomorphism on $\mathbb{E}$ that fixes $\mathbb{Q}$, since fixing $\mathbb{Q}$ implies

$$\varphi\left(\sqrt{2}\cdot\sqrt{2}\right) = \varphi(2) = 2,$$

while the homomorphism property implies

$$\varphi\left(\sqrt{2}\cdot\sqrt{2}\right) = \varphi\left(\sqrt{2}\right)\varphi\left(\sqrt{2}\right) = \sqrt[3]{3}\cdot\sqrt[3]{3} \neq 2,$$

a contradiction.

This example illustrates an important property.

Theorem 9·23. *If $\mathbb{E}$ is a radical extension of $\mathbb{F}$, and $\alpha, \beta \in \mathbb{E}$ such that $\alpha^m, \beta^n \in \mathbb{F}$ but $\alpha^m \neq \beta^m$, then no isomorphism over $\mathbb{E}$ both fixes $\mathbb{F}$ and exchanges α and β.*

Proof. By way of contradiction, suppose there is such an isomorphism φ. Let $q \in \mathbb{F}$ such that $\alpha^m = q$. By substitution and the homomorphism property,

$$\varphi(\beta^m) = [\varphi(\beta)]^m = \alpha^m = q = \varphi(q) = \varphi(\alpha^m).$$

We chose φ to be an isomorphism, hence one-to-one. By definition of one-to-one, we infer that $\alpha^m = \beta^m$, which contradicts the hypothesis that $\alpha^m \neq \beta^m$. □

Question 9·24 .

Let $f \in \mathbb{F}[x]$ be irreducible over $\mathbb{F}$, and $\mathbb{E}$ an extension of $\mathbb{F}$. Show that if $\varphi : \mathbb{E} \to \mathbb{E}$ is an automorphism that fixes $\mathbb{F}$, and $\alpha \in \mathbb{E}$ is a root of f, then $\varphi(\alpha)$ is also a root of f.

In short, we can obtain an isomorphism by permuting $\sqrt[3]{3}$ with other cube roots of three ($\omega\sqrt[3]{3}$, $\omega^2\sqrt[3]{3}$), and we can obtain an isomorphism by permuting $\sqrt{2}$ with other square roots of 2 ($-\sqrt{2}$ only), but we cannot obtain an isomorphism by permuting $\sqrt[3]{3}$ with $\sqrt{2}$. We have shown that

Isomorphisms in the extension field that fix the base field isolate roots of the base field.

Such a fundamental relationship deserves a special name.

Definition 9·25. Let $\mathbb{E}$ be an extension of $\mathbb{F}$. The set of field automorphisms of $\mathbb{E}$ that fixes $\mathbb{F}$ is the **Galois set** of $\mathbb{E}$ over $\mathbb{F}$. We write $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$.

Our first observation of the Galois set is that it's actually a *group*.

Theorem 9·26. *The Galois set of an extension is a group.*

Proof. Let $\mathbb{E}$ be any extension of a field $\mathbb{F}$, and let G be its Galois set. We wish to show that G is a group. Since G consists of automorphisms, which are functions, which satisfy the associative property, the elements of G satisfy the associative property. The identity automorphism ι over $\mathbb{E}$ fixes elements of $\mathbb{E}$, so it likewise fixes elements of $\mathbb{F}$, so $\iota \in G$.

To show that G is closed, let $\varphi, \psi \in G$. Let $a \in \mathbb{F}$; by definition, $\varphi(a) = a$ and $\psi(a) = a$, so $(\varphi \circ \psi)(a) = \varphi(\psi(a)) = a$. We know from before that the composition of one-to-one, onto functions is one-to-one and onto, and the composition of homomorphisms is a homomorphism. Thus, $\varphi \circ \psi \in G$.

It remains to show that G contains the inverses of its elements. Let $\varphi \in G$. Since φ is an automorphism, it has an inverse, ψ, which is also a field automorphism. Let $a \in \mathbb{F}$; by definition, $\varphi(a) = a$, so $\psi(a) = \varphi^{-1}(a) = a$. Hence, ψ fixes $\mathbb{F}$, so that by definition, $\psi \in G$. Since φ was an arbitrary element of G, every element of G has an inverse in G itself.

We have shown that G satisfies the definition of a group. By definition, $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = G$ is a group. □

Our second observation is that the Galois group of a radical extension has a wonderfully simple form.

Theorem 9·27. *Let $p \in \mathbb{N}^+$ be irreducible, and $\mathbb{F}$ a field that contains a primitive p-th root of unity. If $\alpha^p \in \mathbb{F}$, then $\mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F}) \cong \mathbb{Z}_p$.*

One reason we first adjoin a primitive p-th root of unity is the discussion at the end of Section 9·1, where we saw that in order to obtain all the roots of $x^p - a$ we must adjoin not only $\sqrt[p]{a}$, but a primitive p-th root of unity, as well. We will talk about the Galois group of an extension by a primitive p-th root of unity in Exercise 9.30. (See also Exercise 9.12.)

Proof. Assume $\alpha^p \in \mathbb{F}$. For convenience, write $\mathbb{E} = \mathbb{F}(\alpha)$ and $q = \alpha^p$. Let $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. By Theorem 9·23, any $\varphi \in G$ satisfies $\varphi(\alpha) = \beta$ only if β is another pth root of α^p. By Theorem 9·10, $\beta = \omega^m \alpha$ where ω is a primitive p-th root of unity and m lies between 0 and $p-1$, inclusive. Thus, any $\varphi \in G$ has p choices for where to map.

Can we have that many, though? In other words, do all such choices lead to an isomorphism that fixes $\mathbb{F}$? We claim that they do. To see why, let $0 \leqslant i < p-1$ and define, for any $m \in \mathbb{N}^+$, $\varphi\left(\sum_{j=0}^{m} b_j \alpha^j\right) = \sum_{j=0}^{m} b_j (\omega^i \alpha)^j$. It is clear that φ fixes $\mathbb{F}$, since any $a \in \mathbb{F}$ can be written as $a + 0 \cdot \alpha$, and by definition $\varphi(a + 0 \cdot \alpha) = a + 0(\omega^i \alpha) = \alpha$. To see why φ is a homomorphism, observe that for any $a, b, c, d \in \mathbb{F}$, we have

$$\varphi\left(\left(\sum_{j=0}^{p-1} b_j \alpha^j\right)\left(\sum_{k=0}^{p-1} c_k \alpha^k\right)\right) = \varphi\left(\sum_{j=0}^{2p-2}\left[\sum_{k+\ell=j}(b_k c_\ell)\right]\alpha^j\right)$$
$$= \sum_{j=0}^{2p-2}\left[\sum_{k+\ell=j}(b_k c_\ell)\right](\omega^i\alpha)^j$$

and

$$\begin{aligned}\varphi\left(\sum_{j=0}^{p-1} b_j\alpha^j\right)\varphi\left(\sum_{k=0}^{p-1} c_k\alpha^k\right) &= \left[\sum_{j=0}^{p-1} b_j\left(\omega^i\alpha\right)^j\right]\left[\sum_{k=0}^{p-1} c_k\left(\omega^i\alpha\right)^k\right]\\ &= \sum_{j=0}^{p-1}\sum_{k=0}^{p-1} b_jc_k\left(\omega^i\alpha\right)^{j+k}\\ &= \sum_{j=0}^{2p-2}\left[\sum_{k+\ell=j}\left(b_kc_\ell\right)\right]\left(\omega^i\alpha\right)^j.\end{aligned}$$

(The j and k in the last line are *not* the same as the j and k in the one before it.)

Is φ one-to-one? The definition of φ guarantees that $\varphi(ax) = a\varphi(x)$ for any $a \in \mathbb{F}$ and any $x \in \mathbb{E}$, so a problem can arise only if $\varphi\left(\omega^j\alpha\right) = \varphi\left(\omega^k\alpha\right)$ for some $0 \leqslant j, k < p$. Recall that $\mathbb{F}$ contains a primitive pth root of unity ω, so $\varphi\left(\omega^j\alpha\right) = \varphi(\omega)^j\varphi(\alpha) = \omega^j\left(\omega^i\alpha\right) = \omega^{ij}\alpha$. Likewise, $\varphi\left(\omega^k\alpha\right) = \omega^{ik}\alpha$. By substitution, $\omega^{ij}\alpha = \omega^{ik}\alpha$; multiply both sides by $\omega^{-i}\alpha^{-1}$ to obtain $\omega^j = \omega^k$. In other words, φ remains one-to-one.

Is φ onto? As before, we need merely ensure that for any $k = 0, \ldots, p-1$ we can find $j \in \{0, \ldots, p-1\}$ such that $\varphi\left(\omega^j\alpha\right) = \omega^k\alpha$. To that end, let $k \in \{0, \ldots, p-1\}$. By substitution, $\varphi\left(\omega^{k-i}\alpha\right) = \omega^i\omega^{k-i}\alpha = \omega^k\alpha$. Since k was arbitrary, φ is onto.

Since i was arbitrary, we conclude that, for any choice of $i = 0, \ldots, p-1$, the choice of $\varphi(\omega\alpha) = \omega^i\alpha$ is an isomorphism, and so there are at least p isomorphisms in G.

We had already found that there are at most p isomorphisms in G; we have now found that there are at least that many. Together, this means $|G| = p$. Recall that p is irreducible; up to isomorphism, there is only one group of order p, $\mathbb{Z}_p$ (see below). Hence, $\mathrm{Gal}(\mathbb{E}/\mathbb{F}) \cong \mathbb{Z}_p$. □

Question 9·28 .

Let p be irreducible. Explain why $\mathbb{Z}_p$ is the only group of order p, up to isomorphism. (Hint: Use a corollary to Langrange's Theorem to determine the group's structure. It then becomes straightforward to describe an isomorphism from any group of order p to $\mathbb{Z}_p$.)

Question 9·29 .

Show that when p is irreducible, every non-identity element of Ω_p is a primitive root of unity.

Question 9·30 .

Suppose that ω is a primitive p-th root of unity, where $p > 2$ and p is irreducible. Show that if $\omega \notin \mathbb{F}$, then $\mathrm{Gal}(\mathbb{F}(\omega)/\mathbb{F}) \cong \mathbb{Z}_{p-1}$.

Solving polynomials by radicals

We want to know whether we can solve a polynomial over $\mathbb{Q}$ by radicals; that is, if for any $f \in \mathbb{Q}[x]$ we can construct a radical extension $\mathbb{E} = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ containing all the roots of f. We can certainly construct *some* extension field $\mathbb{E}$ containing all the roots of f using quotient

groups, and our study of permutations of the roots had led us to develop the notion of the Galois group of an extension field, $\mathrm{Gal}(\mathbb{E}/\mathbb{Q})$. We now have to put everything together.

We concluded the last section with the observation that the Galois group of a radical extension by one root of irreducible degree is isomorphic to $\mathbb{Z}_p$. Let's look at the example $f = (x^2-2)(x^3-3)$. Putting ω as a primitive cube root of unity as before, we extend $\mathbb{Q}$ in parts, called a **tower of fields** or a **tower of extensions**, obtaining

$$\mathbb{Q} \subsetneq \mathbb{Q}\left(\sqrt{2}\right) \subsetneq \mathbb{Q}\left(\sqrt{2},\omega\right) \subsetneq \mathbb{Q}\left(\sqrt{2},\omega,\sqrt[3]{3}\right) = \mathbb{E}.$$

If we write $\mathbb{F}_0 = \mathbb{Q}$, $\mathbb{F}_1 = \mathbb{Q}\left(\sqrt{2}\right)$, $\mathbb{F}_2 = \mathbb{Q}\left(\sqrt{2},\omega\right)$, and $\mathbb{F}_3 = \mathbb{E}$, what can we say about $\mathrm{Gal}(\mathbb{F}_3/\mathbb{F}_i)$ for $i = 0,1,2,3$?

We shall adopt the convention that we add a primitive p-th root of unity before adding $\sqrt[p]{a}$, unless a primitive root of unity is already in the field. We also remind the reader that we consider only algebraic extensions, as that is the focus of our inquiry; that is, if $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E}$, then $\mathbb{K}$ is an algebraic extension of $\mathbb{F}$, and $\mathbb{E}$ is an algebraic extension of $\mathbb{K}$.

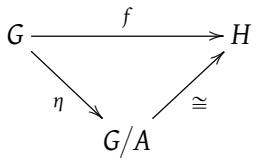
Theorem 9·31. *If $\mathbb{F} \subsetneq \mathbb{F}(\alpha) \subsetneq \mathbb{E}$ is a tower of extensions of $\mathbb{F}$, where $\mathbb{F}(\alpha)$ is a radical extension of degree p, p is irreducible, and*

- *α is a primitive p-th root of unity, or*
- *$\mathbb{F}$ contains a primitive p-th root of unity,*

then

- $\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) \lhd \mathrm{Gal}(\mathbb{E}/\mathbb{F})$, *and*
- *the corresponding quotient group is abelian.*

Proof of Theorem 9·31. The basic idea is to use the Isomorphism Theorem for Groups (Theorem 4·159 on page 175). For a homomorphism f from G onto H, with $A = \ker f$, we have the following diagram.



(Fact 4·153 guarantees that $\ker f$ is a normal subgroup of G.) Suppose we set $H = \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$. Depending on whether α is a primitive p-th root of unity or $\mathbb{F}$ contains a primitive p-th root of unity, $H = \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ is isomorphic either to $\mathbb{Z}_p$ (Theorem 9·27) or $\mathbb{Z}_{p-1}$ (Question 9.30). If we can find a way to set $G = \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ and map G *onto* H in such a way that $\ker f = \mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$, we would first have

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) = \ker f \quad \lhd \quad G = \mathrm{Gal}(\mathbb{E}/\mathbb{F}),$$

and by the Isomorphism Theorem

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F})/\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha)) \quad \cong \quad \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F}) = H,$$

so that, since H is abelian, the quotient group is abelian, as desired.

To this end, define $f : \mathrm{Gal}(\mathbb{E}/\mathbb{F}) \to \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ by *restriction to* $\mathbb{F}(\alpha)$, which means that f assigns each $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$ to $\tau \in \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ so long as $\tau(x) = \sigma(x)$ for every $x \in \mathbb{F}(\alpha)$.

Is f well-defined? Assume that f can map σ to either τ or $\widehat{\tau}$. By definition, $\tau(x) = \sigma(x) = \widehat{\tau}(x)$ for every $x \in \mathbb{F}(\alpha)$. However, the domain of τ and $\widehat{\tau}$ is precisely $\mathbb{F}(\alpha)$, so $\tau = \widehat{\tau}$. Hence, f is indeed well-defined.

Is f a homomorphism? Let $\sigma, \widehat{\sigma} \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$, $\tau = f(\sigma)$, $\widehat{\tau} = f(\widehat{\sigma})$, and $\mu = f(\sigma\widehat{\sigma})$. To show that f is a homomorphism, we have to show that $(\tau\widehat{\tau})(x) = \mu(x)$ for each $x \in \mathbb{F}(\alpha)$. So, let $x \in \mathbb{F}(\alpha)$. By definition, $\widehat{\tau}(x) = \widehat{\sigma}(x)$, so substitution gives us $(\tau\widehat{\tau})(x) = \tau(\widehat{\tau}(x)) = \tau(\widehat{\sigma}(x))$. Since $\widehat{\sigma}$ is an automorphism, $\widehat{\sigma}(x) \in \mathbb{F}(\alpha)$, so by definition, $\tau(\widehat{\sigma}(x)) = \sigma(\widehat{\sigma}(x))$. On the other hand, the definition of μ tells us that $\mu(x) = (\sigma\widehat{\sigma})(x) = \sigma(\widehat{\sigma}(x))$. We just saw that this was the same as $(\tau\widehat{\tau})(x)$, and x was arbitrary in $\mathbb{F}(\alpha)$; thus, $f(\sigma) f(\widehat{\sigma}) = \tau\widehat{\tau} = \mu = f(\sigma\widehat{\sigma})$, and we are indeed dealing with a homomorphism.

Is f onto? Let $\tau \in \mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$, and define

$$\sigma(x) = \begin{cases} \tau(x), & x \in \mathbb{F}(\alpha); \\ x, & \text{otherwise.} \end{cases}$$

You will show in Question 9.33 that $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$, and it is clear from the definition of σ that $f(\sigma) = \tau$. Thus, f is indeed onto.

So, what is $\ker f$? By definition, $\sigma \in \ker f$ if and only if $f(\sigma)$ is the identity homomorphism ι of $\mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$. An identity homomorphism maps every element to itself; in this case, $\iota(x) = x$ for all $x \in \mathbb{F}(\alpha)$. Thus, $\sigma \in \ker f$ if and only if $\sigma(x) = x$ for all $x \in \mathbb{F}(\alpha)$. This implies that σ is an automorphism of $\mathbb{E}$ that fixes not only $\mathbb{F}$, but $\mathbb{F}(\alpha)$, as well! In other words, $\sigma \in \mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$! Since σ was arbitrary, $\ker f = \mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$.

We have shown that f is a function from $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ onto $\mathrm{Gal}(\mathbb{F}(\alpha)/\mathbb{F})$ whose kernel is $\mathrm{Gal}(\mathbb{E}/\mathbb{F}(\alpha))$. As explained in the first paragraph of the proof, this completes the theorem. □

We rely on the following corollary in subsequent sections.

Corollary 9·32. *Let $\mathbb{F} \subsetneq \mathbb{F}(\alpha_1) \subsetneq \mathbb{F}(\alpha_1, \alpha_2) \subsetneq \cdots \mathbb{F}(\alpha_1, \ldots, \alpha_n)$ be a tower of radical extensions of irreducible degree, where we always add a primitive p-th root of unity before any other p-th root. There exist subgroups $G_1, \ldots, G_n$ of $\mathrm{Gal}(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F})$ such that*

$$\begin{aligned} \{\mathfrak{R}\} = G_0 &\lhd G_1 \\ G_1 &\lhd G_2 \\ &\vdots \\ G_{n-1} &\lhd G_n = \mathrm{Gal}(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}) \end{aligned}$$

and the corresponding quotient groups are abelian.

Proof. Apply repeatedly the preceding theorem with $\mathbb{E} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$, $\alpha_{\text{theorem}} = \alpha_k$, and

$\mathbb{F}_{\text{theorem}} = \mathbb{F}(\alpha_1, \ldots, \alpha_{k-1})$ to build the abelian quotient groups

$$\begin{gathered}\text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}\right)/\text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1)\right) \\ \text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1)\right)/\text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1, \alpha_2)\right) \\ \vdots \\ \text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1, \ldots, \alpha_{n-1})\right)/\text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1, \ldots \alpha_n)\right).\end{gathered}$$

From these groups, the following assignments satisfy the claim:

$$\begin{aligned}G_0 &= \text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1, \ldots, \alpha_n)\right) \\ G_1 &= \text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1, \ldots, \alpha_{n-1})\right) \\ &\vdots \\ G_{n-1} &= \text{Gal}\left(\mathbb{F}(\alpha_1, \ldots, \alpha_n)/\mathbb{F}(\alpha_1)\right).\end{aligned}$$

□

Question 9·33 .

Suppose $\mathbb{E} \supsetneq \mathbb{K} \supsetneq \mathbb{F}$ is a tower of fields. Let $\tau \in \text{Gal}(\mathbb{K}/\mathbb{F})$. Define $\sigma : \mathbb{E} \to \mathbb{E}$ by

$$\begin{cases}\sigma(x) = \tau(x), & x \in \mathbb{K}; \\ \sigma(x) = x, & \text{otherwise.}\end{cases}$$

Show that $\sigma \in \text{Gal}(\mathbb{E}/\mathbb{F})$.

9·4 "Solvable" groups

We found in the previous section that the Galois groups corresponding to each step of a tower of radical extensions form abelian quotient groups. We study this property in some detail in this section, and start by generalizing the property to arbitrary groups.

Definition 9·34. If a group G contains subgroups $G_0, G_1, \ldots, G_n$ such that

- $G_0 = \{я\}$;
- $G_n = G$;
- $G_{i-1} \lhd G_i$; and
- G_i/G_{i-1} is abelian,

then G is a **solvable group**. The chain of subgroups $G_0, \ldots, G_n$ is called a **normal series**.

Example 9·35. Any finite abelian group G is solvable: let $G_0 = \{я\}$ and $G_1 = G$. Subgroups of an abelian group are always normal, so $G_0 \lhd G_1$. In addition, $X, Y \in G_1/G_0$ implies that $X = x\{я\}$ and $Y = y\{я\}$ for some $x, y \in G_1 = G$. Since G is abelian,

$$XY = (xy)\{я\} = (yx)\{я\} = YX.$$

Example 9·36. The group D_3 is solvable. To see this, let $n = 2$ and $G_1 = \langle \rho \rangle$:

- By Exercise 4.127 on page 161, $\{\iota\} \lhd G_1$. To see that $G_1/\{\iota\}$ is abelian, note that for any $X, Y \in G_1/\{\iota\}$, we can write $X = x\{\iota\}$ and $Y = y\{\iota\}$ for some $x, y \in G_1$. By definition of G_1, we can write $x = \rho^a$ and $y = \rho^b$ for some $a, b \in \mathbb{Z}$. We can then fall back on the commutative property of addition in $\mathbb{Z}$ to show that

$$\begin{aligned} XY &= (xy)\{\iota\} = \rho^{a+b}\{\iota\} \\ &= \rho^{b+a}\{\iota\} = (yx)\{\iota\} = YX. \end{aligned}$$

- By Exercise 4.133 and the fact that $|G_1| = 3$ and $|G_2| = 6$, we know that $G_1 \lhd G_2$. The same exercise tells us that G_2/G_1 is abelian.

Question 9·37 .
Explain why Ω_n is solvable for any $n \in \mathbb{N}^+$.

Question 9·38 .
Show that Q_8 is solvable.

A rather surprising property of solvable groups is that their subgroups and quotient groups are also solvable. Showing that quotient groups are solvable is a little easier, so we start with that first.

Theorem 9·39. *Every quotient group of a solvable group is solvable.*

Proof. Let G be a solvable group and $A \lhd G$. We need to show that G/A is solvable. Since G is solvable, choose a normal series $G_0, \ldots, G_n$. For each $i = 0, \ldots, n$, put

$$A_i = \{gA : g \in G_i\}.$$

We claim that the chain $A_0, A_1, \ldots, A_n$ likewise satisfies the definition of a solvable group.

First, we show that $A_{i-1} \lhd A_i$ for each $i = 1, \ldots, n$. First we must show that $A_{i-1} < A_i$. It is a subset because any $X \in A_{i-1}$ has the form xA where $x \in G_{i-1} \subseteq G_i$, so $x \in G_i$ and thus $xA \in A_i$. To show that it is a subgroup, let $X, Y \in A_{i-1}$, and $x, y \in G_{i-1}$ such that $X = xA$ and $Y = yA$. By substitution and coset arithmetic, $XY^{-1} = (xA)(yA)^{-1} = (xy^{-1})A$. Recall that G_{i-1} is a subgroup of G_i; by the Subgroup Theorem, $xy^{-1} \in G_{i-1}$, so $XY^{-1} = (xy^{-1})A \in A_{i-1}$

Now we show that A_{i-1} is a normal subgroup. Let $X \in A_i$; by definition, $X = xA$ for some $x \in G_i$. We have to show that $XA_{i-1} = A_{i-1}X$. Let $Y \in A_{i-1}$; by definition, $Y = yA$ for some $y \in G_{i-1}$. Recall that $G_{i-1} \lhd G_i$, so there exists $\widehat{y} \in G_{i-1}$ such that $xy = \widehat{y}x$. Let $\widehat{Y} = \widehat{y}A$; since $\widehat{y} \in G_{i-1}$, $\widehat{Y} \in A_{i-1}$. Using substitution and the definition of coset arithmetic, we have

$$XY = (xy)A = (\widehat{y}x)A = \widehat{Y}X \in A_{i-1}X.$$

Since Y was arbitrary in A_{i-1}, $XA_{i-1} \subseteq A_{i-1}X$. A similar argument shows that $XA_{i-1} \supseteq A_{i-1}X$, so the two are equal. Since X is an arbitrary coset of A_{i-1} in A_i, we conclude that $A_{i-1} \lhd A_i$.

Second, we show that A_i/A_{i-1} is abelian. Let $X, Y \in A_i/A_{i-1}$. By definition, we can write $X = SA_{i-1}$ and $Y = TA_{i-1}$ for some $S, T \in A_i$. Again by definition, there exist $s, t \in G_i$ such that $S = sA$ and $T = tA$. We know that

$$\begin{aligned} XY = YX &\Leftrightarrow (SA_{i-1})(TA_{i-1}) = (TA_{i-1})(SA_{i-1}) \\ &\Leftrightarrow (ST)A_{i-1} = (TS)A_{i-1} \\ &\Leftrightarrow (ST)^{-1}(TS) \in A_{i-1} \\ &\Leftrightarrow T^{-1}S^{-1}TS \in A_{i-1}. \end{aligned}$$

By substitution and coset arithmetic,

$$T^{-1}S^{-1}TS = \left(t^{-1}A\right)\left(s^{-1}A\right)(tA)(sA) = \left(t^{-1}s^{-1}ts\right)A.$$

Recall that $G_{i-1} \lhd G_i$ and G_i/G_{i-1} is abelian, so

$$(tG_{i-1})(sG_{i-1}) = (sG_{i-1})(tG_{i-1}) \quad \Leftrightarrow \quad (ts)G_{i-1} = (st)G_{i-1} \quad \Leftrightarrow \quad t^{-1}s^{-1}ts \in G_{i-1}.$$

By substitution,

$$T^{-1}S^{-1}TS \in A_{i-1}.$$

Following the above chain of equivalences back to their beginning, we have $XY = YX$. Since X and Y were arbitrary in the quotient group A_i/A_{i-1}, we conclude that it is abelian.

We have constructed a normal series in G/A; it follows that G/A is solvable. □

Question 9·40 .

In Question 9.38 you showed that Q_8 is solvable. From Theorem 9·39 you know the quotient group $Q_8/\langle -\mathbf{1}\rangle$ is also solvable. List a normal series.

The following result is also true:

Theorem 9·41. *Every subgroup of a solvable group is solvable.*

To prove Theorem 9·41, we need the definition of the commutator from Questions 2.30 on page 53 and 4.151 on page 170, and a few properties of commutator subgroups. If you skipped those before, you should go back and do them now, to familiarize yourself with the idea.

Definition 9·42. Let G be a group. The **commutator subgroup** G' of G is the intersection of all subgroups of G that contain $[x, y]$ for all $x, y \in G$.

Notice that $G' < G$ by Question 4.10.

Notation 9·43. We wrote G' as $[G, G]$ in Question 4.151.

Question 9·44 .

Compute G' for Q_8. Then compute $(G')'$ (we call this $G^{(2)}$ further below) and $\left((G')'\right)'$. Keep going until you can go no further; how do you know you can go no further?

Lemma 9·45. *For any group G, $G' \lhd G$. In addition, G/G' is abelian.*

Proof. You showed that $G' \lhd G$ in Question 4.151. To show that G/G' is abelian, let $X, Y \in G/G'$. Write $X = xG'$ and $Y = yG'$ for appropriate $x, y \in G$. By definition, $XY = (xy)\,G'$. Let $g' \in G'$; by definition, g' is in every group that contains all the commutators of G. Closure ensures that the product of g' with another element of G' is also in G'; certainly the commutator $[x,y]$ is in G', so $[x,y]\,g' \in G'$. Write $z = [x,y]\,g'$. Substitution and properties of groups allows to infer

$$[x,y]\,g' = z \quad \Longrightarrow \quad \left(x^{-1}y^{-1}xy\right) g' = z \quad \Longrightarrow \quad (xy)\,g' = (yx)\,z.$$

Thus, $(xy)\,g' \in (yx)\,G'$. Since g' was arbitrary, $(xy)\,G' \subseteq (yx)\,G'$. Similar reasoning shows that $(xy)\,G' \supseteq (yx)\,G'$, which gives us equality. Substitution gives us

$$XY = (xy)\,G' = (yx)\,G' = YX.$$

We conclude that G/G' is abelian. □

Lemma 9·46. *If $H \subseteq G$, then $H' \subseteq G'$.*

Proof. You do it! See Question 9.47. □

Question 9·47 .
Show that if $H \subseteq G$, then $H' \subseteq G'$.

Notation 9·48. Define $G^{(0)} = G$ and $G^{(i)} = \left(G^{(i-1)}\right)'$; that is, $G^{(i)}$ is the commutator subgroup of $G^{(i-1)}$.

Lemma 9·49. *A group is solvable if and only if $G^{(n)} = \{\mathfrak{R}\}$ for some $n \in \mathbb{N}$.*

Proof. ($\Longleftarrow$) Suppose that $G^{(n)} = \{\mathfrak{R}\}$ for some $n \in \mathbb{N}$. By Lemma 9·45, the subgroups form a normal series; that is,

$$\{\mathfrak{R}\} = G^{(n)} \lhd G^{(n-1)} \lhd \cdots \lhd G^{(0)} = G$$

and $G^{(n-i)}/G^{(n-(i-1))}$ is abelian for each $i = 0, \ldots, n-1$. As this is a normal series, we have shown that G is solvable.

($\Longrightarrow$) Suppose that G is solvable. Let $G_0, \ldots, G_n$ be a normal series for G. We claim that $G^{(n-i)} \subseteq G_i$. If this claim were true, then $G^{(n-0)} \subseteq G_0 = \{\mathfrak{R}\}$, and we would be done. We proceed by induction on $n - i \in \mathbb{N}$.

Inductive base: If $n - i = 0$, then $G^{(n-i)} = G = G_n$. Also, $i = n$, so $G^{(n-i)} = G_n = G_i$, as claimed.

Inductive hypothesis: Assume that the assertion holds for $n - i$.

Inductive step: By definition, $G^{(n-i+1)} = \left(G^{(n-i)}\right)'$. By the inductive hypothesis, $G^{(n-i)} \subseteq G_i$; by Lemma 9·46, $\left(G^{(n-i)}\right)' \subseteq G_i'$. Hence

$$G^{(n-i+1)} \subseteq G_i'. \tag{9.1}$$

We now show that $G_i' \subseteq G_{i-1}$. Recall from the properties of a normal series that G_i/G_{i-1} is abelian; for any $x, y \in G_i$, we have

$$\begin{aligned}(xy)\,G_{i-1} &= (xG_{i-1})\,(yG_{i-1}) \\ &= (yG_{i-1})\,(xG_{i-1}) = (yx)\,G_{i-1}.\end{aligned}$$

By equality of cosets, $(yx)^{-1}(xy) \in G_{i-1}$ (Lemma 4·103 on page 152); in other words, $[x,y] = x^{-1}y^{-1}xy \in G_{i-1}$. Since x and y were arbitrary in G_i, we have $G_i' \subseteq G_{i-1}$. Along with (9.1), this implies that $G^{(n-(i-1))} = G^{(n-i+1)} \subseteq G_{i-1}$.

We have shown the claim; thus, $G^{(n)} = \{\mathfrak{R}\}$ for some $n \in \mathbb{N}$. □

We can now prove Theorem 9·41.

Proof of Theorem 9·41. Let $H < G$. Assume G is solvable; by Lemma 9·49, $G^{(n)} = \{\mathfrak{R}\}$. By Lemma 9·46, $H^{(i)} \subseteq G^{(i)}$ for all $n \in \mathbb{N}$, so $H^{(n)} \subseteq \{\mathfrak{R}\}$. By the definition of a group, $H^{(n)} \supseteq \{\mathfrak{R}\}$, so the two are equal. By the same lemma, H is solvable. □

Question 9·50 .

In the textbook *God Created the Integers…* the theoretical physicist Stephen Hawking collects reprints of some of the greatest mathematical results in history, adding some commentary. For an excerpt from Evariste Galois' *Memoirs,* Hawking sums up the main result this way.

> To be brief, Galois demonstrated that the general polynomial of degree n could be solved by radicals if and only if every subgroup N of the group of permutations S_n is a normal subgroup. Then he demonstrated that every subgroup of S_n is normal for all $n \leqslant 4$ but not for any $n > 5$. —p. 105

Unfortunately, Hawking's explanation is completely wrong, and this exercise leads you towards an explanation as to why.[1] Recall from Fact 7·30 on page 267 that S_3 is isomorphic to D_3; you can work with whichever group is more comfortable for you.

(a) Find all six subgroups of S_3.

(b) It is known that the general polynomial of degree 3 can be solved by radicals. According to the quote above, what must be true about all the subgroups of S_3?

(c) Why is Hawking's explanation of Galois' result "obviously" wrong?

Question 9·51 .

Show that S_4 is solvable, and explain why this means any degree-four polynomial can be solved by radicals.

9·5 The Theorem of Abel and Ruffini

In this section, we use the characterization of solution by radicals in Theorem 9·31 and Definition 9·34 to show that some polynomials cannot be solved by radicals. The basic idea is that

[1] Perhaps Hawking was trying to simplify what Galois actually showed, and went too far. (I've done much worse, on occasion.) You will see the actual result in the next section.

S_5 is not a solvable group, and we can find a degree-5 polynomial whose Galois group is S_5. Before we dive into that, though, we need an important fact about the order of a group.

A "reverse-Lagrange" Theorem

Lagrange's Theorem tells us that the order of any element g of a group G must divide the order of a group; that is, $\text{ord}(g) \mid |G|$. You might wonder whether the reverse is true; that is, if m is an integer that divides $|G|$, can we always find $g \in G$ such that $\text{ord}(g) = m$? The easy answer is, "Of course not;" after all, we could find $g \in G$ such that $\text{ord}(g) = |G|$, and every group would be cyclic. Nevertheless, some interesting properties *do* hold, and one of them is critical to the result we want.

Cauchy's Theorem. *Let $p \in \mathbb{N}^+$ be irreducible, and let G be a group. If $p \mid |G|$, then we can find $g \in G$ such that* $\text{ord}(g) = p$.

The property is not true in general, as you can show:

Question 9·52 .
Find a finite group G where $m = |G|$, m is not irreducible, and no $g \in G$ has $\text{ord}(g) = m$.

We start with the case where G is abelian, as this is a special case of the more general problem.

Lemma 9·53. *Cauchy's Theorem is true if G is abelian.*

Proof. Suppose that G is an abelian group, $p \in \mathbb{N}^+$ is irreducible, and $p \mid |G|$. We proceed by induction on $|G|$.

Inductive base: If $|G| = 1$, then *no* irreducible number divides $|G|$, and the theorem is "vacuously" true.

Inductive hypothesis: Let $n \in \mathbb{N}^+$, and suppose that any abelian group whose size is $n < |G|$, and where $p \mid n$, contain at least one element whose order is p.

Inductive step: Let $g \in G \setminus \{\text{я}\}$. If $p \mid \text{ord}(g)$, then let $d = \text{ord}(g)/p$, and the group

$$\left\langle g^d \right\rangle = \left\{ g^d, \left(g^d\right)^2, \ldots, \left(g^d\right)^{p-1}, \left(g^d\right)^p = g^{\text{ord}(g)} = \text{я} \right\}$$

will have order p. Otherwise, $p \nmid \text{ord}(g)$. Let $Q = G/\langle g \rangle$; the size of Q is, by definition, the number of cosets of $\langle g \rangle$, which is $|G|/\text{ord}(g)$. Since $p \mid |G|$ but $p \nmid \langle g \rangle$, an application of Lagrange's Theorem shows that $p \mid |Q|$. By hypothesis, G is abelian, so all its subgroups are normal; specifically, $\langle g \rangle$ is normal. Thus, Q is also a group; since $g \neq \text{я}$, the size of Q is less than the size of G, so the inductive hypothesis applies; Q contains an element of order p; call this element X. Let $x \in G$ such that $X = x\langle g \rangle$. Let $m = \text{ord}(x)$ in G. By definition, $x^m = \text{я}$, so

$$X^m = x^m \langle g \rangle = \text{я} \langle g \rangle = \langle g \rangle.$$

Hence X^m is the identity in Q. The order of X is p, so by Exercise 3.56, $p \mid m$. Choose $d \in \mathbb{N}^+$ such that $pd = m$, and then x^d will have order p, just as g^d had order p above. □

We now prove the general case.

Proof of Cauchy's Theorem. As with the abelian case, we proceed by induction, with the inductive base using the same reasoning. We proceed directly to the inductive step.

If G is abelian, then Lemma 9·53 gives us the result, so assume that G is not abelian. Let $Z(G)$ denote the **center** of G,

$$Z(G) = \{g \in G : xg = gx \ \forall x \in G\}.$$

You will show in Question 9.54 that $Z(G)$ is a subgroup of G. Notice that $Z(G)$ is abelian *by definition,* so if $p \mid |Z(G)|$, then Lemma 9·53 gives us an element of order p, and we are done.

Assume, therefore, that $p \nmid |Z(G)|$. For each $x \in G$, define $C_x = \{g \in G : gx = xg\}$. We call C_x the **centralizer** of x; you will show in Exercise 9.55 that this is a subgroup of G. Since $p \nmid |Z(G)|$, $Z(G) \neq G$, so we can find $x \in G \backslash Z(G)$ that does not commute with every element of G, and $|C_x| < |G|$. If $p \mid |C_x|$, the inductive hypothesis applies.

Assume, therefore, that p does not divide the size of any centralizer. Consider G/C_x; since $p \mid |G|$ but $p \nmid |C_x|$, Lagrange's Theorem tells us that $p \mid |G/C_x|$. At this point, we meet up with our old friend conjugation (Definitions 2·29 and 4·143); let x^G be the set of all conjugations of x by some $g \in G$; that is,

$$x^G = \left\{gxg^{-1} : g \in G\right\}.$$

We claim that the set of all these x^G partition G. They certainly cover G, since $x = exe^{-1} \in x^G$, so $x \in x^G$ always. To see that distinct subsets are disjoint, let $x, y \in G$, and suppose $y \in x^G$. By definition, there exists $g \in G$ such that $y = gxg^{-1}$. We can rewrite this expression as $x = g^{-1}yg$, so $x \in y^G$, as well. Moreover, let $z \in x^G$; by definition, we can find $h \in G$ such that

$$z = hxh^{-1} = h\left(g^{-1}yg\right)h^{-1} = \left(hg^{-1}\right)y\left(gh^{-1}\right) = \left(hg^{-1}\right)y\left(hg^{-1}\right)^{-1},$$

so $z \in y^G$. Since z was arbitrary in x^G, $x^G \subseteq y^G$. A similar argument shows that $x^G \supseteq y^G$, so the two must be equal. We have shown that if two subsets are not disjoint, then they are not distinct; thus, if they are distinct, then they are also disjoint. As claimed, the x^G partition G. Use this partition to define $\mathcal{P} \subseteq G$ such that $\cup_{x \in \mathcal{P}} x^G = G$, and for any distinct $x, y \in \mathcal{P}$, $x^G \neq y^G$, so $x^G \cap y^G = \varnothing$. From the partition we can see that $\sum_{x \in \mathcal{P}} |x^G| = |G|$.

We also claim that each x^G satisfies $|x^G| = |G/C_x|$. Why? Let $x \in G$; by definition, for any $y \in x^G$, we can find $g \in G$ such that $gxg^{-1} = y$. Let $\varphi : x^G \to G/C_x$ by $\varphi(y) = gC_x$. We claim that φ is a one-to-one, onto function. We first check that it is a function, since it is possible that more than one $g \in G$ gives us $gxg^{-1} = y$. So, let $g, h \in G$ such that $gxg^{-1} = y = hxh^{-1}$. Rewrite this as $(h^{-1}g)\,x\,(g^{-1}h) = x$, or $(h^{-1}g)\,x\,(h^{-1}g)^{-1} = x$, so $h^{-1}g \in C_x$. The Lemma on coset equality then gives us $hC_x = gC_x$, as needed; φ is, indeed, a function. Is it one-to-one? Suppose $\varphi(y) = \varphi(z)$; let $g \in G$ such that $\varphi(y) = \varphi(z) = gC_x$. By definition of φ, $gxg^{-1} = y$ and $gxg^{-1} = z$; substitution shows us that $y = z$. So, φ is, indeed, one-to-one. Is it onto? For any $gC_x \in G/C_x$, simply let $y = gxg^{-1}$, and by definition, both $y \in x^G$ and $\varphi(y) = gC_x$. So, φ is, indeed, onto. We have found a one-to-one, onto function from x^G to G/C_x; this implies that the two have the same size.

We can finally show what we set out to show. We have constructed $\mathcal{P} \subseteq G$ such that $\sum_{x \in \mathcal{P}} |x^G| = |G|$. For any $x \in Z(G)$, we have

$$x^G = \left\{gxg^{-1} : g \in G\right\} = \left\{gg^{-1}x : g \in G\right\} = \{x\}.$$

In other words, each element of $Z(G)$ has its own set in the partition. That means we can rewrite the sum as

$$|G| = |Z(G)| + \sum_{x \in \mathcal{P} \backslash Z(G)} \left|x^G\right|.$$

We have also seen that $|x^G| = |G/C_x|$ for all $x \in G$, so by substitution,

$$|G| = |Z(G)| + \sum_{x \in \mathcal{P} \backslash Z(G)} |G/C_x|. \tag{9.2}$$

(This important fact is called the **class equation**.) Rewrite this as

$$|G| - \sum_{x \in \mathcal{P} \backslash Z(G)} |G/C_x| = |Z(G)|. \tag{9.3}$$

Recall that if $p \nmid |C_x|$ for each $x \in \mathcal{P} \backslash Z(G)$, then $p \mid |G/C_x|$ for the same x. We have assumed that p does not divide the size of any centralizer, so p must divide the size of *every* G/C_x. By hypothesis, $p \mid |G|$, so p divides the left hand side of 9.3. It must divide the right hand side, as well, which means $p \mid |Z(G)|$, a contradiction.

The only assumptions we made that were not required by the hypothesis were that $p \nmid |Z(G)|$ and $p \nmid |C_x|$ for any x. One of these assumptions must be false, but if so, the fact that their size is smaller than that of G means that the induction hypothesis holds, and we can find $g \in G$ such that $\text{ord}(g) = p$. □

Question 9·54 .
Show that the center $Z(G)$ of a group G is a subgroup of G.

Question 9·55 .
Show that the centralizer C_x of an element x in a group G is a subgroup of G.

Question 9·56 .
Is either $Z(G)$ or $C_x(G)$ guaranteed to be a normal subgroup? Either show why they are, or provide a counterexample.

Question 9·57 .
Compute the class equation for D_3.

Question 9·58 .
Compute the class equation for Q_8.

We cannot solve the quintic by radicals

To show that some polynomials cannot be solved by radicals, we begin with a generalization of the fact that the purely radical roots of a polynomial can only be mapped to other roots of the same radical; that is, we can map $\sqrt[4]{3} \longrightarrow -\sqrt[4]{3}$, but not to $\sqrt{2}$.

Lemma 9·59. *If α and β are roots of an irreducible polynomial $f \in \mathbb{F}[x]$, then there exists a unique isomorphism $\sigma : \mathbb{F}(\alpha) \to \mathbb{F}(\beta)$ with $\sigma(\alpha) = \beta$ and that fixes $\mathbb{F}$.*

Proof. Let $m = \deg f$. Let $\sigma : \mathbb{F}(\alpha) \to \mathbb{F}(\beta)$ by $\sigma\left(\sum_{j=0}^{m-1} a_j\alpha^j\right) = \sum_{j=0}^{m-1} a_j\beta^j$. It is clear from the definition that σ is one-to-one and onto, but is σ a homomorphism? For the sum, this is easy:

$$\begin{aligned}\sigma\left(\sum_{j=0}^{m-1} a_j\alpha^j + \sum_{j=0}^{m-1} b_j\alpha^j\right) &= \sigma\left(\sum_{j=0}^{m-1} \left(a_j + b_j\right)\alpha^j\right)\\ &= \sum_{j=0}^{m-1} \left(a_j + b_j\right)\beta^j\\ &= \sum_{j=0}^{m-1} a_j\beta^j + \sum_{j=0}^{m-1} b_j\beta^j\\ &= \sigma\left(\sum_{j=0}^{m-1} a_j\alpha^j\right) + \sigma\left(\sum_{j=0}^{m-1} b_j\alpha^j\right).\end{aligned}$$

For the product, it is only a little harder:

$$\sigma\left(\sum_{j=0}^{m-1} a_j\alpha^j \cdot \sum_{j=0}^{m-1} b_j\alpha^j\right) = \sigma\left(\sum_{j=0}^{2m-2}\left[\sum_{k+\ell=j}\left(a_kb_\ell\right)\right]\alpha^j\right) = \sum_{j=0}^{2m-2}\left(\sum_{k+\ell=j} a_kb_\ell\right)\beta^j,$$

while

$$\sigma\left(\sum_{j=0}^{m-1} a_j\alpha^j\right)\cdot\sigma\left(\sum_{j=0}^{m-1} b_j\alpha^j\right) = \sum_{j=0}^{m-1} a_j\beta^j \cdot \sum_{j=0}^{m-1} b_j\beta^j = \sum_{j=0}^{2m-2}\left(\sum_{k+\ell=j} a_kb_\ell\right)\beta^j,$$

where the j's in the last equality do not have the same meaning in the left and right expressions.

To show that σ is unique, consider how the isomorphism can map roots. Let $\tau : \mathbb{F}(\alpha) \to \mathbb{F}(\beta)$ be any isomorphism that fixes $\mathbb{F}$. By Question 9.24, $\tau(\alpha)$ must be a root of f. Since τ must fix $\mathbb{F}$, this completely defines τ as a homomorphism, and in addition, it shows that $\tau = \sigma$, since there is no room for distinction. □

Lemma 9·60. *A_5 is not solvable.*

Proof. In a moment we will argue that any normal subgroup that is not $\{\mathfrak{R}\}$ actually contains all the three-cycles, which generate A_5. That leaves only $\{\mathfrak{R}\}$ and A_5 as normal subgroups, so the only way to generate a tower of radical extensions would be using $\{\mathfrak{R}\} \subsetneq A_5$. However, $A_5/\{\mathfrak{R}\} \cong A_5$, which is not abelian! So if the only normal subgroups of A_5 are $\{\mathfrak{R}\}$ and A_5 itself, A_5 cannot be solvable.

We turn our attention to the claim. Let H be a non-trivial normal subgroup of A_5. We first claim that H contains at least one three-cycle. To see why, let $\sigma \in H \setminus \{(1)\}$. Since H is normal, $\tau\sigma\tau^{-1} \in H$ for any $\tau \in A_5$. Consider the possible simplifications.

- You will show in Question that if H contains a two-cycle or a four-cycle, then it also contains a three-cycle. The argument is very similar to the next two.
- If $\sigma = (a\ b)\ (c\ d)$, let $\tau = (a\ b)\ (c\ e)$. Notice that $\tau = \tau^{-1}$. The conjugation tells us that

$$[(a\ b)\ (c\ e)]\ [(a\ b)\ (c\ d)]\ [(a\ b)\ (c\ e)] = (a\ b)\ (d\ e) \in H.$$

  The closure of H implies that it must also contain $(a\ b)\ (c\ d)\ (a\ b)\ (d\ e) = (c\ d\ e)$.
- If $\sigma = (a\ b\ c\ d\ e)$, let $\tau = (a\ b\ c)$. Notice that $\tau^{-1} = (a\ c\ b)$. The conjugation tells us that

$$(a\ b\ c)\ (a\ b\ c\ d\ e)\ (a\ c\ b) = (a\ d\ e\ b\ c) \in H.$$

  The closure of H implies that it must also contain $(a\ b\ c\ d\ e)^2\ (a\ d\ e\ b\ c) = (b\ e\ d)$.

Either way, H contains a three-cycle.

Now we claim that H contains *all* the three-cycles. Suppose H contains $(a\ b\ c)$. By conjugation, it also contains

- $(b\ c\ d)\ (a\ b\ c)\ (b\ d\ c) = (a\ c\ d)$,
- $(b\ c\ e)\ (a\ b\ c)\ (b\ e\ c) = (a\ c\ e)$,
- $(b\ d\ c)\ (a\ b\ c)\ (b\ c\ d) = (a\ d\ b)$,
- $(b\ d)\ (c\ e)\ (a\ b\ c)\ (b\ d)\ (c\ e) = (a\ d\ e)$,
- $(b\ e\ c)\ (a\ b\ c)\ (b\ c\ e) = (a\ e\ b)$,
- $(a\ c\ d)\ (a\ b\ c)\ (a\ d\ c) = (b\ d\ c)$,
- $(a\ d)\ (c\ e)\ (a\ b\ c)\ (a\ d)\ (c\ e) = (b\ e\ d)$,
- $(a\ c\ e)\ (a\ b\ c)\ (a\ e\ c) = (b\ e\ c)$, and
- $(a\ d)\ (b\ e)\ (a\ b\ c)\ (a\ d)\ (b\ e) = (c\ d\ e)$.

Since H is closed, it also contains the inverses of these elements, so H contains at least twenty three-cycles. A counting argument tells us that there are in fact $5!/3! = 20$ three-cycles, so H contains all the three-cycles.

We leave it to the reader to show that A_5 is generated by all the three-cycles; see Question 9.61. □

Question 9·61 .

Show that if a subgroup H of A_5 contains all the three-cycles, then in fact $H = A_5$.

Question 9·62 .

We can also show that A_5 is not solvable by considering its commutators. As usual, let A_5' denote the commutator subgroup of A_5.

(a) Show that $(a\ b\ c) \in A_5'$ for any distinct $a, b, c \in \{1, 2, 3, 4, 5\}$.

(b) Show that $(a\ b)\ (c\ d) \in A_5'$ for any distinct $a, b, c, d \in \{1, 2, 3, 4, 5\}$.

(c) Show that $(a\ b\ c\ d\ e) \in A_5'$ for any distinct $a, b, c, d, e \in \{1, 2, 3, 4, 5\}$.

(d) Explain why this shows that $A_5' = A_5$.

(e) Explain why this shows that A_5 is not solvable.

Corollary 9·63. *S_5 is not solvable.*

Proof. If S_5 were solvable, then Theorem 9·41 would imply that A_5 is solvable. We just saw that A_5 is *not* solvable, so S_5 cannot be solvable, either. □

We turn our attention to finding a polynomial whose Galois group is S_5.

Lemma 9·64 (Eisenstein's Criterion)**.** *Let $f = a_m x^m + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$, and p an irreducible integer. If*

- *$p \mid a_i$ for each $i = 0, \ldots, m-1$,*
- *$p \nmid a_m$, and*
- *$p^2 \nmid a_0$,*

then f is irreducible, even when viewed in $\mathbb{Q}[x]$.

Proof. Suppose f factors in $\mathbb{Z}[x]$ as $f = gh$. It will also factor when considered as a polynomial of $\mathbb{Z}_p[x]$, with the same gh. Assume that p divides every coefficient of f except the leading coefficient, so $f = a_m x^m$ as a polynomial in $\mathbb{Z}_p[x]$, so $g = bx^\beta$ and $h = cx^\gamma$. Observe that p divides the constant terms of g and h, which means that $p^2 \mid a_0$. This contradicts the third criterion, so if f factors in $\mathbb{Z}[x]$, then we cannot satisfy all three criteria.

To complete the proof, we need to show that if f factors in $\mathbb{Q}[x]$, then it also factors in $\mathbb{Z}[x]$. Suppose $f = gh$ is a factorization of f in $\mathbb{Q}[x]$. Rewrite this factorization as $f = d\widehat{g}\widehat{h}$, where $d \in \mathbb{N}^+$ is the least common denominator of the coefficients of, g and h, obtaining an integer factorization of an integer polynomial. Rewrite the factorization again as $f = d'g'h'$, where d' is the product of d and the greatest common divisors of the coefficients of $\widehat{g}$ and of $\widehat{h}$. Notice that d' must be an integer, as d cannot divide g' or h'. We have thus obtained a factorization of f into integer polynomials. □

Question 9·65 .

Use the product rule of Calculus to show that $x = a$ is a repeated root of a polynomial f if and only $f'(a) = 0$.

Theorem 9·66. *There exists a quintic polynomial over $\mathbb{Q}$ that is not solvable by radicals.*

Proof. Let $f(x) = x^5 - 4x + 2$. Using Eisenstein's Criterion and the irreducible integer $p = 2$, we see that f is irreducible over $\mathbb{Q}$. Extend $\mathbb{Q}$ to a field $\mathbb{E}$ that contains all the roots of f.

Since we are working over the real numbers, we resort briefly to calculus. The maxima and minima of $f(x) = x^5 - 4x + 2$ occur when $0 = f'(x) = 5x^4 - 4$; these are $x = \pm\sqrt[4]{4/5}$. If we substitute these values of x into f, we find that

$$f\left(-\sqrt[4]{\frac{4}{5}}\right) \approx -1 + 4 + 2 > 0 \quad \text{and} \quad f\left(\sqrt[4]{\frac{4}{5}}\right) \approx 1 - 4 + 2 < 0.$$

Since neither critical point is also a root, there are no repeated roots (see Question 9.65), so f makes exactly two turns on the real plane, so it can have exactly three roots $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\backslash\mathbb{Q}$. Once we extend $\mathbb{Q}$ with those roots, f factors as

$$f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)\left(x^2 + ax + b\right),$$

where $a, b \in \mathbb{R}$. Since f has no more *real* roots, the quadratic polynomial has complex roots; call them β_1 and β_2. We know from the quadratic formula that if $\beta_1 = c + di$, then $\beta_2 = c - di$.

Now consider the automorphisms of the final extension field $\mathbb{E}$.

- One automorphism is defined homomorphically by $\varphi(i) = -i$; this corresponds to an exchange of the complex roots, or, a transposition in S_5. Of course, it's not enough to claim it's an automorphism that fixes $\mathbb{Q}$; we must actually show this. It *is* clear that φ fixes not only $\mathbb{Q}$, but non-complex elements of $\mathbb{E}$, as well, as mapping $\pm i \to \mp i$ does not affect them in the slightest. You showed in Question 9.20 that φ is a ring isomorphism in $\mathbb{C}$; the same argument applies to $\mathbb{E}$, as well.

- We claim that when $\mathrm{Gal}(\mathbb{E}/\mathbb{Q})$ is viewed as a subgroup of S_5, there must also be a 5-cycle. To see why, consider how we can extend the identity isomorphism $\iota : \mathbb{Q} \to \mathbb{Q}$ to an automorphism on $\mathbb{Q}(\alpha)$, where α is any one of the roots of f. The elements of $\mathbb{F} = \mathbb{Q}[x]/\langle f \rangle$ can be written using the basis $\{1, x + I, \ldots, x^4 + I\}$, and $\mathbb{Q}(\alpha) \cong \mathbb{F}$, so when we view $\mathbb{E}$ as an extension of $\mathbb{Q}(\alpha)$, *each* element that we adjoin can be seen as having a coefficient in $\mathbb{F}$, which has dimension 5. Using similar reasoning, elements of $\mathbb{E}$ can be seen as an extension of $\mathbb{Q}$ with a basis containing $5m$ elements, for some $m \in \mathbb{N}^+$. By Lemma 9·59, there are $5m$ unique isomorphisms extending ι to $\mathbb{E}$, one for each element of the basis of $\mathbb{E}$. Hence, $|\mathrm{Gal}(\mathbb{E}/\mathbb{F})| = 5m$. What matters here is that the size of the group is divisible by 5; we can now apply Cauchy's Theorem to show that $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ has an element of order 5; in other words, a 5-cycle.

Once we have a two-cycle and a five-cycle in $\mathrm{Gal}(\mathbb{E}/\mathbb{Q})$, we can show that $\mathrm{Gal}(\mathbb{E}/\mathbb{Q}) \cong S_5$ (Question 9.67). We know from Corollary 9·63 that S_5 is not solvable. Apply the contrapositive of Theorem 9·31 to see that f cannot be solved by radicals. □

Question 9·67 .

Suppose a subgroup H of S_5 has a two-cycle and a five-cycle. Show that $H = S_5$.

9·6 The Fundamental Theorem of Algebra

Carl Friedrich Gauß proved the Fundamental Theorem of Algebra in his doctoral thesis.

The Fundamental Theorem of Algebra. *Every $f \in \mathbb{C}[x]$ has a root in $\mathbb{C}$.*

Although it deals with an algebraic topic (the roots of univariate polynomial equations), proving it requires at least a few non-trivial results from analysis, and it can be proved without any algebraic ideas at all. This has led some to joke that the theorem is neither fundamental nor algebraic.

We will describe an algebraic proof of the Fundamental Theorem, based on ideas from Galois theory; this argument is basically found in Chapter 7 of [1]. Of course, Galois would not have made the argument we produce below. Since we need some analytical ideas first, we turn to them, without dwelling on why they are true — you can consult a text on calculus or analysis.

Background from Calculus

Every first-semester calculus student encounters the following fact.

The Intermediate Value Theorem. *Let f be a continuous function on $[a, b]$. For every y-value between $f(a)$ and $f(b)$, we can find $c \in (a, b)$ such that $f(c) = y$.*

Intuitively speaking, continuity means that f has no holes or asymptotes, so *of course* it would pass through y. However, this is not so easy to prove; the precise definition of continuity is that you can evaluate the limit at every point by substitution ($\lim_{x\to a} f(x) = f(a)$), so it takes a little more work than you would imagine at first glance. This is a class in algebra, not analysis, so we move on.

Theorem 9·68. *Polynomials over $\mathbb{C}$ are continuous.*

This one is not quite so intuitive, unless you have worked extensively with polynomials whose coefficients are complex. It is not difficult, but again, it is analytical in nature, so we move on.

Corollary 9·69. *Let $f \in \mathbb{R}[x]$. If $\deg f$ is odd, then f has a root in $\mathbb{R}$.*

This one is worth considering briefly; again, we rely on ideas from calculus.

Proof. Let $n = \deg f$, and consider

$$\lim_{x\to\infty} \frac{f(x)}{x^n} = \lim_{x\to\infty} \frac{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0}{x^n} = \lim_{x\to\infty} \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right) = a_n.$$

Let $\varepsilon > 0$. By definition, there exists $N \in \mathbb{R}$ such that for all $x > N$, $\left|a_n - \frac{f(x)}{x}\right| < \varepsilon$. Thus, for all these x, we have

$$-\varepsilon < a_n - \frac{f(x)}{x} < \varepsilon \quad \Longrightarrow x(a_n + \varepsilon) > f(x) > x(a_n - \varepsilon).$$

In other words, for all $x > N$, $f(x)$ has the same sign as a_n. A similar argument shows that we can find $M \in \mathbb{R}$ such that for all $x < M$, $f(x)$ has the same sign as $-a_n$. By definition of degree, $a_n \neq 0$, so f has at least one positive value, and at least one negative value. Apply continuity and the Intermediate Value Theorem to see that f has a root between these two points. □

Some more algebra

Now for two algebraic ideas. The first is *separability,* which has to do with how a polynomial factors in its extension field. The second is the first of the famous *Sylow Theorems.*

Definition 9·70. Suppose $\mathbb{E} = \mathbb{F}(\alpha)$ is an extension field. Let f be an irreducible polynomial over $\mathbb{F}$ such that α is a root of f. We say that α is **separable** over $\mathbb{F}$ if f factors in $\mathbb{E}$ as $(x-\alpha)\,g(x)$, and $g(\alpha) \neq 0$.

Theorem 9·71. *Extensions of $\mathbb{C}$ are separable.*

Proof. This is a consequence of Calculus. If $f = (x-a)^m \cdot g$, then $f' = m(x-a)^{m-1} g + (x-a)^m g'$. The derivative of a complex polynomial is also a complex polynomial, and the Euclidean algorithm gives us a gcd which has $p = (x-a)^{m-1}$ as a factor. If f is irreducible, the gcd of f and f' must be a constant, so $m = 1$. □

(The proof above can fail in a field of nonzero characteristic, but in this chapter we have assumed that this is not the case.)

Theorem 9·72. *Let $\mathbb{E}$ be an algebraic extension of $\mathbb{C}$. The degree of $\mathbb{E}$ over $\mathbb{C}$ is $|\mathrm{Gal}(\mathbb{E}/\mathbb{C})|$.*

Proof. We proceed by induction on $[\mathbb{E}:\mathbb{C}]$ (the degree of $\mathbb{E}$ over $\mathbb{C}$).

Inductive base: If $[\mathbb{E}:\mathbb{C}] = 1$, then $\mathbb{E} = \mathbb{C}$, so the only element of $\mathrm{Gal}(\mathbb{E}/\mathbb{C})$ is the identity. Hence $[\mathbb{E}:\mathbb{C}] = |\mathrm{Gal}(\mathbb{E}/\mathbb{C})|$.

Inductive hypothesis: Let $n \in \mathbb{N}^+$, and assume that if $[\mathbb{E}:\mathbb{C}] \leqslant n$, then $\mathrm{Gal}(\mathbb{E}/\mathbb{C}) = n$.

Inductive step: Let $f \in \mathbb{C}[x]$ such that $\mathbb{E}$ is the algebraic extension by the roots of f, and $[\mathbb{E}:\mathbb{C}] = n+1$. Let q be an irreducible factor of f, and choose g such that $f = qg$; if $\deg q = 1$, then the root of q is already in $\mathbb{C}$. Hence, we may assume without loss of generality that $\deg q > 1$. Let α be any root of q, and $\varphi \in \mathrm{Gal}(\mathbb{E}/\mathbb{C})$. In Question 9.24 you showed that $\varphi(\alpha)$ is another root of q. By Theorem 9·71, extensions of $\mathbb{C}$ are separable, so the choice of mappings for φ is determined entirely by q. Hence, $|\mathrm{Gal}(\mathbb{C}(\alpha)/\mathbb{C})| = \deg q = [\mathbb{C}(\alpha):\mathbb{C}]$. Apply the inductive hypothesis to $[\mathbb{E}:\mathbb{C}(\alpha)]$ to obtain

$$[\mathbb{E}:\mathbb{C}] = [\mathbb{E}:\mathbb{C}(\alpha)]\,[\mathbb{C}(\alpha):\mathbb{C}] = |\mathrm{Gal}(\mathbb{E}/\mathbb{C}(\alpha))| \cdot |\mathrm{Gal}(\mathbb{C}(\alpha)/\mathbb{C})|.$$

At this point we define a homomorphism that maps elements of $\mathrm{Gal}(\mathbb{E}/\mathbb{C})$ to elements of $\mathrm{Gal}(\mathbb{C}(\alpha)/\mathbb{C})$ by restriction, similar to the proof of Theorem 9·31. (The difference here is that we are not working with radical extensions, so we cannot guarantee $\mathrm{Gal}(\mathbb{C}(\alpha)/\mathbb{C})$ is abelian.) As before, the kernel will be $\mathrm{Gal}(\mathbb{E}/\mathbb{C}(\alpha))$. The kernel is a normal subgroup, so Lagrange's Theorem tells us

$$|\mathrm{Gal}(\mathbb{E}/\mathbb{C}(\alpha))| \cdot |\mathrm{Gal}(\mathbb{C}(\alpha)/\mathbb{C})| = |\mathrm{Gal}(\mathbb{E}/\mathbb{C})|,$$

which completes the proof. □

We now turn to the First Sylow Theorem, which generalizes Cauchy's Theorem that if an irreducible p divides $|G|$, then G contains an element of order p.

First Sylow Theorem. *Let G be a group, and $p \in \mathbb{N}^+$ be irreducible. If $|G| = p^m q$ where $p \nmid q$, then G has a subgroup of size p^i for each $i \in \{1, \ldots, m\}$.*

Proof. We proceed by induction on the size of G. The *inductive basis* follows from Cauchy's Theorem, so for the *inductive hypothesis,* assume that for any group H of order smaller than $|G|$, such that $|H| = p^m r$ and $p \nmid r$, we can find a subgroup A of size p^m. We need to show that we can also find a subgroup of size p^{m+1}.

Recall the class equation (9.3),

$$|G| - \sum_{x \in \mathcal{P} \setminus Z(G)} |G/C_x| = |Z(G)|.$$

We consider two cases.

Case 1: If p divides $|Z(G)|$, then Cauchy's Theorem tells us that $Z(G)$ has a normal subgroup A of size p. Elements of $Z(G)$ commute with all elements of G, so A is a normal subgroup of G. Hence, G/A is a quotient group. By Lagrange's Theorem,

$$|G/A| = \frac{|G|}{|A|} = \frac{p^m q}{p} = p^{m-1} q,$$

and since $m > 1$, p divides $|G/A|$. By hypothesis, G/A has a subgroup of size p^{m-1}. Call it B.

Recall the natural homomorphism $\mu : G \to G/A$ by $\mu(g) = gA$. This homomorphism is onto G/A, so let

$$H = \{g \in G : \mu(g) \in B\}.$$

We claim that $H < G$; to see why, let $x, y \in H$. A property of homomorphisms is that $\mu(y^{-1}) = \mu(y)^{-1} \in B$, so now closure and properties of homomorphisms guarantee that $\mu(xy^{-1}) = \mu(x)\mu(y)^{-1} \in B$.

We claim that $|H| = p^m$. Why? An argument similar to that of the Isomorphism Theorem shows that $B \cong H/\ker\mu$, so $|B| = |H| / |\ker\mu|$, and $|\ker\mu| = |A|$, so $|H| = |A||B| = p \cdot p^{m-1} = p^m$, as desired.

Case 2: Suppose $p \nmid |Z(G)|$. We claim that $p \nmid |G/C_x|$ for some $x \in G$. To see why, assume by way of contradiction that it divides all of them. By hypothesis, p divides $|G|$; p then divides the left-hand side of the class equation above, so p must divide the right hand side, $|Z(G)|$, a contradiction.

The centralizer of an element is a subgroup of G. By Lagrange's Theorem, $|G/C_x| = |G| / |C_x|$. Rewrite this as $|G/C_x||C_x| = |G|$. By hypothesis, p^m divides the right hand side, but $p \nmid |G/C_x|$, so the definition of a prime number forces $p^m \mid |C_x|$.

On the other hand, $x \notin Z(G)$, so $C_x \neq G$, so $|C_x| < |G|$. The inductive hypothesis applies, and we can find a subgroup A of C_x of size p^m. A subgroup of C_x is also a subgroup of G, so A is a desired subgroup of G whose order is p^m. □

Proof of the Fundamental Theorem

Let $f \in \mathbb{C}[x]$. Let $\mathbb{E}$ be the field that contains all the roots of f. We claim that $\mathbb{E} = \mathbb{C}$.

By unique factorization and the Factor Theorem, f can have only finitely many roots, so $\mathbb{E}$ is a finite extension of $\mathbb{C}$, itself a finite extension of $\mathbb{R}$. Hence, $\mathbb{E}$ is also a finite extension of $\mathbb{R}$. We claim that $\mathbb{E}$ is an odd-degree extension of $\mathbb{R}$; if not, we would be able to find an odd-degree polynomial $f \in \mathbb{R}[x]$ that is irreducible. By the corollary to the Intermediate Value Theorem, however, odd-degree polynomials over $\mathbb{R}$ *must* have a root in $\mathbb{R}$, a contradiction.

Hence, $\mathbb{E}$ must be an even extension of $\mathbb{R}$. If it is a degree-2 extension, then the quadratic formula suggests that $\mathbb{C} \supseteq \mathbb{E} \supseteq \mathbb{C}$, so $\mathbb{C} = \mathbb{E}$. The remaining possibilities fall into two cases: $[\mathbb{E} : \mathbb{R}] = 2^m$ (a pure power of 2) or $[\mathbb{E} : \mathbb{R}] = 2^m q$ for some odd q.

We consider the second possibility first. Suppose the degree of $\mathbb{E}$ over $\mathbb{R}$ is $2^m q$, where $m, q \in \mathbb{N}^+$ and $2 \nmid q$. Let $G = \text{Gal}(\mathbb{E}/\mathbb{R})$ be its Galois group; notice that $|G| = 2^m q$. By the First Sylow Theorem, G has a subgroup H of size 2^m. By Lagrange's Theorem, $|G/H| = q$. This corresponds to an intermediate field $\widehat{\mathbb{E}}$ such that

- the degree of $\mathbb{E}$ over $\widehat{\mathbb{E}}$ is 2^m, and
- the degree of $\widehat{\mathbb{E}}$ over $\mathbb{R}$ is q.

Since $2 \nmid q$, $\widehat{\mathbb{E}}$ is an odd-degree extension of $\mathbb{R}$, and we already dealt with that. Hence $q = 1$, and the only possibility that remains is$|G| = 2^m$, a pure power of 2.

Of course, $\mathbb{C} = \mathbb{R}\left[\sqrt{-1}\right]$ is an intermediate field between $\mathbb{E}$ and $\mathbb{R}$. Its degree over $\mathbb{R}$ is 2, so the degree of $\mathbb{E}$ over $\mathbb{C}$ is 2^{m-1}. Let f be an irreducible polynomial of degree $m-1$ over $\mathbb{C}$. We claim that $m = 1$; to see why, assume the contrary, and proceed by induction on m. If $m = 2$, then the quadratic formula shows us that the roots of $f = ax^2 + bx + c$ are

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

We claim that the square roots of complex numbers are also complex. To see why, consider $z = a + bi$, where $a, b \in \mathbb{R}$. Let $\alpha = \arctan(b/a)$ and $r = a^2 + b^2$. You will show in Question 9.73 that $z = r(\cos\alpha + i\sin\alpha)$. Let

$$w = \sqrt{r}\left(\cos\frac{\alpha}{2} + i\sin\frac{\alpha}{2}\right);$$

notice that

$$w^2 = \left(\sqrt{r}\right)^2\left[\left(\cos^2\frac{\alpha}{2} - \sin^2\frac{\alpha}{2}\right) + 2i\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\right].$$

Apply the double-angle formulas to get

$$w^2 = r\left[\cos\left(2\cdot\frac{\alpha}{2}\right) + i\sin\left(2\cdot\frac{\alpha}{2}\right)\right] = r(\cos\alpha + i\sin\alpha) = z.$$

Since z was arbitrary in $\mathbb{C}$, we see that square roots of complex numbers are also complex.

Assume, therefore, that for some $n \in \mathbb{N}^+$, if the degree of an extension field over $\mathbb{C}$ is 2^n, then the extension field *is* $\mathbb{C}$. Let $\mathbb{F}$ be an extension of $\mathbb{C}$ of degree 2^{n+1}. As before, we can construct an extension field $\widehat{\mathbb{F}}$ of $\mathbb{C}$ of degree 2^n, so that the degree of $\mathbb{F}$ over $\widehat{\mathbb{F}}$ is 2. By the inductive hypothesis, $\widehat{\mathbb{F}} = \mathbb{C}$. Hence the degree of $\mathbb{F}$ over $\mathbb{C}$ is 2, which the inductive base tells us means $\mathbb{F} = \mathbb{C}$.

By induction, then, $\mathbb{E} = \mathbb{C}$.

Question 9·73 .

Let $z \in \mathbb{C}$, and choose $a, b \in \mathbb{R}$ such that $z = a + bi$. Let $\alpha = \arctan(b/a)$ and $r = a^2 + b^2$. Show that $z = r(\cos\alpha + i\sin\alpha)$.

Chapter 10

Roots of polynomial systems

This chapter is about the roots of systems of polynomial equations, such as

$$\begin{cases} x^2 + y^3 &= 4 \\ xy &= 1 \end{cases}.$$

Rather than investigate the *computation* of roots, we consider the *analysis* of roots *we have not enumerated* explicitly, and the tools used to compute that analysis. In particular, we want to know when the roots to a multivariate system of polynomial equations exists. Techniques described here allow us to answer the following questions:

1. Does the system have any solutions in $\mathbb{C}$?
2. If so,
   (a) Are there infinitely many, or finitely many?
      i. If finitely many, exactly how many?
      ii. If infinitely many, what is the "dimension" of the solution set?
   (b) Are any of the solutions in $\mathbb{R}$?

We refer to these as ***five natural questions about the roots of a polynomial system***. We start off reviewing them for *linear* systems, but you should already have seen that in linear algebra, so we emphasize "review." We then analyze how the nature of a non-linear, multivariate *monomial* hampers this strategy with non-linear, multivariate *polynomials,* before concluding with a foray into Hilbert's Nullstellensatz and Gröbner bases, fundamental results and tools of commutative algebra and algebraic geometry.

It shouldn't surprise you that polynomial systems appear in many contexts. A chemist once emailed me about a problem he was studying that involved microarrays. Microarrays measure gene expression, and he was trying to model them using this system of equations:

$$\begin{aligned} axy - b_1x - cy + d_1 &= 0 \\ axy - b_2x - cy + d_2 &= 0 \\ axy - b_2x - b_1y + d_3 &= 0 \end{aligned} \tag{10.1}$$

where $a, b_1, b_2, c, d_1, d_2, d_3 \in \mathbb{N}$ are known constants and $x, y \in \mathbb{R}$ were unknown. The chemist wanted to find values for x and y that made all the equations true.

This already is an interesting, well-studied problem, but the chemist's fancy software didn't *always* solve the system. He didn't understand whether it was because there was something wrong with his numbers, or with the system itself. All he knew is that for some values of the coefficients, the system gave him a solution, but for other values the system found no solution. The reason turned out to be the software's reliance on *numerical techniques* to look for a solution, which can fail *even when a solution exists.*

Techniques described in this chapter showed that no *real* solution existed; all solutions were *complex*. The software's numerical techniques wasn't designed to discover such solutions, and this is why it failed.

10·1 Gaussian elimination

A generic system of m linear equations in n variables looks like

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where the a_{ij} and b_i are elements of a field $\mathbb{F}$. Although it is typically taught with $\mathbb{F} = \mathbb{R}$, linear algebra can be done over *any* field $\mathbb{F}$, such as a finite field!

Example 10·1. A linear system with $m = 3$, $n = 5$, and coefficients in $\mathbb{Z}_{13}$ is

$$\begin{aligned} 5x_1 + x_2 + 7x_5 &= 7 \\ x_3 + 11x_4 + 2x_5 &= 1 \\ 3x_1 + 7x_2 + 8x_3 &= 2. \end{aligned}$$

An equivalent system, with the same solutions, is

$$\begin{aligned} 5x_1 + x_2 + 7x_5 + 6 &= 0 \\ x_3 + 11x_4 + 2x_5 + 12 &= 0 \\ 3x_1 + 7x_2 + 8x_3 + 11 &= 0. \end{aligned}$$

Our **standard form** will typically describe a system as a list or sequence of the left-hand sides of the second form above,

$$\left\{ \begin{array}{c} 5x_1 + x_2 + 7x_5 + 6, \\ x_3 + 11x_4 + 2x_5 + 12, \\ 3x_1 + 7x_2 + 8x_3 + 11 \end{array} \right\}.$$

Gaussian elimination obtains a "triangular system" equivalent to the original. By "equivalent", we mean that $(a_1, \ldots, a_n) \in \mathbb{F}^n$ is a solution to the triangular system if and only if it is a solution to the original system as well. Algorithm 10.1 describes one way to apply the method.

Algorithm 10.1 Gaussian elimination

inputs
  $F = (f_1, f_2, \ldots, f_m)$, a list of linear polynomials in n variables, with coefficients from a field $\mathbb{F}$.
outputs
  $G = (g_1, g_2, \ldots, g_m)$, a list of linear polynomials in n variables, in triangular form, whose roots are precisely the roots of F.
do
  Let $G := F$
  for $i = 1, 2, \ldots, m-1$ **do**
    Rearrange $g_i, g_{i+1}, \ldots, g_m$ so that for each $k < \ell$, $g_\ell = 0$, or $\mathrm{lv}(g_k) \geqslant \mathrm{lv}(g_\ell)$
    if $g_i \neq 0$ **then**
      Denote the coefficient of $\mathrm{lv}(g_i)$ by a
      for $j = i+1, i+2, \ldots m$ **do**
        if $\mathrm{lv}(g_j) = \mathrm{lv}(g_i)$ **then**
          Denote the coefficient of $\mathrm{lv}(g_j)$ by b
          Replace g_j with $ag_j - bg_i$
  return G

Definition 10·2. Let $G = (g_1, g_2, \ldots, g_m)$ be a list of linear polynomials in n variables. For each $i = 1, 2, \ldots, m$ designate the **leading variable of** g_i, as the smallest-indexed variable of non-zero coefficient. Write $\mathrm{lv}(g_i)$ for this variable.

Remark. The leading variable of the zero polynomial, $\mathrm{lv}(0)$, is undefined.

The ordering for leading variables guarantees $x_1 > x_2 > \ldots > x_n$, something like a dictionary. We refer to it as the **lexicographic term ordering**. In the same way, we order $x > y > z$; if other variables appear, we state the ordering explicitly.

Example 10·3. Using the example from 10·1,

$$\mathrm{lv}(5x_1 + x_2 + 7x_5 + 6) = x_1,$$
$$\mathrm{lv}(x_3 + 11x_4 + 2x_5 + 12) = x_3.$$

Definition 10·4. A list of linear polynomials F is in **triangular form** if for each $i < j$,

- $f_i = 0$ implies $f_j = 0$, while
- $f_i, f_j \neq 0$ implies $\mathrm{lv}(f_i) > \mathrm{lv}(f_j)$.

Example 10·5. Using the example from 10·1,the list

$$F = \left(5x_1 + x_2 + 7x_5 + 6, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11 \right)$$

is not in triangular form, since $\mathrm{lv}(f_1) = \mathrm{lv}(f_3) = x_1$, whereas we want $\mathrm{lv}(f_1) > \mathrm{lv}(f_3)$.

The list

$$G = \left(x_1 + 6, \quad 0, \quad x_2 + 3x_4 \right)$$

is also not in triangular form, because g_2 is zero while $g_3 \neq 0$.

However, the list

$$H \quad = \quad \begin{pmatrix} x_1 + 6, & x_2 + 3x_4, & 0 \end{pmatrix}$$

is in triangular form, because $h_3 = 0$ and $\mathrm{lv}\,(h_1) > \mathrm{lv}\,(h_2)$.

Theorem 10·6. *Algorithm 10.1 terminates correctly.*

Proof. All the loops of the algorithm are explicitly finite, so the algorithm terminates. To show that it terminates correctly, we must show both that G is triangular and that its roots are the roots of F.

That G is triangular: We first claim that the ith iteration of the outer loop terminates with G in *i-subtriangular form*; by this we mean that

- the list $(g_1, \ldots, g_i)$ is in triangular form; *and*
- for each $j = 1, \ldots, i$ if $g_j \neq 0$ then the coefficient of $\mathrm{lv}\left(g_j\right)$ in $g_{i+1}, \ldots, g_m$ is 0.

For example, a system in 2-subtriangular form looks like this, where a "*" indicates a non-zero coefficient of a variable:

$$\begin{array}{ccccc} * & * & * & * & * \\ & * & * & * & * \\ & & * & * & * \\ & & * & * & * \\ & & * & * & * \end{array}$$

Proving this first subclaim is straightforward; after all, line 8 ensures that all the zero polynomials occur at the end of the list, and $\mathrm{lv}\,(g_i) \geqslant \mathrm{lv}\left(g_{i+j}\right)$ for any $j \geqslant 1$, while lines 13 and 14 ensure that if $g_i \neq 0$ then $\mathrm{lv}\,(g_i) > \mathrm{lv}\left(g_{i+j}\right)$ for any $j \geqslant 1$.

Having established that subclaim, we now observe that G is in triangular form if and only if G is in i-subtriangular form for all $i = 1, 2, \ldots, m$. Again, this is straightforward, and establishes that G is in triangular form after at most m iterations.

Showing that G is equivalent to F is only a little harder. The combinations of F that produce G are all linear; that is, for each $j = 1, \ldots, m$ there exist $c_{i,j} \in \mathbb{F}$ such that

$$g_j = c_{1,j} f_1 + c_{2,j} f_2 + \cdots + c_{m,j} f_m.$$

Hence if $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ is a common root of F, it is also a common root of G. For the converse, observe from the algorithm that there exists some i such that $f_i = g_1$; then there exists some $j \in \{1, \ldots, m\} \setminus \{i\}$ and some $a, b \in \mathbb{F}$ such that $f_j = a g_1 - b g_2$; and so forth. Hence the elements of F are also a linear combination of the elements of G, and a similar argument shows that the common roots of G are common roots of F. □

Remark 10·7. There are other ways to define both triangular form and Gaussian elimination. The approach we have taken assists us in the development of later ideas.

Example 10·8. We illustrate Gaussian elimination for the system of equations described in Example 10·1.

- We start with the input,

$$F \quad = \quad \left(\quad 5x_1 + x_2 + 7x_5 + 6, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11 \quad \right).$$

- Line 6 tells us to set $G = F$, so now

$$G \quad = \quad \left(\quad 5x_1 + x_2 + 7x_5 + 6, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11 \quad \right).$$

- We enter the *outer* loop on i:
  - In the first iteration, $i = 1$.
  - We rearrange G, obtaining

$$G \quad = \quad \left(\quad 5x_1 + x_2 + 7x_5 + 6, \quad 3x_1 + 7x_2 + 8x_3 + 11, \quad x_3 + 11x_4 + 2x_5 + 12 \quad \right).$$

  - Since $g_i \neq 0$, Line 10 tells us to denote a as the coefficient of $\mathrm{lv}\,(g_i)$, so $a = 5$.
  - We now enter the *inner* loop on j:
    - In the first iteration, $j = 2$.
    - As $\mathrm{lv}\left(g_j\right) = \mathrm{lv}\,(g_i)$, Line 13 tells us to denote b as the coefficient of $\mathrm{lv}\left(g_j\right)$, so $b = 3$.
    - Replace g_j with

$$\begin{aligned} ag_j - bg_i &= 5\left(3x_1 + 7x_2 + 8x_3 + 11\right) \\ &\quad - 3\left(5x_1 + x_2 + 7x_5 + 6\right) \\ &= 32x_2 + 40x_3 - 21x_5 + 37. \end{aligned}$$

      Recall that the field is $\mathbb{Z}_{13}$, so we can rewrite this as

$$6x_2 + x_3 + 5x_5 + 11.$$

      We now have

$$G \quad = \quad \left(\quad 5x_1 + x_2 + 7x_5 + 8, \quad 6x_2 + x_3 + 5x_5 + 11, \quad x_3 + 11x_4 + 2x_5 + 12 \quad \right).$$

  - We continue with the inner loop on j:
    - In the second iteration, $j = 3$.
    - Since $\mathrm{lv}\left(g_j\right) \neq \mathrm{lv}\,(g_i)$, we proceed no further.
  - Now $j = 3 = m$, and the inner loop is finished.

- We continue with the outer loop on i:
  - In the second iteration, $i = 2$.
  - We do not rearrange G, as it is already in the form indicated. (In fact, it is in triangular form already, but the algorithm does not "know" this yet.)

  - Since $g_i \neq 0$, Line 10 tells us to denote a as the coefficient of $\mathrm{lv}\,(g_i)$; since $\mathrm{lv}\,(g_i) = x_2$, $a = 6$.
  - We now enter the *inner* loop on j:
    * In the first iteration, $j = 2$.
    * Since $\mathrm{lv}\,(g_j) \neq \mathrm{lv}\,(g_i)$, we do not proceed with this iteration.
  - Now $j = 3 = m$, and the inner loop is finished.

- Now $i = 2 = m - 1$, and the outer loop is finished.
- We return G, which is in triangular form!

Once we have the triangular form of a linear system, it is easy to answer the five natural questions.

Theorem 10·9. *Let $G = (g_1, g_2, \ldots, g_m)$ is a list of nonzero linear polynomials in n variables over a field $\mathbb{F}$. If G is in triangular form, then each of the following holds.*

(A) G has common solutions if and only if none of the g_i is a nonzero constant.

(B) G has finitely many common solutions if and only if $\mathbb{F}$ has nonzero characteristic, G has common solutions, and $m = n$. In this case, there is exactly one solution.

(C) G has common solutions of dimension $d > 1$ if and only if $\mathbb{F}$ has characteristic zero, G has common solutions, and $d = n - m$.

A proof of Theorem 10·9 can be found in any textbook on linear algebra, although probably not in one place.

Example 10·10. Continuing with the system that we have used in this section, we found that a triangular form of

$$F \quad = \quad \left(\quad 5x_1 + x_2 + 7x_5 + 6, \quad x_3 + 11x_4 + 2x_5 + 12, \quad 3x_1 + 7x_2 + 8x_3 + 11 \quad \right)$$

is

$$G \quad = \quad \left(\quad 5x_1 + x_2 + 7x_5 + 6, \quad 6x_2 + x_3 + 5x_5 + 11, \quad x_3 + 11x_4 + 2x_5 + 12 \quad \right).$$

Theorem 10·9 implies that

(A) G has a solution, because none of the g_i is a constant.

(B) G has finitely many solutions, because the characteristic (13) is nonzero.

(C) If the characteristic were zero, it would have infinitely many solutions of dimension $d = n - m = 2$, as the number of polynomials ($m = 3$) is *not* the same as the number of variables ($n = 5$). (A field of characteristic zero always has infinitely many elements.)

Lexicographic order allows us to parametrize the solution set easily. Let $s, t \in \mathbb{Z}_{13}$ be arbitrary, and let $x_4 = s$ and $x_5 = t$. Back-substituting in S, we have:

- From $g_3 = 0$, $x_3 = 2s + 11t + 1$.
- From $g_2 = 0$,
$$6x_2 = 12x_3 + 8t + 12. \tag{10.2}$$
The Euclidean algorithm helps us derive the multiplicative inverse of 6 in $\mathbb{Z}_{13}$; we get 11. Multiplying both sides of (10.2) by 11, we have
$$x_2 = 2x_3 + 10t + 9.$$
Recall that we found $x_3 = 2s + 11t + 1$, so
$$x_2 = 2(2s + 11t + 1) + 10t + 9 = 4s + 6t + 11.$$
- From $g_1 = 0$,
$$5x_1 = 12x_2 + 6x_5 + 7.$$
Repeating the process that we carried out in the previous step, we find that
$$x_1 = 7s + 7.$$

We can verify this solution by substituting it into the original system:

$$\begin{aligned}
f_1: \quad & 5(7s + 7) + (4s + 6t + 11) + 7t + 6 \\
& = 39s + 13t + 52 \\
& = 0 \\
f_2: \quad & (2s + 11t + 1) + 11s + 2t + 12 \\
& = 0 \\
f_3: \quad & 3(7s + 7) + 7(4s + 6t + 11) + 8(2s + 11t + 1) + 11 \\
& = (8s + 8) + (2s + 3t + 12) + (3s + 10t + 8) + 11 \\
& = 0.
\end{aligned}$$

Before proceeding to the next section, study the proof of Theorem 10·6 carefully. Think about how we might relate these ideas to non-linear polynomials.

Question 10·11 .

A *homogeneous linear system* is one where none of the polynomials has a constant term: that is, $b_i = 0$ for $i = 1, \ldots, m$. Explain why homogeneous systems always have at least one solution.

Question 10·12 .

Find the triangular form of the following linear systems, and use it to find the common solutions of the corresponding system of equations (if any).

(a) $f_1 = 3x + 2y - z - 1$, $f_2 = 8x + 3y - 2z$, and $f_3 = 2x + z - 3$; over the field $\mathbb{Z}_7$.

(b) $f_1 = 5a + b - c + 1$, $f_2 = 3a + 2b - 1$, $f_3 = 2a - b - c + 1$; over the same field.

(c) The same system as (a), over the field $\mathbb{Q}$.

Question 10·13 .
In linear algebra you also used matrices to solve linear systems, by rewriting them in echelon (or triangular) form. Do the same with system (a) of the Question 10.12.

Question 10·14 .
Does Algorithm 10.1 also terminate correctly if the coefficients of F are not from a field, but from an integral domain? If so, and if $m = n$, can we then solve the resulting triangular system G for the roots of F as easily as if the coefficients were from a field? Why or why not?

10·2 Monomial orderings

As in the linear case, we would like to find a triangular form for *non*-linear polynomial systems. We expect that we shall have to cancel monomials. Consider the example we mentioned at the beginning of the chapter,

$$\begin{cases} x^2 + y^3 &= 4 \\ xy &= 1 \end{cases}.$$

This translates to $F = (x^2 + y^3 - 4, xy - 1)$. We will need to cancel the leading monomials of multiples of these polynomials. (We explain why later.) But, which monomial is the "leading" monomial of $f_1 = x^2 + y^3 - 4$?

- If the leading monomial is x^2, then the smallest multiples that cancel leading terms[1] give us
$$y\left(x^2 + y^3 - 4\right) - x\left(xy - 1\right) = y^4 - 4y + x.$$
- If the leading monomial is y^3, then the smallest multiples that cancel leading terms give us
$$x\left(x^2 + y^3 - 4\right) - y\left(xy - 1\right) = xy^3 - 4x + y.$$

These different results lead to different bases!

Before proceeding, we must ask ourselves how to identify the "most important" monomial in this more general setting. With linear polynomials, it was relatively easy; we picked the variable with the smallest index. You could rearrange the variables if you wanted (choose y as the leading variable, rather than x) and you'd still end up with the same basis. That doesn't work in the polynomial case; there are different options for ordering terms, and we consider them now.

Remark. We assume variables to be both prime and irreducible, so that every term of $\mathbb{X}$ has a unique factorization into variables. For instance, $xy \nmid z$ and $x^2y^3 \neq yz^2$. When you do want these relations to be true, build an ideal containing $xyq - z$ and $x^2y^3 - yz^2$.

The lexicographic ordering

Our first ordering generalizes the lexicographic ordering described in Definition 1·29 on page 13.

[1] We explain the whys and wherefores of "smallest multiples that cancel leading terms" in the next section.

Definition 10·15. Let $t, u \in \mathbb{X}$. The **lexicographic ordering** orders $t > u$ if

- $\deg_{x_1} t > \deg_{x_1} u$, or
- $\deg_{x_1} t = \deg_{x_1} u$ and $\deg_{x_2} t > \deg_{x_2} u$, or
- $\dots$
- $\deg_{x_i} t = \deg_{x_i} u$ for $i = 1, 2, \dots, n-1$ and $\deg_{x_n} t > \deg_{x_n} u$.

Another way of saying this is that $t > u$ iff there exists i such that

- $\deg_{x_j} t = \deg_{x_j} u$ for all $j = 1, 2, \dots, i-1$, and
- $\deg_{x_i} t > \deg_{x_i} u$.

The **leading monomial** of a non-zero polynomial p is any monomial t such that $t > u$ for all other terms u of p. *The leading monomial of 0 is left undefined.*

Notation 10·16. We denote the leading monomial of a polynomial p as $\operatorname{lm}(p)$.

Example 10·17. Using the lexicographic ordering over x, y,

$$\begin{aligned} \operatorname{lm}\left(x^2 + y^3 - 4\right) &= x^2 \\ \operatorname{lm}(xy - 1) &= xy \\ \operatorname{lm}\left(y^4 - 4y + x\right) &= x. \end{aligned}$$

Recall that $\mathbb{X}$ is the set of all monomials in the variables $x_1, \dots, x_n$.

Fact 10·18. *The lexicographic ordering on* $\mathbb{X}$

(A) is a linear ordering;

(B) is **compatible with divisibility***: for any $t, u \in \mathbb{X}$, if $t \mid u$, then $t \leqslant u$;*

(C) is **compatible with multiplication***: for any $t, u, v \in \mathbb{X}$, if $t < u$, then for any monomial v over $\mathbb{X}$, $tv < uv$;*

(D) orders $1 \leqslant t$ for any $t \in \mathbb{X}$; and

(E) is a well ordering.

Proof. For (A), suppose that $t \neq u$. Then there exists i such that $\deg_{x_i} t \neq \deg_{x_i} u$. Pick the smallest i for which this is true. We now have $\deg_{x_j} t = \deg_{x_j} u$ for $j = 1, 2, \dots, i-1$. If $\deg_{x_i} t < \deg_{x_i} u$, then $t < u$; otherwise, $\deg_{x_i} t > \deg_{x_i} u$, so $t > u$.

For (B), $t \mid u$ iff $\deg_{x_i} t \leqslant \deg_{x_i} u$ for all $i = 1, 2, \dots, m$. Hence $t \leqslant u$.

For (C), assume that $t < u$. Let i be such that $\deg_{x_j} t = \deg_{x_j} u$ for all $j = 1, 2, \ldots, i-1$ and $\deg_{x_i} t < \deg_{x_i} u$. For any $\forall j = 1, 2, \ldots, i-1$, we have

$$\begin{aligned}\deg_{x_j}(tv) &= \deg_{x_j} t + \deg_{x_j} v \\ &= \deg_{x_j} u + \deg_{x_j} v \\ &= \deg_{x_j} uv\end{aligned}$$

and

$$\begin{aligned}\deg_{x_i}(tv) &= \deg_{x_i} t + \deg_{x_i} v \\ &< \deg_{x_i} u + \deg_{x_i} v = \deg_{x_i} uv.\end{aligned}$$

Hence $tv < uv$.

(D) is a special case of (B).

For (E), let $M \subseteq \mathbb{X}$. We proceed by induction on the number of variables n.

For the inductive base, if $n = 1$ then the monomials are ordered according to the exponent on x_1, which is a natural number. Let E be the set of all exponents of the monomials in M; then $E \subseteq \mathbb{N}$. By the Well-Ordering Principle, E has a least element; call it e. By definition of E, e is the exponent of some monomial m of M. Since $e \leqslant \alpha$ for any other exponent $x^\alpha \in M$, m is a least element of M.

For the inductive hypothesis, assume that for all $i < n$, the set of monomials in i variables is well-ordered.

For the inductive step, let N be the set of all monomials in $n-1$ variables such that for each $t \in N$, there exists $m \in M$ such that $m = t \cdot x_n^e$ for some $e \in \mathbb{N}$. By the inductive hypothesis, N has a least element; call it t. Let

$$P = \left\{ t \cdot x_n^e : \ t \cdot x_n^e \in M \ \exists e \in \mathbb{N} \right\}.$$

All the elements of P are equal in the first $n-1$ variables: their exponents are the exponents of t. Let E be the set of all exponents of x_n for any monomial $u \in P$. As before, $E \subseteq \mathbb{N}$. Hence E has a least element; call it e. By definition of E, there exists $u \in P$ such that $u = t \cdot x_n^e$; since $e \leqslant \alpha$ for all $\alpha \in E$, u is a least element of P.

Finally, let $v \in M$. Since t is minimal in N, either there exists i such that

$$\deg_{x_j} u = \deg_{x_j} t = \deg_{x_j} v \quad \forall j = 1, \ldots, i-1$$
$$\text{and}$$
$$\deg_{x_i} u = \deg_{x_i} t < \deg_{x_i} v,$$

or

$$\deg_{x_j} u = \deg_{x_j} t = \deg_{x_j} v \quad \forall j = 1, 2, \ldots, n-1$$

In the first case, $u < v$ by definition. In the second case, e is minimal in E, and

$$\deg_{x_n} u = e \leqslant \deg_{x_n} v,$$

in which case $u \leqslant v$. Hence u is a least element of M.

Since M is arbitrary in $\mathbb{X}$, every subset of $\mathbb{X}$ has a least element. Hence $\mathbb{X}$ is well-ordered by the lexicographic order. □

Monomial diagrams

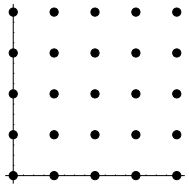
Monomial diagrams, essentially lattices, can only represent phenomena surrounding monomials in a bivariate polynomial ring $\mathbb{F}[x, y]$. We can however infer properties that hold true with an arbitrary number of variables, as well.

Definition 10·19. Let $t \in \mathbb{X}$. Its **exponent vector** $(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ satisfies $\alpha_i = \deg_{x_i} t$.

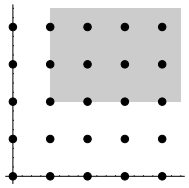
Let $t \in \mathbb{F}[x, y]$ be a monomial, and (α, β) its exponent vector. That is,

$$t = x^\alpha y^\beta.$$

If we consider (α, β) as a point in the x-y plane, the set of all monomials in two variables forms a lattice:

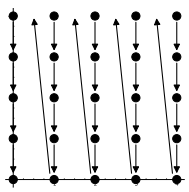


If $t \mid u$, then the point corresponding to u lies above and/or to the right of the point corresponding to t, but *never* below or to the left of it. The points corresponding to monomials divisible by xy^2 lie within the shaded region of the following diagram:



These diagrams come in handy when visualizing certain features of an ideal. For instance, we can sketch vectors on a monomial diagram that show the ordering of the monomials.

Example 10·20. In the lex ordering, the smallest monomial is 1. The next smallest is always y. Lex ordering ensures $x > y, y^2, y^3, \ldots$, so the next monomial after y is y^2, followed by y^3, etc. Once we have marked every pure power of y, the next monomial is x. Lex ordering ensures $x^2 > xy, xy^2, xy^3, \ldots$, so the next monomial after x is xy, followed by xy^2, etc. The following diagram illustrates this with each arrow pointing from one term to the next-smaller, or else to the "top" of a column of infinitely many monomials smaller than it:



This diagram illustrates an important and useful fact.

Fact 10·21. *Unless $t = x_n^a$, the lexicographic order places infinitely many monomials smaller than t.*

That's bad news in a computational sense, as it makes it impossible to guarantee from a leading monomial how many monomials a polynomial has: even if $\operatorname{lm}(f) = x$, f could have 1, 10, 100, 1000, or even more monomials. No one likes to work with polynomials that large, *not even computers!*

The graded reverse lexicographic ordering

Let's try ordering the monomials in a way that allows us to cap the size of a polynomial. For instance, given a monomial t, we might like to guarantee that all monomials have the same degree or smaller.

Definition 10·22. For a monomial t, the **total degree** of t is the sum of the exponents, denoted $\operatorname{tdeg}(t)$. For two monomials t, u, a **total-degree ordering** orders $t < u$ whenever $\operatorname{tdeg}(t) < \operatorname{tdeg}(u)$.

Example 10·23. The total degrees of x^3y^2 and xy^5 are 5 and 6, respectively, so $x^3y^2 < xy^5$. However, we cannot order x^3y^2 and x^2y^3 by total degree alone, because $\operatorname{tdeg}(x^3y^2) = \operatorname{tdeg}(x^2y^3)$ but $x^3y^2 \neq x^2y^3$.

Ties in the total degree force us to refine this approach. One way is the following.

Definition 10·24. For monomials t, u the **graded reverse lexicographic ordering**, or **grevlex**, orders $t < u$ whenever

- $\operatorname{tdeg}(t) < \operatorname{tdeg}(u)$, or
- $\operatorname{tdeg}(t) = \operatorname{tdeg}(u)$ *and* there exists $i \in \{1, \ldots, n\}$ such that for all $j = i+1, \ldots, n$
  - $\deg_{x_j} t = \deg_{x_j} u$, and
  - $\deg_{x_i} t > \deg_{x_i} u$.

To break a total-degree tie, grevlex reverses the lexicographic ordering in a double way: it searches *backwards* for the *smallest* degree, and designates the winner as the larger monomial.

Example 10·25. Under grevlex, $x^3y^2 > x^2y^3$ because the total degrees are both 5 and $y^2 < y^3$.

Question 10·26 .

Define $\pi_{\leqslant i}$ as the map from $\mathbb{X}$ to itself that "projects" a monomial in n variables to a monomial in i variables. For example,

$$\pi_{\leqslant 3}\left(x_1^5 x_2^4 x_4 x_5^2\right) = x_1^5 x_2^4.$$

Think of $\pi_{\leqslant i}$ as "chopping" variables x_{i+1}, x_{i+2}, …, x_n off the monomial. More formally, if $0 < i \leqslant n$, then

$$\pi_{\leqslant i} : \mathbb{X}[m] \to \mathbb{X}[i] \quad \text{by} \quad \pi_{\leqslant i}\left(x_1^{a_1} \cdots x_n^{a_n}\right) = x_1^{a_1} \cdots x_i^{a_i}.$$

Show that the definition of the grevlex ordering is equivalent to the following:

Definition 10·27 (Alternate definition of grevlex)**.** We say that $t < u$ if there exists i such that $\operatorname{tdeg}(\pi_{\leqslant k}(t)) = \operatorname{tdeg}(\pi_{\leqslant k}(u))$ for $k = 1, 2, \ldots, i-1$ but $\operatorname{tdeg}(\pi_{\leqslant i}(t)) < \operatorname{tdeg}(\pi_{\leqslant i}(u))$.

Fact 10·28. *The graded reverse lexicographic ordering*

(A) is a linear ordering;

(B) is compatible with divisibility;

(C) is compatible with multiplication;

(D) orders $1 \leqslant t$ *for any* $t \in \mathbb{X}$*; and*

(E) is a well ordering.

Proof. Let $t, u \in \mathbb{X}$.

Linear ordering? Assume $t \neq u$; by definition, there exists $i \in \mathbb{N}^+$ such that $\deg_{x_i} t \neq \deg_{x_i} u$. Choose the largest such i, so that $\deg_{x_j} t = \deg_{x_j} u$ for all $j = i+1, \ldots, n$. Then $t < u$ if $\deg_{x_i} t < \deg_{x_i} u$; otherwise $u < t$.

Compatible with divisibility? Assume $t \mid u$. If $t = u$, then we're done. Otherwise, $t \neq u$. We can't have $\operatorname{tdeg}(t) > \operatorname{tdeg}(u)$, as that would contradict the hypothesis that $t \mid u$! Hence $\operatorname{tdeg}(t) \leqslant \operatorname{tdeg}(u)$. If $\operatorname{tdeg}(t) < \operatorname{tdeg}(u)$, then $t < u$, and we're done. Otherwise, $t \neq u$ implies there exists $i \in \{1, \ldots, n\}$ such that $\deg_{x_i} t \neq \deg_{x_i} u$. Choose the largest such i, so that $\deg_{x_j} t = \deg_{x_j} u$ for $j = i+1, \ldots, n$. By definition, $t < u$.

Question 10·29 .
Why is it that if $t \mid u$, then $\operatorname{tdeg}(t) \leqslant \operatorname{tdeg}(u)$? Show the details that I've glossed over in the paragraph above.

Proof of Theorem 10·28 (continued). Compatible with multiplication? Assume $t < u$, and let $v \in \mathbb{X}$. By definition, $\operatorname{tdeg}(t) < \operatorname{tdeg}(u)$ or there exists $i \in \{1, 2, \ldots, n\}$ such that $\deg_{x_i} t > \deg_{x_i} u$ and $\deg_{x_j} t = \deg_{x_j} u$ for all $j = i+1, \ldots, n$. In the first case,

$$\begin{aligned} \operatorname{tdeg}(tv) &= \operatorname{tdeg}(t) + \operatorname{tdeg}(v) \\ &< \operatorname{tdeg}(u) + \operatorname{tdeg}(v) = \operatorname{tdeg}(uv). \end{aligned}$$

In the second case,

$$\deg_{x_i} tv = \deg_{x_i} t + \deg_{x_i} v > \deg_{x_i} u + \deg_{x_i} v = \deg_{x_i} uv$$

while for $j = i+1, \ldots, n$

$$\deg_{x_j} tv = \deg_{x_j} t + \deg_{x_j} v = \deg_{x_j} u + \deg_{x_j} v = \deg_{x_j} uv.$$

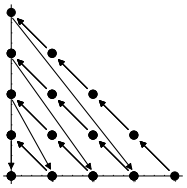
In either case, $tv < uv$ as needed.

Question 10·30 .
In the first case above, why is $\operatorname{tdeg}(tv) = \operatorname{tdeg}(t) + \operatorname{tdeg}(v)$? We skipped over that detail.

Proof of Theorem 10·28 (continued). (D) is again a special case of (B), but we can also argue from the fact that $\text{tdeg}(1) = 0$ while $\text{tdeg}(t) > 0$ for any non-constant monomial $t \in \mathbb{X}$.

We defer the proof of (E) until Fact 10·35. □

Example 10·31. Let's diagram the grevlex ordering. Again, the smallest monomial is 1, followed by y. Here's where things change; in the *grevlex* order, the monomial after y is x, *not* y^2; after all, $\text{tdeg}(x) < \text{tdeg}(y^2)$. Following x is y^2, then xy, then x^2, in that order, rounding out the degree-two monomials. We then have the degree-three monomials y^2, xy^2, x^2y, and x^3, again in that order. This leads to the following monomial diagram:



Fact 10·32. *Let $t \in \mathbb{X}$. In the grevlex order, there are finitely many monomials smaller than t.*

Question 10·33 .
Explain why Fact 10·32 is true.

Admissible orderings

Propositions 10·18 and 10·28 show that lex and grevlex share some common properties that are both convenient to the multiplication of monomials, and consistent with monomials in only one variable. We can distill these properties and identify the ones from which the others flow.

Definition 10·34. An **admissible ordering** $<$ on $\mathbb{X}$ is a linear ordering which is compatible with divisibility and multiplication.

By definition, properties (A), (B), and (C) of Proposition 10·18 hold for an admissible ordering. What of the others?

Fact 10·35. *The following properties of an admissible ordering all hold.*

(A) $1 \leqslant t$ *for all* $t \in \mathbb{X}$.

(B) The set $\mathbb{X}$ of all monomials in n variables is well-ordered by any admissible ordering. That is, every subset M of $\mathbb{X}$ has a least element.

Proof. Let $<$ be any admissible ordering. For (A), you do it!

Question 10·36 .
Show that for any admissible ordering and any $t \in \mathbb{X}$, $1 \leqslant t$.

Proof of Fact 10·35 (continued). For (B), let $M \subseteq \mathbb{X}$ and let A be the smallest absorbing subset of $\mathbb{X}$ that contains M. (Recall from Section 4·2 that the absorption property means that for any $t \in \mathbb{X}$ and any $u \in A$, $tu \in A$ also.)

We claim that A has finitely many monomials that are *not* divisible by another element of A. Why? Hark back to Ideal Nim. Let F (the *F*orbidden *F*rontier) be defined as the set of monomials *not* in A, $F = \mathbb{X} \backslash A$. If A had an infinite set S of monomials *not* divisible by other elements of A, then two player of Ideal Nim could play a game defined by F where they chose elements of S, playing for ever. *Dickson's Lemma tells us this cannot happen!*

So A has a finite set of monomials not divisible by other elements of A; call this set T. Since A is the *smallest* absorbing subset of $\mathbb{X}$ that contains M, $T \subseteq M$; otherwise, we'd have a monomial $t \in T \backslash M$ that we could remove from T, and the resulting absorbing subset would still contain M. A linear ordering can always sort finitely many elements, so the admissible ordering allows us to identify a smallest element of T; call it t. Let $u \in M$; by definition, $u \in A$, so we can find $v \in T$ such that v divides u. Since $t \leqslant v$, we use compatibility with divisibility to see that $t \leqslant v \leqslant u$. We chose u as an arbitrary element of M, so t is minimal in M. We chose M as an arbitrary subset of $\mathbb{X}$, so $\mathbb{X}$ is well-ordered by $<$. □

Question 10·37 .

The **graded lexicographic order**, which we will denote by **gralex**, orders $t < u$ if

- $\operatorname{tdeg}(t) < \operatorname{tdeg}(u)$, or
- $\operatorname{tdeg}(t) = \operatorname{tdeg}(u)$ and the lexicographic ordering would place $t < u$.

(a) Order x^2y, xy^2, and z^5 by gralex.

(b) Show that gralex is an admissible order.

(d) Sketch a monomial diagram that shows how gralex orders $\mathbb{X}$.

We conclude this section by showing two properties of admissible orderings that we need for polynomial arithmetic.

Fact 10·38. *Let $f, g \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. Each of the following holds:*

(A) $\operatorname{lm}(fg) = \operatorname{lm}(f) \cdot \operatorname{lm}(g)$

(B) $\operatorname{lm}(f \pm g) \leqslant \max(\operatorname{lm}(f), \operatorname{lm}(g))$

Proof. For convenience, write $t = \operatorname{lm}(f)$ and $u = \operatorname{lm}(g)$.

(A) Any monomial of fg can be written as the product of two monomials vw, where v is a monomial of f and w is a monomial of g. If $v \neq \operatorname{lm}(f)$, then the definition of a leading monomial implies that $v < t$. Proposition 10·18 implies that

$$vw \leqslant tw,$$

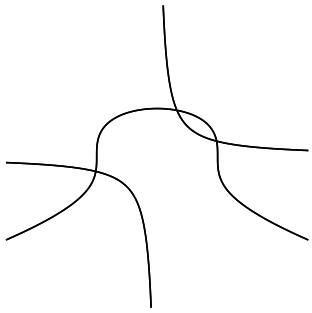


Figure 10·1: Plots of $x^2 + y^3 = 4$ and $xy = 1$

with equality only if $v = t$. The same reasoning implies that

$$vw \leqslant tw \leqslant tu,$$

with equality only if $w = u$. Hence

$$\operatorname{lm}(fg) = tu = \operatorname{lm}(f)\operatorname{lm}(g).$$

(B) Any monomial of $f \pm g$ is a monomial of f or of g. Hence $\operatorname{lm}(f \pm g)$ is a monomial of f or of g. The maximum of these is $\max(\operatorname{lm}(f), \operatorname{lm}(g))$. Hence $\operatorname{lm}(f \pm g) \leqslant \max(\operatorname{lm}(f), \operatorname{lm}(g))$. □

We typically use Fact (10·38) without explicitly referencing it, since familiarity with polynomial arithmetic typically allows one to recognize it through experience.

10·3 A triangular form for polynomial systems

Throughout this section, assume an admissible ordering of monomials.

Consider the following system of equations:

$$\begin{aligned} x^2 + y^3 &= 4 \\ xy &= 1. \end{aligned}$$

A picture can help us analyze the roots; Figure 10·1 shows the curves that correspond to these equations. Common solutions occur at the curves' intersections. We see three intersections in the real plane: two in the first quadrant, one in the fourth.

Unfortunately, the graph does not show whether complex solutions exist. (In fact, there are two.) In any case, plotting graphs for three variables is difficult; plotting more than three, effectively impossible. While it's relatively easy to solve the system above, it isn't "triangular" system in the sense that the last equation is only in one variable. So we can't solve for one variable immediately and then go backwards. We can solve for y in terms of x, but not for an exact value of y.

Question 10·39 .
Manipulate the given equations until one of them is in terms of one variable alone y. Use a computer algebra system to try to find an exact value of y.

A matrix point of view

Another way of seeing that the system isn't triangular is to consider a matrix whose rows are *degree-one* multiples of $f_1 = x^2 + y^3 - 4$ and $f_2 = xy - 1$, and whose columns are coefficients of monomials. If we order the monomials according to lex, we have the following matrix:

$$\begin{pmatrix} x^3 & x^2y & x^2 & xy^3 & xy^2 & xy & x & y^4 & y^3 & y & 1 & \\ & \mathbf{1} & & & & & \mathbf{-1} & & & & & \mathbf{xf_2} \\ & & & & & 1 & & & & \mathbf{-1} & & yf_2 \\ & & & & & & 1 & & & & -1 & f_2 \\ 1 & & & 1 & & & -4 & & & & & xf_1 \\ & \mathbf{1} & & & & & & \mathbf{1} & & \mathbf{-4} & & \mathbf{yf_1} \\ & & 1 & & & & & & 1 & & -4 & f_1 \end{pmatrix}.$$

Look at the rows labeled by yf_1 and xf_2; the leading terms' coefficients appear in the same column (x^2y). Triangularize those rows to get a new polynomial:

$$f_3 = y\left(x^2 + y^3 - 4\right) - x\left(xy - 1\right) = x + y^4 - 4y. \tag{10.3}$$

We have transformed the matrix into a new form:

$$\begin{pmatrix} x^3 & x^2y & x^2 & xy^3 & xy^2 & xy & x & y^4 & y^3 & y & 1 & \\ & & & & & & \mathbf{1} & \mathbf{1} & & \mathbf{-4} & & \mathbf{yf_1 - xf_2} \\ & & & & & 1 & & & & -1 & & yf_2 \\ & & & & & & 1 & & & & -1 & f_2 \\ 1 & & & 1 & & & -4 & & & & & xf_1 \\ & \mathbf{1} & & & & & & \mathbf{1} & & \mathbf{-4} & & \mathbf{yf_1} \\ & & 1 & & & & & & 1 & & -4 & f_1 \end{pmatrix}.$$

The leading monomials no longer cancel — in *this* matrix! With degree-one multiples of f_3, however, we find ourselves in another pickle:

$$\begin{pmatrix} x^2 & xy^4 & xy & x & y^5 & y^4 & y^3 & y^2 & y & 1 & \\ \mathbf{1} & \mathbf{1} & \mathbf{-4} & & & & & & & & \mathbf{xf_3} \\ & & \mathbf{1} & & \mathbf{1} & & & \mathbf{-4} & & & \mathbf{yf_3} \\ & & & 1 & & 1 & & & -4 & & f_3 \\ & 1 & & & & & -1 & & & & y^3f_2 \\ & & \mathbf{1} & & & & & & & \mathbf{-1} & \mathbf{f_2} \\ \mathbf{1} & & & & & & \mathbf{1} & & & \mathbf{-4} & \mathbf{f_1} \end{pmatrix}.$$

(You will see why we included $y^4 f_2$ in a moment.) This matrix has *two* cancellations. Their triangularization gives us

$$\begin{pmatrix} x^2 & xy^4 & xy & x & y^5 & y^4 & y^3 & y^2 & y & 1 & \\ & \mathbf{-1} & \mathbf{4} & & & & \mathbf{1} & & & \mathbf{-4} & \mathbf{f_1 - xf_3} \\ & & & & \mathbf{-1} & & & \mathbf{4} & & \mathbf{-1} & \mathbf{f_2 - yf_3} \\ & & & 1 & & 1 & & & -4 & & f_3 \\ & \mathbf{1} & & & & & \mathbf{-1} & & & & \mathbf{y^3 f_2} \\ & & 1 & & & & & & & -1 & f_2 \\ 1 & & & & & & 1 & & & -4 & f_1 \end{pmatrix},$$

which through cancellation in column xy^4 reduces to

$$\begin{pmatrix} x^2 & xy^4 & xy & x & y^5 & y^4 & y^3 & y^2 & y & 1 & \\ & & \mathbf{4} & & & & & & & \mathbf{-4} & \mathbf{f_1 - xf_3 + y^4 f_2} \\ & & & & \mathbf{-1} & & & \mathbf{4} & & \mathbf{-1} & \mathbf{f_2 - yf_3} \\ & & & 1 & & 1 & & & -4 & & f_3 \\ & 1 & & & & -1 & & & & & y^4 f_2 \\ & & \mathbf{1} & & & & & & & \mathbf{-1} & \mathbf{f_2} \\ 1 & & & & & & 1 & & & -4 & f_1 \end{pmatrix},$$

and through cancellation in column xy to

$$\begin{pmatrix} x^2 & xy^4 & xy & x & y^5 & y^4 & y^3 & y^2 & y & 1 & \\ & & & & & & & & & \mathbf{0} & \mathbf{f_1 - xf_3 + y^4 f_2 + 4f_2} \\ & & & & \mathbf{-1} & & & \mathbf{4} & & \mathbf{-1} & \mathbf{f_2 - yf_3} \\ & & & 1 & & 1 & & & -4 & & f_3 \\ & & 1 & & & -1 & & & & & y^4 f_2 \\ & & & 1 & & & & & & -1 & f_2 \\ 1 & & & & & & 1 & & & -4 & f_1 \end{pmatrix},$$

This gives us two new polynomials,

$$\begin{aligned} f_4 &= -y^5 + 4y^2 - 1 \\ f_5 &= 0, \end{aligned}$$

but the latter is irrelevant. However, it is apparent that cancellation will continue, since the leading monomials of yf_4, xf_4, and $x^2 f_4$ cancel with the leading monomials of multiples of f_2, f_3, and f_1, respectively. For that matter, the leading monomials of $x^2 f_2$ and xyf_1 also cancel, as do others.

Will this ever end?

An ideal point of view

Recall that all elements of a polynomial ideal share the generators' roots. The operations we perform by subtracting multiples of rows of the matrix above produce elements of the ideal, so they share those roots. Under the lex ordering, every non-constant monomial is new (x, y^5, y^4, y^2, y); none appears in the original polynomials (x^2, y^3, xy)! Contrast this to the linear case; cancelling leading variables always gives a monomial that appears in one or both of the originals! This difference is largely due to the facts that

- we cancel *variables* using *scalar multiplication*; but
- we cancel *monomials* using *monomial multiplication*.

Thus, standard Gaussian elimination won't work here, inasmuch as we need to reconsider what "triangular form" means in this case.

The primary issue to resolve is the one we observed immediately after computing the subtraction polynomial of equation (10.3): we built a polynomial f_3 whose leading term x was not divisible by the leading term of either f_1 or f_2. We built f_3 as

$$f_3 = yf_1 - xf_2;$$

by the Ideal Theorem, ideals absorb multiplication and are closed under subtraction, so

$$f_3 \in \left\langle x^2 + y^3 - 4, xy - 1 \right\rangle.$$

While f_3 is in the ideal, we wouldn't have guessed that from its leading monomial , which is *not* divisible by the leading monomials of the ideal's basis. We'd like a basis for the ideal that does not suffer from this problem.

Definition 10·40. Let G be a basis of an ideal I. We call it a **Gröbner basis of** I if for every $p \in I$, we can find $g \in G$ such that $\operatorname{lm}(g) \mid \operatorname{lm}(p)$.

It isn't obvious at the moment that a finite basis of this kind exists, let alone how we could decide whether a basis has that form. On the other hand, we can certainly conclude that

$$\left(x^2 + y^3 - 4, xy - 1\right)$$

is *not* a Gröbner basis, because $f_3 = x + y^4 - 4y$ violates the definition of a Gröbner basis, and $f_3 \in \langle x^2 + y^3 - 4, xy - 1\rangle$.

Buchberger's algorithm

How did we find f_3, f_4, and f_5? The matrices directed our attention to *subtraction polynomials,* using the smallest multiples whose leading monomials cancel. Let $t, u \in \mathbb{X}$. Write $t = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ and $u = x_1^{\beta_1} x_2^{\beta_2} \cdots x_n^{\beta_n}$. Any common multiple of t and u must have the form

$$v = x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}$$

where $\gamma_i \geqslant \alpha_i$ and $\gamma_i \geqslant \beta_i$ for each $i = 1, 2, \ldots, n$. We can thus identify a **least common multiple**

$$\operatorname{lcm}(t, u) = x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}$$

where $\gamma_i = \max(\alpha_i, \beta_i)$ for each $i = 1, 2, \ldots, n$. It really is the *least* because no common multiple can have a smaller degree in any of the variables, and so it is smallest by the definition of the lexicographic ordering.

Lemma 10·41. *Let $f, g \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, with $\operatorname{lm}(f) = t$ and $\operatorname{lm}(g) = u$. The smallest multiples of f and g whose leading terms cancel have $\operatorname{lcm}(t, u)$ as their leading terms.*

Proof. Since $\operatorname{lcm}(t,u)$ is defined to have exponents *no smaller* than those of t and u, it is straightforward to find monomials v and w such that $tv = uw = \operatorname{lcm}(t,u)$. Hence, $\operatorname{lm}(vf) = \operatorname{lm}(wg) = \operatorname{lcm}(t,u)$; we need merely multiply vf and wg by appropriate field elements to get cancellation. For instance, if $c = \operatorname{lc}(f)$ and $d = \operatorname{lc}(g)$, then $c^{-1}vf - d^{-1}wg$ does the job.

It remains to show that the *smallest* multiples that cancel have leading term $\operatorname{lcm}(t,u)$. Let vf and wg be *any* multiples such that $c^{-1}vf - d^{-1}wg$ cancels the leading terms. For each variable $x \in \{x_1, \ldots, x_n\}$,

$$\deg_x(\operatorname{lm}(vf)) = \deg_x v + \deg_x t \quad \text{and} \quad \deg_x(\operatorname{lm}(wg)) = \deg_x w + \deg_x u;$$

cancellation implies $\operatorname{lm}(vf) = \operatorname{lm}(wg)$, so $\deg_x(\operatorname{lm}(vf)) = \deg_x(\operatorname{lm}(wg))$, giving us

$$\deg_x v + \deg_x t = \deg_x w + \deg_x u.$$

Suppose $\deg_x t < \deg_x u$; the degree of a monomial is nonnegative, so

$$\deg_x v = \deg_x w + (\deg_x u - \deg_x t) \geqslant \deg_x u - \deg_x t.$$

By substitution,

$$\deg_x v + \deg_x t \geqslant (\deg_x u - \deg_x t) + \deg_x t = \deg_x u = \deg_x \operatorname{lcm}(t,u).$$

Similarly, if $\deg_x t > \deg_x u$, we would find

$$\deg_x w + \deg_x u \geqslant \deg_x t = \deg_x \operatorname{lcm}(t,u).$$

Recall that x was arbitrary in $\{x_1, \ldots, x_n\}$, so in fact $\deg_x v + \deg_x t \geqslant \deg_x \operatorname{lcm}(t,u)$ for each variable x, guaranteeing that

$$\operatorname{lcm}(t,u) \quad \text{divides} \quad vt,$$

and similarly $\operatorname{lcm}(t,u)$ divides wu. As an admissible ordering is compatible with divisibility,

$$\operatorname{lcm}(t,u) \quad \leqslant \quad vt,\ wu.$$

This shows that any multiples of f, g whose leading terms cancel have leading terms no smaller than $\operatorname{lcm}(t,u)$, as claimed. □

Definition 10·42. Let f, g be two polynomials, and $v, w \in \mathbb{X}$ such that $\operatorname{lm}(vf) = \operatorname{lm}(wg) = \operatorname{lcm}(\operatorname{lm}(f), \operatorname{lm}(g))$. Choose c, d such that $cvf - dwg$ cancels the leading terms. We call $cvf - dwg$ an ***S*-polynomial of f and g**, and write $\operatorname{spol}(f,g) = cvf - dwg$.

(The *S* stands for "subtraction.") Some *S*-polynomials occur after a row of the matrix has already had its first cancellation, as in the computation of f_4 and f_5 above. We want to distinguish those from the first cancellation, so we reserve the term *S*-polynomial for the initial cancellation in a row, and refer to subsequent cancellations as **top-reductions**.

Question 10·43 .
Let $f = x^5 + 2x^4 + 2x^2y + 3y^3 - 4x + 2$ and $g = y^6 - 4y^4 + 3y^3 - 2y^2 + y$. We use the grevlex order in this example.

(a) Show that spol (f, g) reduces to zero.

(b) Pay attention to the *quotients* you used when reducing f and g to zero. What do you notice about them?

(c) Notice that lm $(f) = x^5$ and lm $(g) = y^6$ have no common divisors aside from 1. Suppose that f and g are *any* two polynomials whose leading monomials have no common divisors. Show that spol (f, g) reduces to zero.

We now have the machinery we need to identify and compute a Gröbner basis.

Buchberger's Characterization and Buchberger's Algorithm. *Let $F = \{f_1, f_2, \ldots, f_m\} \subsetneq \mathbb{F}[x_1, x_2, \ldots, x_n]$, and $I = \langle f_1, f_2, \ldots, f_m \rangle$ the ideal generated by F.*

(A) (Buchberger's Characterization) F is a Gröbner basis of I if and only if all the S-polynomials of F reduce to zero.

(B) (Buchberger's Algorithm) If F is not a Gröbner basis of I, we can compute a Gröbner basis G, by setting $G = F$ initially, then constructing and top-reducing S-polynomials, adding to G the reduced forms of those that do not reduce to zero, until all the S-polynomials of G reduce to zero. This process takes finitely many S-polynomials.

Example 10·44. Recall

$$F = \left(x^2 + y^3 - 4, xy - 1\right).$$

We already know it is not a Gröbner basis, as its one S-polynomial is

$$\begin{aligned} S &= \text{spol}\,(f_1, f_2) \\ &= y\left(x^2 + y^3 - 4\right) - x\,(xy - 1) \\ &= x + y^4 - 4y, \end{aligned}$$

and if we adopt the lex order, lm $(S) = x$, which neither leading term of F divides.

Buchberger's Algorithm tells us to compute and top-reduce S-polynomials, adding the reduced forms to G. We start with $G = \{f_1, f_2\}$ and add f_3, f_4, as computed earlier. At this point, something interesting happens; most of the remaining S-polynomials top-reduce to zero. We already saw that spol (f_1, f_3) reduced to zero; while we do not show it, spol (f_1, f_4) also reduces to zero, as does spol (f_3, f_4). One reason we skip them is that we can actually detect this *without* computing those S-polynomials; see Question 10.43. As for spol (f_2, f_4), we get

$$\begin{aligned} \text{spol}\,(f_2, f_4) &= y^4\,(xy - 1) + x\left(-y^5 + 4y^2 - 1\right) \\ &= 4xy^2 - x - y^4. \end{aligned}$$

We first reduce this via f_2:

$$\operatorname{spol}(f_2, f_4) + 4yf_2 = -x - y^4 + 4y,$$

then via f_3:

$$\operatorname{spol}(f_2, f_4) + 4yf_2 + f_3 = 0.$$

We have reached the point where all of G's S-polynomials top-reduce to zero. Buchberger's characterization states that we have a Gröbner basis, so Buchberger's Algorithm can terminate safely.

Question 10·45 .

Show that

$$G \quad = \quad \left(\quad xy - 1, \quad x + y^3 - 4y, \quad y^4 - 4y^2 + 1 \quad \right)$$

is a Gröbner basis with respect to the lexicographic ordering.

Question 10·46 .

Show that G of Question 10.45 is *not* a Gröbner basis with respect to the grevlex ordering. The Gröbner basis property depends on the choice of term ordering!

Question 10·47 .

Show that for any non-constant polynomial f, $F = (f, f + 1)$ is not a Gröbner basis.

Question 10·48 .

Show that every list of monomials is a Gröbner basis.

It remains to prove Buchberger's Characterization and Buchberger's Algorithm. We need the following lemma, which allows us to replace polynomials that are "too large" with smaller polynomials.

Lemma 10·49. *Let $p, f_1, f_2, \ldots, f_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$. Let $F = (f_1, f_2, \ldots, f_m)$. If p top-reduces to zero with respect to F, then there exist $q_1, q_2, \ldots, q_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that each of the following holds:*

(A) $p = q_1 f_1 + q_2 f_2 + \cdots + q_m f_m$*; and*

(B) for each $k = 1, 2, \ldots, m$, $q_k = 0$ *or* $\operatorname{lm}(q_k) \operatorname{lm}(g_k) \leqslant \operatorname{lm}(p)$.

Question 10·50 .

Let

$$p = 4x^4 - 3x^3 - 3x^2y^4 + 4x^2y^2 - 16x^2 + 3xy^3 - 3xy^2 + 12x$$

and $F = (x^2 + y^2 - 4, xy - 1)$.

(a) Show that p reduces to zero with respect to F.

(b) Show that there exist $q_1, q_2 \in \mathbb{F}[x, y]$ such that $p = q_1 f_1 + q_2 f_2$.

(c) Generalize the argument of (b) to prove Lemma 10·49.

By rewriting polynomials that are "too large" as smaller polynomials, Lemma 10·49 leads us to the desired form.

Proof of Buchberger's Characterization and Buchberger's Algorithm. Assume first that F is a Gröbner basis, and let $f, g \in F$. Then

$$\operatorname{spol}(f,g) \in \langle f,g\rangle \subseteq \langle f_1, f_2, \ldots, f_m\rangle.$$

The definition of a Gröbner basis implies that there exists $k_1 \in \{1,2,\ldots,m\}$ such that f_{k_1} top-reduces $\operatorname{spol}(f,g)$ to a new polynomial, say r_1. If r_1 is not zero, then by definition we can find $k_2 \in \{1,2,\ldots,m\}$ such that f_{k_2} top-reduces r_1 to a new polynomial, say r_2. Repeating this iteratively, we obtain a chain of polynomials $r_1, r_2, \ldots$ such that r_ℓ top-reduces to $r_{\ell+1}$ for each $\ell \in \mathbb{N}$. From Proposition 10·38, we see that

$$\operatorname{lm}(r_1) > \operatorname{lm}(r_2) > \cdots.$$

Recall that the monomials are well-ordered under an admissible ordering, so any set of monomials has a least element, including the set $R = \{\operatorname{lm}(r_1), \operatorname{lm}(r_2), \ldots\}$. The chain of top-reductions cannot continue indefinitely. It cannot conclude with a non-zero polynomial r_{last}, since:

- top-reduction keeps each r_ℓ in the ideal:
  - multiplication by the absorption property, and
  - subtraction by the subring property; hence
- by definition of a Gröbner basis, a non-zero r_{last} must be top-reducible by some element of G.

It must be that $r_{\text{last}} = 0$ so $\operatorname{spol}(f_i, f_j)$ top-reduces to zero.

Now assume every S-polynomial top-reduces to zero modulo F. We want to show any element of I is top-reducible by an element of F. So let $p \in I$; by definition, there exist polynomials $h_1, \ldots, h_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that

$$p = h_1 f_1 + \cdots + h_m f_m.$$

For each i, write $t_i = \operatorname{lm}(f_i)$ and $u_i = \operatorname{lm}(h_i)$. Let $T = \max_{i=1,2,\ldots,m}(u_i t_i)$. We call T the **maximal term of the representation** $h_1, h_2, \ldots, h_m$. If $\operatorname{lm}(p) = T$, we are done, since

$$\operatorname{lm}(p) = T = u_k t_k = \operatorname{lm}(h_k)\operatorname{lm}(f_k) \quad \exists k \in \{1,2,\ldots,m\}$$

and we can top-reduce p by f_k. Otherwise, there must be some cancellation among the leading monomials of each polynomial in the sum on the right hand side. That is,

$$T = \operatorname{lm}(h_{\ell_1} f_{\ell_1}) = \operatorname{lm}(h_{\ell_2} f_{\ell_2}) = \cdots = \operatorname{lm}(h_{\ell_s} f_{\ell_s})$$

for some $\ell_1, \ell_2, \ldots, \ell_s \in \{1, 2, \ldots, m\}$. From Lemma 10·41, we know that these leading terms are multiples of $\operatorname{lcm}(t_{\ell_1}, t_{\ell_2})$, etc. That means we can rewrite the cancellations as multiples of S-polynomials,

$$\begin{aligned}\operatorname{lc}(h_{\ell_1})\operatorname{lm}(h_{\ell_1})f_{\ell_1} + \cdots + \operatorname{lc}(h_{\ell_s})\operatorname{lm}(h_{\ell_s})f_{\ell_s} = \\ = \sum_{1\leqslant a<b\leqslant s} c_{a,b}u_{a,b}\operatorname{spol}\left(f_{\ell_a}, f_{\ell_b}\right)\end{aligned}$$

where for each a, b we have $c_{a,b} \in \mathbb{F}$ and $u_{a,b} \in \mathbb{M}$. Let

$$S = \sum_{1\leqslant a<b\leqslant s} c_{a,b}u_{a,b}\operatorname{spol}\left(f_{\ell_a}, f_{\ell_b}\right).$$

Observe that

$$\left[\operatorname{lm}(h_{\ell_1})f_{\ell_1} + \operatorname{lm}(h_{\ell_2})f_{\ell_2} + \cdots + \operatorname{lm}(h_{\ell_s})f_{\ell_s}\right] - S = 0. \tag{10.4}$$

By hypothesis, each S-polynomial of S top-reduces to zero. This fact, Lemma 10·49, and Fact 10·38 imply that for each a, b we can find $q_\lambda^{(a,b)} \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that

$$\operatorname{spol}\left(f_{\ell_a}, f_{\ell_b}\right) = q_1^{(a,b)}f_1 + \cdots + q_m^{(a,b)}f_m$$

and for each $\lambda = 1, 2, \ldots, m$ we have $q_\lambda^{(a,b)} = 0$ or

$$\begin{aligned}\operatorname{lm}\left(q_\lambda^{(a,b)}\right)\operatorname{lm}(f_\lambda) &\leqslant \operatorname{lm}\left(\operatorname{spol}\left(f_{\ell_a}, f_{\ell_b}\right)\right) \\ &< \operatorname{lcm}\left(\operatorname{lm}\left(f_{\ell_a}\right), \operatorname{lm}\left(f_{\ell_b}\right)\right).\end{aligned} \tag{10.5}$$

Let $Q_1, Q_2, \ldots, Q_m \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ such that

$$Q_k = \begin{cases} \sum_{1\leqslant a<b\leqslant s} c_{a,b}u_{a,b}q_k^{(a,b)}, & k \in \{\ell_1, \ldots, \ell_s\}; \\ 0, & \text{otherwise.}\end{cases}$$

By substitution,

$$S = Q_1f_1 + Q_2f_2 + \cdots + Q_mf_m.$$

In other words,

$$S - (Q_1f_1 + Q_2f_2 + \cdots + Q_mf_m) = 0.$$

By equation (10.5) and Proposition 10·38, for each $k = 1, 2, \ldots, m$ we have $Q_k = 0$ or

$$\begin{aligned}\operatorname{lm}(Q_k)\operatorname{lm}(f_k) &\leqslant \max_{1\leqslant a<b\leqslant s}\left\{\left[u_{a,b}\operatorname{lm}\left(q_k^{(a,b)}\right)\right]\operatorname{lm}(f_k)\right\} \\ &= \max_{1\leqslant a<b\leqslant s}\left\{u_{a,b}\left[\operatorname{lm}\left(q_k^{(a,b)}\right)\operatorname{lm}(f_k)\right]\right\} \\ &\leqslant \max_{1\leqslant a<b\leqslant s}\left\{u_{a,b}\operatorname{lm}\left(\operatorname{spol}\left(f_{\ell_a}, f_{\ell_b}\right)\right)\right\} \\ &< u_{a,b}\operatorname{lcm}\left(\operatorname{lm}\left(f_{\ell_a}\right), \operatorname{lm}\left(f_{\ell_b}\right)\right) \\ &= T.\end{aligned} \tag{10.6}$$

By substitution,

$$\begin{aligned} p &= (h_1 f_1 + h_2 f_2 + \cdots + h_m f_m) - \left(S - \sum_{k \in \{\ell_1, \ldots, \ell_s\}} Q_k f_k \right) \\ &= \left[\sum_{k \notin \{\ell_1, \ldots, \ell_s\}} h_k f_k + \sum_{k \in \{\ell_1, \ldots, \ell_s\}} \left(h_k - \operatorname{lc}(h_k) \operatorname{lm}(h_k) \right) f_k \right] \\ &\quad + \cancelto{0}{\left[\sum_{k \in \{\ell_1, \ldots, \ell_s\}} \operatorname{lc}(h_k) \operatorname{lm}(h_k) f_k - S \right]} \\ &\quad + \sum_{k \in \{\ell_1, \ldots, \ell_s\}} Q_k f_k. \end{aligned}$$

Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_m \in \mathbb{F}[x_1, \ldots, x_n]$ such that

$$\mathcal{Q}_k(x) = \begin{cases} h_k, & k \notin \{\ell_1, \ldots, \ell_s\}; \\ h_k - \operatorname{lc}(h_k) \operatorname{lm}(h_k) + Q_k, & \text{otherwise.} \end{cases}$$

By substitution,

$$p = \mathcal{Q}_1 f_1 + \cdots + \mathcal{Q}_m f_m.$$

If $k \notin \{\ell_1, \ldots, \ell_s\}$, then the choice of T as the maximal term of the representation implies that

$$\operatorname{lm}(\mathcal{Q}_k) \operatorname{lm}(f_k) = \operatorname{lm}(h_k) \operatorname{lm}(f_k) < T.$$

Otherwise, Proposition 10·38 and equation (10.6) imply that

$$\begin{aligned} \operatorname{lm}(\mathcal{Q}_k) \operatorname{lm}(f_k) &\leqslant \max\left(\left(\operatorname{lm}(h_k - \operatorname{lc}(h_k) \operatorname{lm}(h_k)), \operatorname{lm}(Q_k) \right) \operatorname{lm}(f_k) \right) \\ &< \operatorname{lm}(h_k) \operatorname{lm}(f_k) \\ &= T. \end{aligned}$$

What have we done? We have rewritten the original representation of p over the ideal, which had maximal term T, with another representation, which has maximal term smaller than T. This was possible because all the S-polynomials reduced to zero; S-polynomials appeared because $T > \operatorname{lm}(p)$, implying cancellation in the representation of p over the ideal. We can repeat this as long as $T > \operatorname{lm}(p)$, generating a list of monomials

$$T_1 > T_2 > \cdots.$$

The well-ordering of $\mathbb{X}$ implies that this cannot continue indefinitely! Hence there must be a representation

$$p = H_1 f_1 + \cdots + H_m f_m$$

such that for each $k = 1, 2, \ldots, m$ $H_k = 0$ or $\operatorname{lm}(H_k) \operatorname{lm}(f_k) \leqslant \operatorname{lm}(p)$. Both sides of the equation must simplify to the same polynomial, with the same leading variable, so at least

one k has $\operatorname{lm}(H_k)\operatorname{lm}(f_k) = \operatorname{lm}(p)$; that is, $\operatorname{lm}(f_k) \mid \operatorname{lm}(p)$. Since p was arbitrary, F satisfies the definition of a Gröbner basis.

(B) If F is *not* a Gröbner basis, then Buchberger's Algorithm instructs us to continue adding the non-zero top-reductions of S-polynomials in G until all S-polynomials reduce to zero. Can this process continue indefinitely? No! To see why not, let r be an element we have just added to G. It completed top-reduction, so $\operatorname{lm}(g) \nmid \operatorname{lm}(r)$ for all $g \in G$. Remember that monomials correspond to moves in Ideal Nim, so adding r to G corresponds to a legal move of Ideal Nim: $\operatorname{lm}(r)$ is not in the set of points Gone from Gameplay! We know from Dickson's Lemma that this cannot continue indefinitely, so Buchberger's Algorithm likewise cannot continue indefinitely. The algorithm ends only if every S-polynomial reduces to zero, so the algorithm ends with a Gröbner basis, as claimed. □

Question 10·51 .
Using G of Question 10.45, compute a Gröbner basis with respect to the grevlex ordering.

Question 10·52 .
It is usually "faster" to compute a Gröbner basis in a total degree ordering than it is in the lexicographic ordering; monomial diagrams can help explain why.

(a) On a monomial diagram, shade the region containing monomials smaller than x^2y^3 with respect to lex.

(b) On a monomial diagram, shade the region containing monomials smaller than x^2y^3 with respect to grevlex.

(c) Explain how the diagram implies top-reduction of a polynomial with leading monomial x^2y^3 will *probably* take less effort with grevlex than with lex.

Question 10·53 .
For G to be a Gröbner basis, Definition 10·40 requires that every polynomial in the ideal generated by G be top-reducible by some element of G. If polynomials in the basis are top-reducible by other polynomials in the basis, we call them **redundant elements of the basis**.

(a) The Gröbner basis of Question 10.45 has redundant elements. Find a subset G_{min} of G that contains no redundant elements, but is still a Gröbner basis.

(b) Describe the method you used to find G_{min}.

(c) Explain why redundant polynomials are not required to satisfy Definition 10·40. That is, if we know that G is a Gröbner basis, then we could remove redundant elements to obtain a smaller list, G_{min}, which is also a Gröbner basis *of the same ideal*.

Definition 10·54. We call the basis obtained by the process you describe in Question 10.53 a **minimal Gröbner basis** of the ideal.

10·4 Nullstellensatz

In order to apply our triangular form to the solution of zeros, we need a theorem of Hilbert. The theorem goes by its German name, *Nullstellensatz,* which translates roughly as "Theorem (*satz*) on the locations (*stellen*) of zero (*null*)." There are two versions, a *weak* Nullstellensatz, and a "*not-so-weak*" Nullstellensatz. We consider only the weak version. Throughout this section,

- $\mathbb{F}$ is an *algebraically closed* field—that is, the roots of every nonconstant polynomial over $\mathbb{F}$ appear in $\mathbb{F}$;
- $\mathcal{R} = \mathbb{F}[x_1, x_2, \ldots, x_n]$ is a polynomial ring;
- $F \subseteq \mathcal{R}$;
- $V_F \subseteq \mathbb{F}^n$ is the set of common roots of elements of F;[2] and
- $I = \langle F \rangle$.

For example, $\mathbb{C}$ is algebraically closed by the Fundamental Theorem of Algebra, but $\mathbb{R}$ is not, since the roots of $x^2 + 1 \in \mathbb{R}[x]$ are not in $\mathbb{R}$. An interesting and useful consequence of algebraic closure is the following.

Lemma 10·55. *$\mathbb{F}$ is infinite.*

Proof. Let $n \in \mathbb{N}^+$, and $a_1, \ldots, a_n \in \mathbb{F}$. Let $A = \{a_1, \ldots, a_n\} \subseteq \mathbb{F}$ be any list of elements of $\mathbb{F}$. Let $f = (x - a_1) \cdots (x - a_n)$; this is a polynomial in $\mathbb{F}[x_1, \ldots, x_n]$. While it is not the zero polynomial, it is equal to zero at every point used to build f.

Now let $g = f + 1$; it, too, is a polynomial of $\mathbb{F}[x_1, \ldots, x_n]$. However, g has *no* common roots with f, since

$$g(a) = f(a) + 1 = 0 + 1 = 1 \quad \text{for all } a \in \mathbb{F}.$$

Because $\mathbb{F}$ is algebraically closed, we can find a root $b \in \mathbb{F} \backslash A$ of g. In other words, every finite set A of elements of $\mathbb{F}$ lacks at least one element b of f, showing that no finite subset enumerates $\mathbb{F}$, which must be infinite. □

Hilbert's Weak Nullstellensatz. *If $V_F = \varnothing$, then $I = \mathcal{R}$.*

Proof. We proceed by induction on n, the number of variables.

Inductive base: Let $n = 1$. In this case, $\mathcal{R} = \mathbb{F}[x]$. By Theorem 4·53, $\mathcal{R}$ is a principal ideal domain. Thus $I = \langle f \rangle$ for some $f \in \mathcal{R}$. If $V_F = \varnothing$, then f has no roots in $\mathbb{F}$. Theorem 6·42 tells us that every principal ideal domain is a unique factorization domain, so if f is nonconstant, it has a unique factorization into irreducible polynomials. Fact 3·69 tells us that any irreducible factor p of f transforms $\mathcal{R}$ to a field $\mathbb{E} = \mathcal{R}/\langle p \rangle$ containing both $\mathbb{F}$ and a root α of p. Since $\mathbb{F}$ is algebraically closed, $\alpha \in \mathbb{F}$ itself; that is, $\mathbb{E} \cong \mathbb{F}$. But then $\alpha \in \mathbb{F}$, which means both p and, therefore, f have a root in $\mathbb{F}$, contradicting the hypothesis that $V_F = \varnothing$. Our only questionable assumption was that p is an irreducible factor of f; we conclude that f

[2] The notation V_F comes from the term **variety** in algebraic geometry.

has no irreducible factors, which (since we are in a unique factorization domain) means that f is a nonzero constant; that is, $f \in \mathbb{F}$. By the inverse property of fields, $f^{-1} \in \mathbb{F} \subseteq \mathbb{F}[x]$, and absorption implies that $1 = f \cdot f^{-1} \in I$.

Inductive hypothesis: Let $k \in \mathbb{N}^+$, and suppose that in any polynomial ring over a closed field with k variables, $V_F = \varnothing$ implies $I = \mathcal{R}$.

Inductive step: Let $n = k+1$. Assume $V_F = \varnothing$. If F contains a constant polynomial, then we are done; thus, let $f \in F$. Let d be the maximum degree of a term of f. Rewrite f by substituting

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= y_2 + a_2 y_1, \\ &\vdots \\ x_n &= y_n + a_n y_1, \end{aligned}$$

with $a_1, \ldots, a_n \in \mathbb{F}$ specificied below. This can be a little confusing, so let's take an example. □

Example 10·56. Suppose $f = x_1 + x_2^2 x_3$. We rewrite f as

$$\begin{aligned} y_1 + \left(y_2 + a_2 y_1\right)^2 \left(y_3 + a_3 y_1\right)^3 = \\ y_1 + \left(y_2^2 + 2a_2 y_1 y_2 + a_2^2 y_1^2\right)\left(y_3^3 + 3a_3 y_1 y_3^2 + 3a_3^2 y_1^2 y_3 + a_3^3 y_1^3\right). \end{aligned}$$

Take note of the forms within the parentheses.

Proof of the Weak Nullstellensatz, continued. Observe that if $i \neq 1$, then we rewrite x_i^d as $y_i^d + a_2 y_1 y_i^{d-1} \cdots + a_i^d y_1^d$, so if both $1 < i < j$ and $b + c = d$, then

$$\begin{aligned} x_i^b x_j^c &= \left(y_i^b + \cdots + a_i^b y_1^b\right)\left(y_j^c + \cdots + a_j^c y_1^c\right) \\ &= a_i^b a_j^c y_1^{b+c} + g\left(y_1, y_i, y_j\right) \\ &= a_i^b a_j^c y_1^d + g\left(y_1, y_i, y_j\right), \end{aligned}$$

where $\deg_{y_1} g < d$. Thus, we can collect the terms of f as

$$f = c y_1^d + g\left(y_1, \ldots, y_n\right)$$

where $c \in \mathbb{F}$, d is the maximal degree of y_1, and $\deg_{y_1} g < d$. Since $\mathbb{F}$ is infinite, we can find $a_2, \ldots, a_n$ such that $c \neq 0$.

Let $\varphi : \mathcal{R} \longrightarrow \mathbb{F}[y_1, \ldots, y_n]$ by

$$\varphi\left(f\left(x_1, \ldots, x_n\right)\right) = f\left(y_1, y_2 + a_2 y_1, \ldots, y_n + a_n y_1\right);$$

that is, φ substitutes *every* element of $\mathcal{R}$ with the values that we obtained so that f_1 would have the special form above. This is a ring isomorphism (Question 10.59), so $J = \varphi(I)$ is an ideal of $\mathbb{F}[y_1, \ldots, y_n]$. If $V_J \neq \varnothing$, then any $b \in V_J$ can be transformed into an element of V_F (see Question 10.60); hence $V_J = \varnothing$ as well.

Now let $\eta : \mathbb{F}[y_1, \ldots, y_n] \longrightarrow \mathbb{F}[y_2, \ldots, y_n]$ by $\eta(g) = g(0, y_2, \ldots, y_n)$. □

Example 10·57. For instance, $\eta\left(x_1^3 + x_1x_3^2 + x_2^2x_3 + x_4\right) = x_2^2x_3 + x_4$.

Proof of the Weak Nullstellensatz, continued. Again, $K = \eta(J)$ is an ideal, though the proof is different (Question 10.62). We claim that if $V_K \neq \varnothing$, then likewise $V_J \neq \varnothing$. To see why, let $h \in \eta\left(\mathbb{F}\left[y_1, \ldots, y_n\right]\right)$, and suppose $b \in \mathbb{F}^{n-1}$ satisfies $h(b) = 0$. Let g be any element of $\mathbb{F}\left[y_1, \ldots, y_n\right]$ such that $\eta(g) = h$; then

$$g\left(0, b_1, \ldots, b_{n-1}\right) = h\left(b_1, \ldots, b_{n-1}\right) = 0,$$

so that we can prepend 0 to any element of V_K and obtain an element of V_J. Since $V_J = \varnothing$, this is impossible, so $V_K = \varnothing$.

Since $V_K = \varnothing$ and $K \subseteq \mathbb{F}\left[y_2, \ldots, y_n\right]$, the inductive hypothesis finally helps us see that $K = \mathbb{F}\left[y_2, \ldots, y_n\right]$. In other words, $1 \in K$. Since $K \subseteq J$ (see Question 10.62), $1 \in J$. Since $\varphi(f) \in \mathbb{F}$ if and only if $f \in \mathbb{F}$ (Question 10.61), there exists some $f \in \langle F \rangle$ such that $f \in \mathbb{F}$. □

Question 10·58 .
Show that the intersection of two radical ideals is also radical.

Question 10·59 .
Show that φ in the proof of Hilbert's Weak Nullstellensatz is a ring isomorphism.

Question 10·60 .
Show that in the proof of Hilbert's Weak Nullstellensatz, any $b \in V_{\varphi(F)}$ can be rewritten to obtain an element of V_F. *Hint:* Reverse the translation that defines φ.

Question 10·61 .
Show that in the proof of Hilbert's Weak Nullstellensatz, $\varphi(f) \in \mathbb{F}$ if and only if $f \in \mathbb{F}$.

Question 10·62 .
Show that if J is an ideal of $\mathbb{F}\left[y_1, \ldots, y_n\right]$, then η in the proof of Hilbert's Weak Nullstellensatz maps J to an ideal $\eta(J)$ of $\mathbb{F}\left[y_2, \ldots, y_n\right]$. *Hint:* $\mathbb{F}\left[y_2, \ldots, y_n\right] \subsetneq \mathbb{F}\left[y_1, \ldots, y_n\right]$ and $\eta(J) = J \cap \mathbb{F}\left[y_2, \ldots, y_n\right]$ is an ideal of $\mathbb{F}\left[y_2, \ldots, y_n\right]$.

10·5 Elementary applications of Gröbner bases

We turn our attention to posing, and answering, questions that make Gröbner bases interesting. As in Section 10·4,

- $\mathbb{F}$ is an *algebraically closed* field—that is, all polynomials over $\mathbb{F}$ have their roots in $\mathbb{F}$;
- $\mathcal{R} = \mathbb{F}\left[x_1, x_2, \ldots, x_n\right]$ is a polynomial ring;
- $F \subseteq \mathcal{R}$;
- $V_F \subseteq \mathbb{F}^n$ is the set of common roots of elements of F;

- $I = \langle F \rangle$; and
- $G = (g_1, g_2, \ldots, g_m)$ is a Gröbner basis of I with respect to an admissible ordering.

Note that $\mathbb{C}$ is algebraically closed, but $\mathbb{R}$ is not, since the roots of $x^2 + 1 \in \mathbb{R}[x]$ are not in $\mathbb{R}$.

A Gröbner basis of an ideal

Our first questions regards the relationship of a Gröbner basis to its ideal.

Theorem 10·63 (The Ideal Membership Problem)**.** *Let $p \in \mathcal{R}$. The following are equivalent:*

(A) $p \in I$, and

(B) p top-reduces to zero with respect to G.

Proof. That (A) $\Rightarrow$ (B): Assume that $p \in I$. If $p = 0$, then we are done. Otherwise, the definition of a Gröbner basis implies that $\mathrm{lm}(p)$ is top-reducible by some element of G; let r be the result of this top-reduction. By Fact 10·38, $\mathrm{lm}(r_1) < \mathrm{lm}(p)$. By the definition of an ideal, $r_1 \in I$. If $r_1 = 0$, then we are done; otherwise the definition of a Gröbner basis implies that $\mathrm{lm}(p)$ is top-reducible by some element of G. Continuing as above, we generate a list of polynomials $p, r_1, r_2, \ldots$ such that

$$\mathrm{lm}(p) > \mathrm{lm}(r_1) > \mathrm{lm}(r_2) > \cdots.$$

By the well-ordering of $\mathbb{X}$, this list cannot continue indefinitely, so eventually top-reduction must be impossible. As long as $r_i \neq 0$, we can continue this indefinitely, so the chain must terminate with $r_i = 0$.

That (B) $\Rightarrow$ (A): Assume that p top-reduces to zero with respect to $G = \{g_1, \ldots, g_m\}$. By Lemma 10·49, we can find $q_1, \ldots, q_m$ such that $p = q_1 g_1 + \cdots + q_m g_m$. A Gröbner basis is a subset of its ideal, so $g \in I$ for each $g \in G$. By absorption, $q_i g_i \in I$ for $i = 1, \ldots, m$. By closure of addition in subgroups, $p = q_1 g_1 + \cdots + q_m g_m \in I$, as claimed. □

Up to this point, we've considered a Gröbner basis to be a basis of an ideal in the mere sense of divisibility of leading monomials. Is it also a basis in the sense of Definition 4·47; that is, can we write every element of I in terms of the Gröbner basis?

Question 10·64 .

Why do Question 10·49 and Theorem 10·63 show that a Gröbner basis of I is a basis of I in the traditional sense? That is, for every element $f \in I$ we can find $q_1, q_2, \ldots, q_m \in \mathbb{F}[x_1, \ldots, x_m]$ such that $f = q_1 g_1 + q_2 g_2 + \cdots + q_m g_m$?

A Gröbner basis and a variety

In Question 4.59 you showed that

> …the common roots of $f_1, f_2, \ldots, f_m$ are common roots of all polynomials in the ideal I.

In Question 10.64, you showed that $I = \langle G \rangle$. So the common roots of $g_1, g_2, \ldots, g_m$ are common roots of all polynomials in I. Similarly, the common roots of $f_1, f_2, \ldots, f_m$ are common roots of $g_1, g_2, \ldots, g_m$. So we can analyze the roots of a polynomial system F by analyzing the roots of any Gröbner basis G of $\langle F \rangle$. This might seem unremarkable, except that like triangular linear systems, *it is easy to analyze the roots of Gröbner bases!* Our next result gives an easy test for the existence of common roots.

Theorem 10·65. *The following both hold.*

(A) $V_F = V_G$; that is, common roots of F are common roots of G, and vice versa.

(B) F has no common roots if and only if G contains a nonzero constant polynomial.

Proof. (A) Let $\alpha \in V_F$. By definition, $f_i(\alpha_1, \ldots, \alpha_n) = 0$ for each $i = 1, \ldots, m$. By construction, $G \subseteq \langle F \rangle$, so $g \in G$ implies that $g = h_1 f_1 + \cdots + h_m f_m$ for certain $h_1, \ldots, h_m \in \mathcal{R}$. By substitution,

$$\begin{aligned} g(\alpha_1, \ldots, \alpha_n) &= \sum_{i=1}^{m} h_i(\alpha_1, \ldots, \alpha_n) f_i(\alpha_1, \ldots, \alpha_n) \\ &= \sum_{i=1}^{m} h_i(\alpha_1, \ldots, \alpha_n) \cdot 0 \\ &= 0. \end{aligned}$$

That is, α is also a common root of G. In other words, $V_F \subseteq V_G$.

On the other hand, $F \subseteq \langle F \rangle = \langle G \rangle$ by Question 10.64, so a similar argument shows that $V_F \supseteq V_G$. We conclude that $V_F = V_G$.

(B) From (A), F has common roots if and only if G has common roots. If G contains a nonzero constant polynomial g, then no element of $\mathbb{F}$ is a root of g, so $V_G = \varnothing$, and we conflude that $V_F = \varnothing$; or, F has no common roots.

For the converse, we need the Weak Nullstellensatz. If F has no common roots, then $V_F = \varnothing$, and by the Weak Nullstellensatz, $I = \mathcal{R}$. By Question 4.34, $1 \in I$. By definition of a Gröbner basis, there is some $g \in G$ such that $\operatorname{lm}(g) \mid \operatorname{lm}(1)$. This is possible only if g is a constant. □

Once we know common solutions exist, we want to know how many there are.

Theorem 10·66. *There are finitely many complex solutions if and only if for each $i = 1, \ldots, n$ we can find $g \in G$ and $a \in \mathbb{N}$ such that $\operatorname{lm}(g) = x_i^a$.*

Remark 10·67. Theorem 10·66 is related to the strong Nullstellensatz.

Proof. We can find $g \in G$ and $\alpha \in \mathbb{N}$ such that $\operatorname{lm}(g) = x_i^a$ for each $i = 1, 2, \ldots, n$ if and only if $\mathcal{R}/I$ is finite; see Figure 10·2. So the trick is to show a relationship between the residues of $\mathcal{R}/I$ and the common roots of I. The definition $\mathcal{R}/I$ is independent of any monomial ordering, so we can assume the ordering is lexicographic without loss of generality.

Assume first that for each $i = 1, \ldots, n$ we can find $g \in G$ and $a \in \mathbb{N}$ such that $\operatorname{lm}(g) = x_i^a$. Since x_n is the smallest variable, even $x_{n-1} > x_n$, so g must be a polynomial in x_n *alone;* any other variable in a non-leading monomial would contradict the assumption that $\operatorname{lm}(g) = x_n^a$. The Fundamental Theorem of Algebra implies that g has a complex solution. We can back-substitute these solutions into the remaining polynomials, using similar logic. Each back-substitution yields only finitely many solutions. There are finitely many polynomials, so G has finitely many complex solutions.

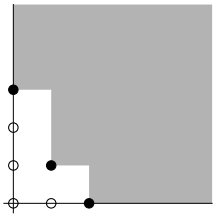


Figure 10·2: This monomial diagram shades the monomials divisible by the leading monomials of a Gröbner basis of I. If $\mathcal{R}/I$ is finite, then we cannot find infinitely many polynomials in $\mathcal{R}$ and outside I. This includes the axes of the monomial diagram, which consist of the monomials $x, x^2, x^3, \ldots$ and $y, y^2, y^3, \ldots$. They *must* reduce into a finite $\mathcal{R}/I$, so the Gröbner basis must have polynomials whose leading monomials divide them: in this case, x^2 and y^3.

Conversely, assume G has finitely many solutions; call them $\alpha^{(1)}, \ldots, \alpha^{(\ell)} \in \mathbb{F}^n$. Let

$$J = \left\langle x_1 - \alpha_1^{(1)}, \ldots, x_n - \alpha_n^{(1)} \right\rangle \bigcap \cdots \bigcap \left\langle x_1 - \alpha_1^{(\ell)}, \ldots, x_n - \alpha_n^{(\ell)} \right\rangle.$$

By Lemma 4·46, J is an ideal. The roots of I and J are related:

Question 10·68 .
Suppose A, B are ideals of $\mathcal{R}$.

(a) Show that $V_{A \cap B} = V(A) \cup V(B)$; that is, the variety of an intersection of ideals is the union of the ideals' varieties.

(b) Explain why this shows that for I and J defined above, $V_I = V_J$.

Proof of Theorem 10·66 (continued). Recall from Question 4.60 the radical of an ideal. Recall from Question 3.14 that $\mathcal{R}$ has no zero divisors, so for any $f \in \sqrt{I}$,

$$f(\alpha) = 0 \qquad \Longleftrightarrow \qquad f^a(\alpha) = 0 \ \exists a \in \mathbb{N}^+.$$

The roots of an ideal and its radical are thus identical; $V_I = V_{\sqrt{I}}$.

Let K be the ideal of polynomials that vanish on V_I. By definition, $I \subseteq \sqrt{I} \subseteq K$. We claim that $\sqrt{I} \supseteq K$ as well. Why? Let $p \in K$ be a nonzero polynomial. Consider the polynomial ring $\mathbb{F}[x_1, \ldots, x_n, y]$ where y is a new variable. Let $A = \langle f_1, \ldots, f_m, 1 - yp \rangle$. Does A have any common roots? We claim that it does not; to see why not, let $a = (a_1, \ldots, a_n, a_y)$ be a hypothetical common root. By definition, $f_1(a) = \cdots = f_m(a) = 0$, so that a is a common root of I, or $a \in V_I$. Elements of K vanish at a, so $p(a) = 0$, so by substitution into $1 - yp$, we have $1 - a_y p(a) = 1 - a_y \cdot 0 = 1$. This contradicts the hypothesis that a is a common root of the f's *and* of $1 - yp$, so A has no common roots.

By the Weak Nullstellensatz, any Gröbner basis of A has a nonconstant polynomial, call it c. By definition of A, there exist $H_1, \ldots, H_{m+1} \in \mathbb{F}[x_1, \ldots, x_n, y]$ such that

$$c = H_1 f_1 + \cdots + H_m f_m + H_{m+1}(1 - yp).$$

Let $h_i = c^{-1} H_i$ and

$$1 = h_1 f_1 + \cdots + h_m f_m + h_{m+1}(1 - yp).$$

Put $y = \frac{1}{p}$ and we have

$$1 = h_1 f_1 + \cdots + h_m f_m + h_{m+1} \cdot 0$$

where each h_i is now in terms of x_1, ..., x_n and $1/p$. Clear the denominators by multiplying both sides by a suitable power a of p, and we have

$$p^a = h_1' f_1 + \cdots + h_m' f_m$$

where each $h_i' \in \mathcal{R}$. Since $I = \langle f_1, \ldots, f_m \rangle$, we see that $p^a \in I$. Thus $p \in \sqrt{I}$. Since p was arbitrary in K, we have $\sqrt{I} \supseteq K$, as claimed.

We have shown[3] that $K = \sqrt{I}$. Since K is the ideal of polynomials that vanish on V_I, $V_K = V_I$; by substitution, $V_{\sqrt{I}} = V_I$; by Question 10.68, we can substitute to $V_{\sqrt{I}} = V_J$. In fact, $V_{\sqrt{I}} = V_{\sqrt{J}}$.

Question 10·69 .
Why can we claim that $V_{\sqrt{I}} = V_{\sqrt{J}}$? *Hint:* If you can show that $J = \sqrt{J}$, it's a matter of substitution.

Proof of Theorem 10·66 (continued). If two radical ideals have identical varieties, then the radical ideals themselves must be identical. Hence $\sqrt{I} = \sqrt{J} = J$. Define $q_j = \prod_{i=1}^{\ell} \left(x_j - a_j^{(i)} \right)$ for $j = 1, \ldots, n$. By definition of J, each $q_j \in J$. Since $\sqrt{I} = J$, suitable choices of $a_1, \ldots, a_n \in \mathbb{N}^+$ give us

$$q_1 = \prod_{i=1}^{\ell} \left(x_1 - \alpha_1^{(i)} \right)^{a_1}, \ldots, q_n = \prod_{i=1}^{\ell} \left(x_n - \alpha_n^{(i)} \right)^{a_n} \in I.$$

Notice that $\operatorname{lm}(q_i) = x_i^{a_i}$ for each i. Since G is a Gröbner basis of I, the definition of a Gröbner basis implies that for each i there exists $g \in G$ such that $\operatorname{lm}(g) \mid \operatorname{lm}(q_i)$. In other words, for each i there exists $g \in G$ and $a \in \mathbb{N}$ such that $\operatorname{lm}(g) = x_i^a$. □

Example 10·70. Recall the system from Example 10·44,

$$F = \left(x^2 + y^2 - 4, xy - 1 \right).$$

In Question 10.45 you computed a Gröbner basis in the lexicographic ordering. You probably obtained a superset of

$$G = \left(x + y^3 - 4y, y^4 - 4y^2 + 1 \right).$$

G is also a Gröbner basis of $\langle F \rangle$. Since G contains no constants, we know that F has common roots. Since $x = \operatorname{lm}(g_1)$ and $y^4 = \operatorname{lm}(g_2)$, we know that there are finitely many common roots.

[3] This, incidentally, is the "full" Nullstellensatz: any f whose roots are the common roots of I appears in $\sqrt{I}$.

We conclude by pointing in the direction of how to find the common roots of a system.

The Elimination Theorem. *Suppose the ordering is lexicographic with $x_1 > x_2 > \cdots > x_n$. For all $i = 1, 2, \ldots, n$, each of the following holds.*

(A) $\widehat{I} = I \cap \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$ *is an ideal of* $\mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$. *(If $i = n$, then $\widehat{I} = I \cap \mathbb{F}$.)*

(B) $\widehat{G} = G \cap \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$ *is a Gröbner basis of the ideal $\widehat{I}$.*

Proof. For (A), let $f, g \in \widehat{I}$ and $h \in \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$. Now $f, g \in I$ as well, we know that $f - g \in I$, and subtraction does not add any terms with factors from $x_1, \ldots, x_{i-1}$, so $f - g \in \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$ as well. By definition of $\widehat{I}$, $f - g \in \widehat{I}$. Similarly, $h \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ as well, so $fh \in I$, and multiplication does not add any terms with factors from $x_1, \ldots, x_{i-1}$, so $fh \in \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$ as well. By definition of $\widehat{I}$, $fh \in \widehat{I}$.

For (B), let $p \in \widehat{I}$. Again, $p \in I$, so there exists $g \in G$ such that $\operatorname{lm}(g)$ divides $\operatorname{lm}(p)$. The ordering is lexicographic, so g cannot have *any* terms with factors from $x_1, \ldots, x_{i-1}$. Thus $g \in \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$. By definition of $\widehat{G}$, $g \in \widehat{G}$. Thus $\widehat{G}$ satisfies the definition of a Gröbner basis of $\widehat{I}$. □

The ideal $\widehat{I}$ is important enough to merit its own terminology.

Definition 10·71. For $i = 1, 2, \ldots, n$ the ideal $\widehat{I} = I \cap \mathbb{F}[x_i, x_{i+1}, \ldots, x_n]$ is called the ***i*th elimination ideal of** I.

The Elimination Theorem suggests that we can find the common roots of F by computing a Gröbner basis G of F with respect to the lexicographic ordering, then:

- find common roots of $G \cap \mathbb{F}[x_n]$;
- back-substitute to find common roots of $G \cap \mathbb{F}[x_{n-1}, x_n]$;
- …
- back-substitute to find common roots of $G \cap \mathbb{F}[x_1, x_2, \ldots, x_n]$.

This is *exactly* how Gaussian elimination worked: reducing a matrix to row-echelon form gives us a polynomial in the bottom row whose solutions we can calculate easily, then back-substitute into previous rows.

Example 10·72. We can find the common solutions of the circle and the hyperbola in Figure 10·1 on page 343 using the Gröbner basis computed in Example 10·70. Since

$$G = \left(x + y^3 - 4y, y^4 - 4y^2 + 1\right),$$

we have

$$\widehat{G} = G \cap \mathbb{C}[y] = \left\{y^4 - 4y^2 + 1\right\}.$$

It isn't hard to find the roots of this polynomial. Let $u = y^2$; the resulting substitution gives us the quadratic equation $u^2 - 4u + 1$ whose roots are

$$u = \frac{4 \pm \sqrt{(-4)^2 - 4 \cdot 1 \cdot 1}}{2} = 2 \pm \sqrt{3}.$$

Back-substituting u into $\widehat{G}$,

$$y = \pm\sqrt{u} = \pm\sqrt{2 \pm \sqrt{3}}.$$

We can now back-substitute y into G to find that

$$\begin{aligned} x &= -y^3 + 4y \\ &= \mp\left(\sqrt{2 \pm \sqrt{3}}\right)^3 \pm 4\sqrt{2 \pm \sqrt{3}}. \end{aligned}$$

Thus there are four common roots, all of them real, illustrated by the four intersections of the circle and the hyperbola.

Question 10·73 .

Determine whether $x^6 + x^4 + 5y - 2x + 3xy^2 + xy + 1$ is an element of the ideal $\langle x^2 + 1, xy + 1\rangle$.

Question 10·74 .

How many solutions does this system have?

$$w + x + y + z, \quad wx + xy + yz + zw, \quad wxy + xyz + yzw + zwx, \quad wxyz - 1.$$

If infinitely many, what is the dimension? (This system is commonly known as the **Cyclic-4** system.)

Question 10·75 .

Consider the system

$$F = \left(\begin{array}{l} xyz + xz + 3y + 3, \\ x^2yz^2 + x^2z^2 - y - 1 \end{array} \right).$$

Bibliography

[1] Benjamin Fine and Gerhard Rosenberger. *The Fundamental Theorem of Algebra*. Undergraduate Texts in Mathematics. Springer-Verlag New York, Inc., 1997.

[2] Niels Lauritzen. *Concrete Abstract Algebra: From Numbers to Gröbner Bases*. Cambridge University Press, Cambridge, 2003.

[3] Joachim von zur Gathen and Jurgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, Cambridge, 1999.